Exact Correct–Decoding Exponent of the Wiretap Channel Decoder [*]

Neri Merhav

Department of Electrical Engineering
Technion - Israel Institute of Technology
Technion City, Haifa 32000, ISRAEL
E–mail: merhav@ee.technion.ac.il

Abstract

The security level of the achievability scheme for Wyner's wiretap channel model is examined from the perspective of the probability of correct decoding, P_c, at the wiretap channel decoder. In particular, for finite–alphabet memoryless channels, the exact random coding exponent of P_c is derived as a function of the total coding rate R_1 and the rate of each sub–code R_2. Two different representations are given for this function and its basic properties are provided. We also characterize the region of pairs of rates (R_1, R_2) of full security in the sense of the random coding exponent of P_c, in other words, the region where the exponent of this achievability scheme is the same as that of blind guessing at the eavesdropper side. Finally, an analogous derivation of the correct–decoding exponent is outlined for the case of the Gaussian channel.

Index Terms: Wiretap channel, random coding exponent, information–theoretic security, secrecy.

[*] This research was supported by the Israel Science Foundation (ISF), grant no. 412/12.

1 Introduction

In his seminal paper on the wiretap channel, Wyner [24], studied a model of secure communication across a physically degraded broadcast channel, without any secret key, where the legitimate user receives the output of the better channel and the eavesdropper receives the output of the noisier channel. In that paper, Wyner characterized the optimum trade–off between reliable coding rates for the legitimate user and the equivocation at the wiretapper, which was defined in terms of the conditional entropy of the source given the output of the bad channel, observed by the wiretapper. As a byproduct, he also established the notion of the *secrecy capacity*, which is the maximum coding rate that still enables perfect secrecy, where the equivocation is (asymptotically) equal to the unconditional entropy of the source, thus rendering the information accessible to the eavesdropper virtually useless. By using good codes with rates that approach the secrecy capacity, the channel is fully utilized in the sense that the additional noise at the bad channel output (beyond that of the good channel), is harnessed for securing the message in the best possible way. The idea of the construction of a good code for the wiretapped channel is essentially the same as in the concept of binning. One creates a relatively large randomized code, that is reliably decodable at the legitimate receiver. This code consists of an hierarchy of sub–codes, where each sub-code is reliably decodable individually by the wiretapper. The information that is decodable by the wiretapper is, however, only the one that pertains to the randomization, and thus irrelevant to the source, which is statistically independent.

During the four decades that have passed since [24] was published, the wiretap channel model has been extended in many ways. We mention here only a few. Csiszár and Körner [6] have extended Wyner's setting to a broadcast channel which is not necessarily of a degraded structure. At the same year, Leung–Yan–Cheong and Hellman [8], studied the Gaussian wiretap channel, and have shown that its secrecy capacity is simply the difference between the capacities of the main (legitimate) channel and the wiretap channel. In [19], Ozarow and Wyner studied the so called type II wiretap channel, where the main channel (to the legitimate user) is noiseless, but the wiretapper knows some of the coded bits, and optimal trade-offs were characterized. In [25], the wiretap channel model was generalized to have two parallel broadcast channels, connecting one encoder and one legitimate decoder. According to this model, both channels are wiretapped by non–collaborating wiretappers, and again, optimum trade-offs were given in terms of single–letter expressions. In [26], the model was extended again, in two ways: First, by allowing also a secret key to be shared between the encoder and the legitimate receiver, and secondly, by allowing some distortion in the reproduction of the source at the legitimate receiver. The main coding theorem of [26] gives rise to a separation theorem, asserting that no asymptotic optimality is lost if the encoder first, applies rate–distortion source coding, then it encrypts the compressed bits, and finally, employs a channel code. Approximately a decade ago, the Gaussian wiretap channel model of [8] was further extended in two directions: one is the Gaussian multiple access wiretap channel of [23], and the other is Gaussian interference wiretap channel of [17], [18], where the encoder observes the interference signal as side information.

A comprehensive overview on modern information–theoretic security, in general, and on the wiretap channel, in particular, can be found in [9]. Finally, it should be pointed out that in the last few years, there has also been a research activity in developing constructive coding schemes, which are computationally practical, and at the same time, comply with more stringent security criteria, see, e.g., [2], [3], [10] and references therein.

In this paper, we adopt the large deviations notion of secrecy, that was proposed in [11], and apply it in the context of the wiretap channel.[1] According to this notion, secrecy is measured in terms of the exponential decay rate of the probability of correct decoding, P_c, by the wiretapper. The larger is the exponential decay rate of P_c, the better is the secrecy. In particular, full secrecy, in this sense, amounts to a situation where the exponent of P_c is not improved by the availability of the data accessed by the wiretapper (e.g., the cipher-text, or in the case, the wiretap channel output), compared to the exponent in the absence of this information, which is the exponent of the probability that a blind guess of the transmitted message would be successful. Accordingly, for the achievability scheme of [24], we analyze the random coding exponent of P_c at the wiretapper. It should be pointed out that our analysis, which is based on the type class enumeration method presented in [12, Section 6.3], yields the *exact* random coding exponent (not just a bound), and in particular, we derive two different, but equivalent, single–letter expressions for this exponent, denoted by $E(R_1, R_2)$, which is a function of the rate R_1 of the large code, and the rate R_2 of each

[1] For the sake of simplicity, we adopt Wyner's original model, but our results can be extended to some of the more general models discussed above.

sub–code, where $R_2 = R_1 - R$, R being the information rate conveyed to the legitimate user. Each one of these expressions reveals different properties of the function $E(R_1, R_2)$, which we will explore here. Among these properties, we characterize the region of rates where $E(R_1, R_2) = R_1 - R_2$, which in turn, is the exponent of blind guessing of the transmitted message. This means that in this region, the achievability scheme in [24] is perfectly secure in the large deviations sense of [11].

The remaining part of the paper is organized as follows. In Section 2, we establish notation conventions, provide some background, and formalize the problem. In Section 3, the main theorem of this paper is asserted, discussed and demonstrated. In Section 4, we prove this theorem. Finally, in Section 5, we give a brief outline for an analogous derivation of $E(R_1, R_2)$ for the Gaussian channel.

2 Notation Conventions, Preliminaries and Problem Formulation

2.1 Notation Conventions

Throughout the paper, random variables will be denoted by capital letters, specific values they may take will be denoted by the corresponding lower case letters, and their alphabets, similarly as other sets, will be denoted by calligraphic letters. Random vectors and their realizations will be denoted, respectively, by capital letters and the corresponding lower case letters, both in the bold face font. Their alphabets will be superscripted by their dimensions. For example, the random vector $\boldsymbol{X} = (X_1, \ldots, X_n)$, ($n$ – positive integer) may take a specific vector value $\boldsymbol{x} = (x_1, \ldots, x_n)$ in $\mathcal{X}^n$, the n–th order Cartesian power of $\mathcal{X}$, which is the alphabet of each component of this vector.

Probability distributions associated with sources and channels, will be denoted by the letters P and Q, with subscripts that denote the names of the random variables involved along with their conditioning, if applicable, following the customary notation rules in probability theory. For example, Q_{XZ} stands for a generic joint distribution $\{Q_{XZ}(x, z),\ x \in \mathcal{X},\ z \in \mathcal{Z}\}$, $P_{Z|X}$ denotes the matrix of transition probabilities of the underlying channel from X to Z, $\{P_{Z|X}(z|x),\ x \in \mathcal{X},\ z \in \mathcal{Z}\}$, and so on. Whenever there is no room for confusion, these subscripts may be omitted. Information measures induced by the generic joint distribution Q_{XZ}, or Q for short, will be subscripted by Q, for example, $H_Q(X)$ will denote the entropy of a random variable X drawn by Q, $I_Q(X; Z)$ will denote the corresponding mutual information, etc. When the underlying joint distribution is

$P_{XZ} = P_X \times P_{Z|X}$, this subscript may be omitted. The weighted divergence between two channels, $Q_{Z|X}$ and $P_{Z|X}$, with weight P_X, is defined as

$$D(Q_{Z|X} \| P_{Z|X} | P_X) \stackrel{\Delta}{=} \sum_{x \in \mathcal{X}} P_X(x) \sum_{z \in \mathcal{Z}} Q_{Z|X}(z|x) \ln \frac{Q_{Z|X}(z|x)}{P_{Z|X}(z|x)}. \tag{1}$$

The type class, $\mathcal{T}(P_X)$, associated with a given empirical probability distribution P_X of X, is the set of all $\boldsymbol{x} = (x_1, \ldots, x_n)$, whose empirical distribution is P_X. Similarly, the joint type class of pairs of sequences $\{(\boldsymbol{x}, \boldsymbol{z})\}$ in $\mathcal{X}^n \times \mathcal{Z}^n$, which is associated with an empirical joint distribution Q_{XZ}, will be denoted by $\mathcal{T}(Q_{XZ})$, and so on.

The expectation operator will be denoted by $\boldsymbol{E}\{\cdot\}$. Again, whenever there is room for ambiguity, the underlying probability distribution will appear as a subscript, e.g., $\boldsymbol{E}_Q\{\cdot\}$. Logarithms and exponents will be understood to be taken to the natural base unless specified otherwise. The indicator function will be denoted by $\mathcal{I}(\cdot)$. Sets will normally be denoted by calligraphic letters. The notation $[t]_+$ will stand for $\max\{t, 0\}$. For two positive sequences, $\{a_n\}$ and $\{b_n\}$, the notation $a_n \doteq b_n$ will mean asymptotic equivalence in the exponential scale, that is, $\lim_{n\to\infty} \frac{1}{n} \log(\frac{a_n}{b_n}) = 0$. Similarly, $a_n \stackrel{\cdot}{\le} b_n$ will mean $\limsup_{n\to\infty} \frac{1}{n} \log(\frac{a_n}{b_n}) \le 0$, and so on.

2.2 Preliminaries and Problem Formulation

We begin from a description of Wyner's model of the wiretap channel [24], with some simplification that makes it a pure channel coding model (as opposed to the model in [24], which includes also a source coding component).

Consider two discrete memoryless channels (DMC's), the *main channel*, $P_{Y|X} = \{P_{Y|X}(y|x),\ x \in \mathcal{X},\ y \in \mathcal{Y}\}$ and the *wiretap channel*, $P_{Z|X} = \{P_{Z|X}(z|x),\ x \in \mathcal{X},\ z \in \mathcal{Z}\}$, where $\mathcal{X}$, $\mathcal{Y}$, and $\mathcal{Z}$ are finite alphabets. The main channel serves the legitimate receiver, whereas the wiretap channel, as its name suggests, is at the service of the wiretapper. The wiretap channel is assumed to be a degraded version of the main channel, namely, there exists a channel $P_{Z|Y} = \{P_{Z|Y}(z|y),\ x \in \mathcal{X},\ y \in \mathcal{Y}\}$, such that

$$P_{Z|X}(z|x) = \sum_{y \in \mathcal{Y}} P_{Y|X}(y|x) P_{Z|Y}(z|y). \tag{2}$$

A randomized code of rate R, for this system, is an artificial channel $Q(\boldsymbol{x}|w)$ (subjected to design), that stochastically maps a positive integer $w \in \mathcal{M} = \{0, 1, 2, \ldots, M-1\}$, $M = e^{nR}$ (which designates the message), into a channel input vector $\boldsymbol{x} \in \mathcal{X}^n$. The message w is a realization of a random variable W, uniformly distributed over $\mathcal{M}$. Conceptually, one may think of the randomized mapping from w to $\boldsymbol{x} \in \mathcal{X}^n$ as a deterministic mapping $\boldsymbol{x} = f(w, b)$, where b is a realization of random variable B (drawn using resources of randomness available to the transmitter), independent of W, which is available to the transmitter, but not to the legitimate receiver or the wiretapper. Upon transmitting $\boldsymbol{X}$, the main channel outputs the vector $\boldsymbol{Y} \in \mathcal{Y}^n$ and the wiretap channel produces the vector $\boldsymbol{Z} \in \mathcal{Z}^n$.

One of the goals in [24] was to prove the existence of a stochastic encoder that, on the one hand, would guarantee reliable communication to the legitimate receiver (that is, an arbitrarily small probability of error in estimating W based on $\boldsymbol{Y}$, for large enough n), and on the other hand, provide the largest possible equivocation rate, $\limsup_{n\to\infty} H(W|\boldsymbol{Z})/n$, at the wiretapper side. In that paper, Wyner characterized the optimum trade-off between the achievable information rate R for reliable communication to the legitimate user and the equivocation rate. In particular, he has also established the notion of the *secrecy capacity* as the largest reliable information rate R for which the best achievable asymptotic equivocation rate is still as large as $\lim_{n\to\infty} H(W)/n = R$, namely, full secrecy in terms of equivocation.

The achievability scheme in [24] is based on random coding: select independently at random $M_1 = e^{nR_1}$ codewords in $\mathcal{X}^n$ using a product distribution $\prod_{i=1}^n P_X(x_i)$, where R_1 is chosen slightly smaller than $I(X;Y)$, the mutual information associated with $P_{XY} = P_X \times P_{Y|X}$. Partition the resulting codebook $\mathcal{C} = \{\boldsymbol{x}_0, \boldsymbol{x}_1, \ldots, \boldsymbol{x}_{M_1-1}\}$ into $M = M_1/M_2 = e^{nR}$ sub-codes $\{\mathcal{C}_w\}_{w=0}^{M-1}$, each of size $M_2 = e^{nR_2} = e^{n(R_1-R)}$, where R_2 is less than $I(X;Z)$, the mutual information induced by $P_{XZ} = P_X \times P_{Z|X}$. The partition of the large codebook into sub-codes is arbitrary. As in [24], we take it to be defined by $\mathcal{C}_w = \{\boldsymbol{x}_{wM_2}, \boldsymbol{x}_{wM_2+1}, \ldots, \boldsymbol{x}_{(w+1)M_2-1}\}$, $w = 0, 1, 2, \ldots, M-1$. Given $\mathcal{C}$, the stochastic encoder is defined by the channel

$$Q(\boldsymbol{x}|w) = \begin{cases} \frac{1}{M_2} & \boldsymbol{x} \in \mathcal{C}_w \\ 0 & \text{elsewhere} \end{cases} \tag{3}$$

In other words, $f(w, b) = \boldsymbol{x}_{wM_2+b}$, where b is a realization of a random variable B, which is uniformly distributed over $\{0, 1, 2, \ldots, M_2 - 1\}$. It is shown in [24] that this construction satisfies the direct part of the coding theorem, for the appropriate choice of P_X.

In this paper, we analyze the security of this achievability scheme (with a minor modification described below) from the viewpoint of the probability of correct decoding, P_c, at the wiretapper side, which is defined as

$$P_c \stackrel{\Delta}{=} \frac{1}{M}\sum_{w=0}^{M-1} \Pr\{\hat{w}(\boldsymbol{Z}) = w | W = w\}, \tag{4}$$

where $\hat{w}(\boldsymbol{z})$ is the decoded message based on the wiretap channel output $\boldsymbol{z}$. The ensemble average of P_c will be denoted by $\overline{P_c}$. In the interesting range of the operation of this communication system, $\overline{P_c}$ tends to zero exponentially rapidly, and our goal is to characterize the corresponding *correct–decoding exponent*, that is, the exponential decay rate,

$$E(R_1, R_2) \stackrel{\Delta}{=} \lim_{n\to\infty}\left[-\frac{\ln \overline{P_c}}{n}\right], \tag{5}$$

as a function of R_1 and R_2, where we assume that the wiretapper employs the optimum decoder in this setting, which is given by

$$\hat{w}(\boldsymbol{z}) = \arg\max_{w} P(\boldsymbol{z}|\mathcal{C}_w), \tag{6}$$

where

$$P(\boldsymbol{z}|\mathcal{C}_w) = \frac{1}{M_2}\sum_{\boldsymbol{x}\in\mathcal{C}_w} P(\boldsymbol{z}|\boldsymbol{x}) = \frac{1}{M_2}\sum_{i=wM_2}^{(w+1)M_2-1} P(\boldsymbol{z}|\boldsymbol{x}_i). \tag{7}$$

For a random coding distribution, we take the uniform distribution within a given type class $\mathcal{T}(P_X)$, rather than the above mentioned corresponding product distribution. It should be pointed out that our analysis can fairly easily be generalized to more complicated ensembles, which include hierarchical structures, similarly as those in the construction of ensembles of codes for the broadcast channel (see, e.g., [5, p. 565, proof of Theorem 15.6.2]). However, for the sake of simplicity of the exposition, we prefer to confine ourselves to the structure defined in the achievability part of [24], as described above.

Of course, similarly as in (6), the optimal decoder of the legitimate user seeks the message w that maximizes $P(\boldsymbol{y}|\mathcal{C}_w)$, which is defined similarly to (7), but with $\boldsymbol{z}$ replaced by $\boldsymbol{y}$. In the interesting range of rates, the average probability of error, associated with this decoder, decays exponentially, and the exact random coding exponent can be analyzed using the same methods as in [13] and [22]. However, our focus in this paper is primarily on security aspects, not quite on the random coding error exponent at the legitimate user, and so, we will not delve into this analysis here.

3 The Correct–Decoding Exponent of the Wiretapper

Our main result is the following (see Section 4 for the proof).

Theorem 1 *Consider the achievability scheme of [24] and the ensemble of codes defined in Subsection 2.2. Then,*

$$E(R_1, R_2) = \min\{E_1(R_1, R_2), E_2(R_1, R_2), E_3(R_1)\}, \tag{8}$$

where

$$\begin{aligned} E_1(R_1, R_2) &= R_1 - R_2 + \min_{Q_{Z|X}} \{D(Q_{Z|X} \| P_{Z|X} | P_X) : \ I_Q(X; Z) \leq R_2\} && (9) \\ E_2(R_1, R_2) &= R_1 + \min_{Q_{Z|X}} \{D(Q_{Z|X} \| P_{Z|X} | P_X) - I_Q(X; Z) : \ R_2 \leq I_Q(X; Z) \leq R_1\} && (10) \\ E_3(R_1) &= \min_{Q_{Z|X}} \{D(Q_{Z|X} \| P_{Z|X} | P_X) : \ I_Q(X; Z) \geq R_1\}, && (11) \end{aligned}$$

where $Q = Q_{XZ}$ must satisfy the constraint $Q_X = P_X$. An alternative representation of $E(R_1, R_2)$ is given by

$$\begin{aligned} E(R_1, R_2) = & \min_{\lambda_2 \in [0,1]} \max_{\lambda_1 \in [0,1]} \min_{Q_{Z|X}} \left\{D(Q_{Z|X} \| P_{Z|X} | Q_X) + \right. \\ & \left. (\lambda_1 + \lambda_2 - 1) I_Q(X; Z) + (1 - \lambda_1) R_1 - \lambda_2 R_2 \right\}. \end{aligned} \tag{12}$$

We now discuss the conclusions that can be drawn from Theorem 1 concerning the qualitative behavior of $E(R_1, R_2)$ and we demonstrate an example. It turns out that the two representations of $E(R_1, R_2)$, given in Theorem 1, reveal different properties of this function.

The first important feature is the partition of the plane R_1 vs. R_2 according to the region(s) where $E(R_1, R_2) > 0$, and the region(s) where $E(R_1, R_2) = 0$. The latter corresponds to a situation where the communication system is completely insecure, since $E(R_1, R_2) = 0$ may even correspond to a situation where $\overline{P_c}$ tends to unity as n grows without bound. The conditions for $E(R_1, R_2) = 0$ can easily be deduced from the first representation, given in eq. (8). Assume first that $R_1 - R_2 = R$ is strictly positive. In this case, $E_1(R_1, R_2)$ is always positive, as it is lower bounded by $R_1 - R_2$. $E_2(R_1, R_2)$ can vanish only if $I(X; Z) = R_1$, in which case, the minimizing $Q_{Z|X}$ is $P_{Z|X}$. Finally, $E_3(R_1)$ vanishes iff $I(X; Z) \geq R_1$. Thus, for $R > 0$, the overall correct–decoding $E(R_1, R_2)$ vanishes iff $I(X; Z) \geq R_1$, which makes sense, because in this case, the eavesdropper can even decode reliably the particular codeword that was sent, not only the sub–code $\mathcal{C}_m$ to which it belongs. For $R = 0$

$(R_1 = R_2)$, either $E_1(R_1, R_2)$ or $E_3(R_1)$ always vanishes, and so, $E(R_1, R_1) = 0$ in any case. This is also reasonable, because $R = 0$ means that there is only one sub–code $\mathcal{C}_0$, which is the entire code, so there is actually nothing to decode.

From the second representation of $E(R_1, R_2)$, given in eq. (12), it is easy to see that $E(R_1, R_2)$ is monotonically increasing in R_1 for fixed R_2 and monotonically decreasing in R_2 for fixed R_1. This is expected because as R_1 grows, the eavesdropper has more uncertainty (there are more sub–codes $\{\mathcal{C}_m\}$ that may be confusable for a given $M_2 = e^{nR_2}$), whereas if R_2 increases for fixed R_1, the uncertainty decreases. In [24], R_2 is chosen less than $I(X; Z)$ and R_1 is chosen slightly less than $I(X; Y)$, to achieve the maximum possible equivocation. For fixed R_1, the function $E(R_1, R_2)$ is concave in R_2, as in view of eq. (12), it can be seen as the minimum over a family of affine functions of R_2, parameterized by λ_2. It is not clear, however, if in general, for fixed R_2, the function $E(R_1, R_2)$ has a convexity or a concavity property (if any) in R_1. All the above mentioned properties of $E(R_1, R_2)$ are summarized in Fig. 1.

For $R_2 = 0$, we have the correct–decoding exponent of ordinary maximum likelihood decoding for a code at rate $R_1 = R$. In this case, the minimizing λ_2 always vanishes and the expression boils down to

$$E(R, 0) = \max_{\lambda_1 \in [0,1]} \min_{Q_{Z|X}} \left\{ D(Q_{Z|X} \| P_{Z|X} | Q_X) + \lambda_1 [R - I_Q(X; Z)] \right\}. \tag{13}$$

Eq. (12) lends itself more conveniently to explicit calculations of $E(R_1, R_2)$. According to this expression, the evaluation of $E(R_1, R_2)$ involves three steps of optimization, where the inner most minimization (over $Q_{Z|X}$) is a convex problem, and the two outer ones involve one parameter each. This form is fairly convenient at least in certain examples with a high enough degree of symmetry. We next demonstrate this on the simple example of the binary symmetric channel (BSC).

Example 1. Consider the example of the BSC, that is, $\mathcal{X} = \mathcal{Z} = \{0, 1\}$ and

$$P(z|x) = \begin{cases} 1 - p & z = x \\ p & z \neq x \end{cases} \tag{14}$$

Let $P_X(0) = P_X(1) = 1/2$. Since the minimization over $Q_{Z|X}$ is a convex program, the minimizing channel $Q_{Z|X}$ is a BSC as any non-symmetric channel can be improved by mixing it with its "mirror image" $Q'_{Z|X}(z|x) = Q_{Z|X}(1-z|1-x)$, which is equivalent in terms of both $D(Q_{Z|X} \| P_{Z|X} | P_X)$ and $I_Q(X; Z)$. Thus, the minimization over $Q_{Z|X}$ boils down to minimization over a single parameter

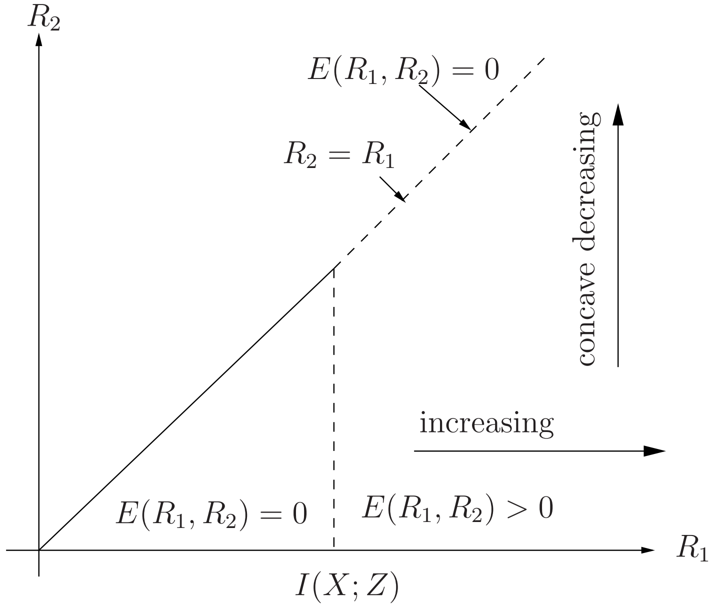

Figure 1: Behavior of $E(R_1, R_2)$ in the plane (R_1, R_2).

of this BSC, which is its crossover probability ϵ. In this case,

$$\begin{aligned}
& D(Q_{Z|X}\|P_{Z|X}|P_X) + (\lambda_1+\lambda_2-1)I_Q(X;Z) \\
= \; & D(\epsilon\|p) + (\lambda_1+\lambda_2-1)[\ln 2 - h(\epsilon)] \\
= \; & \epsilon \ln\frac{1}{p} + (1-\epsilon)\ln\frac{1}{1-p} - h(\epsilon) + (\lambda_1+\lambda_2-1)[\ln 2 - h(\epsilon)] \\
= \; & \epsilon \ln\frac{1}{p} + (1-\epsilon)\ln\frac{1}{1-p} - (\lambda_1+\lambda_2)h(\epsilon) + (\lambda_1+\lambda_2-1)\ln 2 \qquad (15)
\end{aligned}$$

This minimizing ϵ is easily found to be

$$\epsilon^* = \frac{p^{1/(\lambda_1+\lambda_2)}}{p^{1/(\lambda_1+\lambda_2)} + (1-p)^{1/(\lambda_1+\lambda_2)}}, \qquad (16)$$

which gives

$$\begin{aligned}
E(R_1,R_2) \;=\; & \min_{\lambda_2\in[0,1]} \max_{\lambda_1\in[0,1]} \{(\lambda_1+\lambda_2-1)\ln 2 - \\
& (\lambda_1+\lambda_2)\ln\left[p^{1/(\lambda_1+\lambda_2)} + (1-p)^{1/(\lambda_1+\lambda_2)}\right] + (1-\lambda_1)R_1 - \lambda_2 R_2\right\}. \qquad (17)
\end{aligned}$$

This concludes Example 1. □

In [11], a security criterion was defined in terms of the correct–decoding exponent. According to this criterion, a communication system is said to be secure in the correct–decoding exponent sense if the information available to the wiretapper, in our case $\boldsymbol{Z}$, does not help to improve (decrease) the exponent of $\overline{P_c}$ compared to the best achievable exponent in the absence of any information, that is, blind guessing. The exponent of $\overline{P_c}$ in blind guessing is clearly $R = R_1 - R_2$. In the presence of $\boldsymbol{Z}$, of course, $E(R_1, R_2) \le R_1 - R_2$ for all R_1 and R_2 ($R_2 \le R_1$). Let us examine if the equality $E(R_1, R_2) = R_1 - R_2$ is achievable in our system: To this end, we find it convenient to return to the first representation of $E(R_1, R_2)$, that is, eq. (8). The requirement $E(R_1, R_2) = R_1 - R_2$ is equivalent to the requirement that $E_1(R_1, R_2)$, $E_2(R_1, R_2)$ and $E_3(R_1)$ are all at least as large as $R_1 - R_2$. For $E_1(R_1, R_2)$, this requirement is trivially always met. For $E_3(R_1)$, this is simply the condition $R_2 \ge R_1 - E_3(R_1)$ (note that $E_3(R_1)$ vanishes for $R_1 \le I(X;Z)$ and it is monotonically increasing for $R_1 \ge I(X;Z)$). Concerning $E_2(R_1, R_2)$, this condition is equivalent to

$$\min_{Q_{Z|X}} \{D(Q_{Z|X} \| P_{Z|X} | P_X) - I_Q(X;Z) :\ R_2 \le I_Q(X;Z) \le R_1\} \ge -R_2. \tag{18}$$

To see that the set of pairs (R_1, R_2) that satisfy both requirements at the same time (as well as $R_2 \le R_1$) is non–empty, consider the following: Let $Q^*_{Z|X}$ be the unconstrained minimizer of the function $D(Q_{Z|X} \| P_{Z|X} | P_X) - I_Q(X;Z)$. Obviously,

$$D(Q^*_{Z|X} \| P_{Z|X} | P_X) - I_{Q^*}(X;Z) \le D(P_{Z|X} \| P_{Z|X} | P_X) - I_P(X;Z) = -I(X;Z), \tag{19}$$

and so,

$$I_{Q^*}(X;Z) \ge I(X;Z) + D(Q^*_{Z|X} \| P_{Z|X} | P_X) \ge I(X;Z), \tag{20}$$

where the second inequality is strict unless $Q^*_{Z|X} = P_{Z|X}$. Now, select R_1 and R_2 such that

$$R_1 > I_{Q^*}(X;Z) \ge R_2 \ge I_{Q^*}(X;Z) - D(Q^*_{Z|X} \| P_{Z|X} | P_X). \tag{21}$$

These choices clearly guarantee that both $R_1 > I(X;Z)$ and $R_1 > R_2$, which mean that $E(R_1, R_2) > 0$. Next, observe that

$$\begin{aligned} & \min_{Q_{Z|X}} \{D(Q_{Z|X} \| P_{Z|X} | P_X) - I_Q(X;Z) :\ R_2 \le I_Q(X;Z) \le R_1\} \\ = \ & D(Q^*_{Z|X} \| P_{Z|X} | P_X) - I_{Q^*}(X;Z) \\ \ge \ & -R_2, \end{aligned} \tag{22}$$

where the first inequality is due to the choices $R_1 > I_{Q^*}(X;Z) \ge R_2$ (which make the constraints inactive), and the last inequality is from $R_2 \ge I_{Q^*}(X;Z) - D(Q^*_{Z|X} \| P_{Z|X} | P_X)$. Thus, the requirement (18) is satisfied. It remains to show that eq. (21) is not in conflict with the requirement

$R_2 \geq R_1 - E_3(R_1)$, which is the case if we can show that $I_{Q^*}(X;Z) \geq R_1 - E_3(R_1)$. To see this, first recall that for every $Q_{Z|X}$

$$D(Q_{Z|X}\|P_{Z|X}|P_X) - I_Q(X;Z) \geq D(Q^*_{Z|X}\|P_{Z|X}|P_X) - I_{Q^*}(X;Z) \geq -I_{Q^*}(X;Z), \tag{23}$$

namely,

$$I_Q(X;Z) \leq I_{Q^*}(X;Z) + D(Q_{Z|X}\|P_{Z|X}|P_X). \tag{24}$$

Therefore, $I_Q(X;Z) \geq R_1$ implies $I_{Q^*}(X;Z) + D(Q_{Z|X}\|P_{Z|X}|P_X) \geq R_1$, or equivalently,

$$\forall\ Q \text{ with } I_Q(X;Z) \geq R_1, \text{ we have } D(Q_{Z|X}\|P_{Z|X}|P_X) \geq R_1 - I_{Q^*}(X;Z). \tag{25}$$

Since the right–hand side is independent of Q, then minimizing the left–hand side over all $\{Q:\ I_Q(X;Z) \geq R_1\}$ yields

$$\min\{D(Q_{Z|X}\|P_{Z|X}|P_X):\ I_Q(X;Z) \geq R_1\} \geq R_1 - I_{Q^*}(X;Z). \tag{26}$$

But the left–hand side is exactly $E_3(R_1)$, Thus, we have shown that $I_{Q^*}(X;Z)$ is never smaller than $R_1 - E_3(R_1)$. To summarize, the following procedure guarantees the choice of a non–trivial pair of rates that meets the requirements. First, calculate Q^* (which depends only on $P_{Z|X}$), then find $I_{Q^*}(X;Z)$ and $I(X;Z)$, and finally, select

$$R_1 > I_{Q^*}(X;Z), \tag{27}$$

and

$$\max\{I_{Q^*}(X;Z) - D(Q^*_{Z|X}\|P_{Z|X}|P_X), R_1 - E_3(R_1)\} \leq R_2 \leq I_{Q^*}(X;Z). \tag{28}$$

These choices comply with all requirements.

4 Proof of Theorem 1

Let $\boldsymbol{z} \in \mathcal{Z}^n$ be given, and let Q_{XZ} designate the empirical joint distribution pertaining to a (randomly chosen) codeword $\boldsymbol{x}$ together with $\boldsymbol{z}$, where it should be kept in mind that $Q_X = P_X$ by construction. For a given Q_{XZ}, let $N_w(Q_{XZ})$ denote the number of codewords in $\mathcal{C}_w$ whose empirical joint distribution with $\boldsymbol{z}$ is Q_{XZ}, that is

$$N_w(Q_{XZ}) = \sum_{i=wM_2}^{(w+1)M_2-1} \mathcal{I}\{(\boldsymbol{x}_i, \boldsymbol{z}) \in \mathcal{T}(Q_{XZ})\}, \quad w = 0, 1, \ldots, M-1. \tag{29}$$

We also denote

$$f(Q_{XZ}) = \frac{1}{n}\ln\left[\prod_{i=1}^{n} P_{Z|X}(z_i|x_i)\right] = \sum_{x,z} Q_{XZ}(x,z)\ln P_{Z|X}(z|x). \tag{30}$$

The probability of correct decoding, associated with the optimal decoder (6), is then given by

$$\begin{aligned}
P_c &= \frac{1}{M}\sum_{\boldsymbol{z}\in\mathcal{Z}^n}\max_{0\le w\le M-1} P(\boldsymbol{z}|\mathcal{C}_w) && (31)\\
&= \lim_{\beta\to\infty}\frac{1}{M}\sum_{\boldsymbol{z}\in\mathcal{Z}^n}\left[\sum_{w=0}^{M-1} P^{\beta}(\boldsymbol{z}|\mathcal{C}_w)\right]^{1/\beta} && (32)\\
&= \lim_{\beta\to\infty}\frac{1}{M}\sum_{\boldsymbol{z}\in\mathcal{Z}^n}\left[\sum_{w=0}^{M-1}\left(\frac{1}{M_2}\sum_{i=wM_2}^{(w+1)M_2-1} P(\boldsymbol{z}|\boldsymbol{x}_i)\right)^{\beta}\right]^{1/\beta} && (33)\\
&= \lim_{\beta\to\infty}\frac{1}{M_1}\sum_{\boldsymbol{z}\in\mathcal{Z}^n}\left[\sum_{w=0}^{M-1}\left(\sum_{i=wM_2}^{(w+1)M_2-1} P(\boldsymbol{z}|\boldsymbol{x}_i)\right)^{\beta}\right]^{1/\beta} && (34)\\
&= \lim_{\beta\to\infty}\frac{1}{M_1}\sum_{\boldsymbol{z}\in\mathcal{Z}^n}\left[\sum_{w=0}^{M-1}\left(\sum_{\{Q_{X|Z}:\ Q_X=P_X\}} N_w(Q_{XZ})e^{nf(Q_{XZ})}\right)^{\beta}\right]^{1/\beta} && (35)\\
&\doteq \lim_{\beta\to\infty}\frac{1}{M_1}\sum_{\boldsymbol{z}\in\mathcal{Z}^n}\left[\sum_{w=0}^{M-1}\sum_{\{Q_{X|Z}:\ Q_X=P_X\}} [N_w(Q_{XZ})]^{\beta}e^{n\beta f(Q_{XZ})}\right]^{1/\beta} && (36)\\
&= \lim_{\beta\to\infty}\frac{1}{M_1}\sum_{\boldsymbol{z}\in\mathcal{Z}^n}\left[\sum_{\{Q_{X|Z}:\ Q_X=P_X\}}\left(\sum_{w=0}^{M-1}[N_w(Q_{XZ})]^{\beta}\right)e^{n\beta f(Q_{XZ})}\right]^{1/\beta} && (37)\\
&\doteq \lim_{\beta\to\infty}\frac{1}{M_1}\sum_{\boldsymbol{z}\in\mathcal{Z}^n}\sum_{\{Q_{X|Z}:\ Q_X=P_X\}}\left(\sum_{w=0}^{M-1}[N_w(Q_{XZ})]^{\beta}\right)^{1/\beta} e^{nf(Q_{XZ})} && (38)\\
&= \frac{1}{M_1}\sum_{\boldsymbol{z}\in\mathcal{Z}^n}\sum_{\{Q_{X|Z}:\ Q_X=P_X\}}\left\{\max_{0\le w\le M-1} N_w(Q_{XZ})\right\}\cdot e^{nf(Q_{XZ})}. && (39)
\end{aligned}$$

Taking the expectation with respect to (w.r.t.) the ensemble of codes, we have

$$\overline{P_c} \doteq \frac{1}{M_1}\sum_{\boldsymbol{z}\in\mathcal{Z}^n}\sum_{\{Q_{X|Z}:\ Q_X=P_X}\boldsymbol{E}\left\{\max_{0\le w\le M-1} N_w(Q_{XZ})\right\}\cdot e^{nf(Q_{XZ})}. \tag{40}$$

But

$$\boldsymbol{E}\left\{\max_{0\le w\le M-1} N_w(Q_{XZ})\right\} = \sum_{t=1}^{M_2}\Pr\left\{\max_{0\le w\le M-1} N_w(Q_{XZ})\ge t\right\} \tag{41}$$

$$= \sum_{t=1}^{M_2} \Pr \bigcup_{w=0}^{M-1} \{N_w(Q_{XZ}) \geq t\} \tag{42}$$

$$\doteq \sum_{t=1}^{M_2} \min\left\{1, M \cdot \Pr\{N_0(Q_{XZ}) \geq t\}\right\}, \tag{43}$$

where in the last passage we have used the exponential tightness of the union bound (limited by unity) for pairwise independent events [20, Lemma A.2], [21, Lemma 1]. Our next objective then is to assess the behavior of $\Pr\{N_0(Q_{XZ}) \geq t\}$ for a given $1 \leq t \leq M_2$. Now, for a given Q_{XZ}, $N_0(Q_{XZ})$ is a binomial random variable of M_2 independent trials and a probability of success $p \doteq e^{-n[I_Q(X;Z)-\delta_n]}$, where $\delta_n = O((\log n)/n)$. If $p < t/M_2$, the event $\{N_0(Q_{XZ}) \geq t\}$ is a large deviations event, otherwise it occurs with high probability. Accordingly, the Chernoff bound on $\Pr\{N_0(Q_{XZ}) \geq t\}$ is as follows.

$$\begin{aligned}
\Pr\{N_0(Q_{XZ}) \geq t\} &\leq \begin{cases} \exp\{-M_2 D(\frac{t}{M_2}\|p)\} & p < t/M_2 \\ 1 & p \geq t/M_2 \end{cases} \\
&\leq \begin{cases} \exp\{-e^{nR_2} D(te^{-nR_2}\|e^{-n[I_Q(X;Z)-\delta_n]})\} & I_Q(X;Z) > R_2 - \frac{\ln t}{n} + \delta_n \\ 1 & I_Q(X;Z)) \leq R_2 - \frac{\ln t}{n} + \delta_n \end{cases}
\end{aligned} \tag{44}$$

where $D(a\|b)$, for $a, b \in [0,1]$, is the binary divergence function, that is

$$D(a\|b) = a \ln \frac{a}{b} + (1-a) \ln \frac{1-a}{1-b}. \tag{45}$$

Now, for $a \geq b$, the following inequality was proved in [12, pp. 167–168]:

$$D(a\|b) \geq a \left[\ln \frac{a}{b} - 1\right]_+. \tag{46}$$

Thus, the first line of (44) is further upper bounded by

$$\begin{aligned}
&\exp\left\{-e^{nR_2} te^{-nR_2} \left[\ln\left(\frac{te^{-nR_2}}{\exp\{-n[I_Q(X;Z)-\delta_n]\}}\right) - 1\right]_+\right\} \\
= \quad &\exp\left\{-nt\left[I_Q(X;Z) + \frac{\ln t}{n} - R_2 - \delta_n - \frac{1}{n}\right]_+\right\}.
\end{aligned} \tag{47}$$

At this point, we have to distinguish between three cases: (i) $I_Q(X;Z) \leq R_2$, (ii) $R_2 < I_Q(X;Z) \leq R_1$, and (iii) $I_Q(X;Z) > R_1$.

For $I_Q(X;Z) \leq R_2$, we have the following consideration: As long as $t \leq e^{n[R_2 - I_Q(X;Z)]}$, the probability $\Pr\{N_0(Q_{XZ}) \geq t\}$ is nearly 1, and hence so is $\min\{1, M \cdot \Pr\{N_0(Q_{XZ}) \geq t\}\}$. For

$t > e^{n[R_2 - I_Q(X;Z)+\epsilon]}$, the probability $\Pr\{N_0(Q_{XZ}) \geq t\}$ decays double exponentially in n, and hence so does $\min\{1, M \cdot \Pr\{N_0(Q_{XZ}) \geq t\}\}$. Thus, in this case,

$$\boldsymbol{E}\left\{\max_{0\leq w\leq M-1} N_w(Q_{XZ})\right\} \doteq \sum_{t=1}^{M_2} \min\left\{1, M \cdot \Pr\{N_0(Q_{XZ}) \geq t\}\right\} \doteq e^{n[R_2 - I_Q(X;Z)]}. \tag{48}$$

In both cases (ii) and (iii), $N_0(Q_{XZ}) = 0$ with very high probability. Consider a fixed value of t (not growing with n). In this case, according to our general bound (47), $\Pr\{N_0(Q_{XZ}) \geq t\} \stackrel{\cdot}{\leq} e^{-nt[I_Q(X;Z)-R_2]}$.

For $R > I_Q(X;Z) - R_2$, which is case (ii), and $t < \lfloor R/[I_Q(X;Z) - R_2]\rfloor \stackrel{\triangle}{=} t_0$, we have $M \cdot \Pr\{N_0(Q_{XZ}) \geq t\} > 1$, and so, $\min\left\{1, M \cdot \Pr\{N_0(Q_{XZ}) \geq t\}\right\} = 1$. For $t > t_0$, the expression $\min\left\{1, M \cdot \Pr\{N_0(Q_{XZ}) \geq t\}\right\}$ decays exponentially with n, and so, $\boldsymbol{E}\left\{\max_{0\leq w\leq M-1} N_w(Q_{XZ})\right\}$, which is the sum over t, is dominated by t_0, which is a constant.

Finally, in case (iii), $M \cdot \Pr\{N_0(Q_{XZ}) \geq t\} \doteq e^{n\{R - t[I_Q(X;Z)-R_2]\}} < 1$ for all $t \geq 1$, and so,

$$\boldsymbol{E}\left\{\max_{0\leq m\leq M-1} N_m(Q_{XZ})\right\} \doteq \sum_{t=1}^{M_2} e^{n\{R - t[I_Q(X;Z)-R_2]\}} \doteq e^{n\{R - [I_Q(X;Z)-R_2]\}} = e^{n[R_1 - I_Q(X;Z)]}. \tag{49}$$

In summary, we have shown that

$$\boldsymbol{E}\left\{\max_{0\leq w\leq M-1} N_w(Q_{XZ})\right\} \doteq e^{n\Gamma(Q_{XZ}, R_1, R_2)} \tag{50}$$

where

$$\Gamma(Q_{XZ}, R_1, R_2) = \begin{cases} R_2 - I_Q(X;Z) & I_Q(X;Z) \leq R_2 \\ 0 & R_2 < I_Q(X;Z) \leq R_1 \\ R_1 - I_Q(X;Z) & I_Q(X;Z) > R_1 \end{cases} \tag{51}$$

with $Q = Q_{XZ}$ such that $Q_X = P_X$. Finally, we have

$$\begin{aligned} \overline{P_c} &\doteq e^{-nR_1} \sum_{\boldsymbol{z}\in\mathcal{Z}^n} \exp\left\{n \max_{\{Q_{X|Z}:\ Q_X = P_X\}} [\Gamma(Q_{XZ}, R_1, R_2) + f(Q_{XZ})]\right\} & (52) \\ &\doteq e^{-nR_1} \sum_{\mathcal{T}(Q_Z)} |\mathcal{T}(Q_Z)| \cdot \exp\left\{n \max_{\{Q_{X|Z}:\ Q_X = P_X\}} [\Gamma(Q_{XZ}, R_1, R_2) + f(Q_{XZ})]\right\} & (53) \\ &\doteq e^{-nR_1} \max_{Q_Z} e^{nH_Q(Z)} \cdot \exp\left\{n \max_{\{Q_{X|Z}::\ Q_X = P_X\}} [\Gamma(Q_{XZ}, R_1, R_2) + f(Q_{XZ})]\right\} & (54) \\ &= e^{-nE(R_1,R_2)} & (55) \end{aligned}$$

where

$$E(R_1, R_2) = R_1 + \min_{\{Q_{XZ}:\ Q_X = P_X\}} \left[\sum_{x,z} Q_{XZ}(x,z) \ln \frac{1}{P(z|x)} - \Gamma(Q_{XZ}, R_1, R_2) - H_Q(Z)\right] \tag{56}$$

$$= R_1 + \min_{\{Q_{XZ}:\ Q_X = P_X\}} \left[\sum_{x,z} Q_{XZ}(x,z) \ln \frac{Q_Z(z)}{P(z|x)} - \Gamma(Q_{XZ}, R_1, R_2)\right] \tag{57}$$

$$= R_1 + \min_{Q_{Z|X}} \left[D(Q_{Z|X} \| P_{Z|X} | P_X) - I_Q(X;Z) - \Gamma(Q_{XZ}, R_1, R_2)\right] \tag{58}$$

$$= \min\{E_1(R_1,R_2), E_2(R_1,R_2), E_3(R_1)\} \tag{59}$$

with $E_1(R_1,R_2)$, $E_2(R_1,R_2)$ and $E_3(R_1)$ being defined as in Theorem 1. This completes the proof of eq. (8).

Moving on to the proof of eq. (12), observe that

$$\Gamma(Q_{XZ}, R_1, R_2) = [R_2 - I_Q(X;Z)]_+ - [I_Q(X;Z) - R_1]_+. \tag{60}$$

$$\begin{aligned} E(R_1,R_2) &= R_1 + \min_{Q_{Z|X}} \big\{D(Q_{Z|X} \| P_{Z|X} | P_X) - I_Q(X;Z) + \\ &\quad [I_Q(X;Z) - R_1]_+ - [R_2 - I_Q(X;Z)]_+\big\} \\ &= R_1 + \min_{Q_{Z|X}} \min_{\lambda_2 \in [0,1]} \max_{\lambda_1 \in [0,1]} \big\{D(Q_{Z|X} \| P_{Z|X} | P_X) - I_Q(X;Z) + \\ &\quad \lambda_1 [I_Q(X;Z) - R_1] - \lambda_2 [R_2 - I_Q(X;Z)]\big\} \\ &= R_1 + \min_{Q_{Z|X}} \min_{\lambda_2 \in [0,1]} \max_{\lambda_1 \in [0,1]} \big\{D(Q_{Z|X} \| P_{Z|X} | P_X) + \\ &\quad (\lambda_1 + \lambda_2 - 1) I_Q(X;Z) - \lambda_1 R_1 - \lambda_2 R_2\big\} \\ &= R_1 + \min_{\lambda_2 \in [0,1]} \min_{Q_{Z|X}} \max_{\lambda_1 \in [0,1]} \big\{D(Q_{Z|X} \| P_{Z|X} | P_X) + \\ &\quad (\lambda_1 + \lambda_2 - 1) I_Q(X;Z) - \lambda_1 R_1 - \lambda_2 R_2\big\} \end{aligned} \tag{61}$$

Now,

$$\begin{aligned} &D(Q_{Z|X} \| P_{Z|X} | P_X) + (\lambda_1 + \lambda_2 - 1) I_Q(X;Z) \\ = &-\sum_{x,z} Q_{XZ}(x,z) \ln P(z|x) - (\lambda_1 + \lambda_2) H_Q(Z|X) + (\lambda_1 + \lambda_2 - 1) H_Q(Z) \\ = &-\sum_{x,z} Q_{XZ}(x,z) \ln P(z|x) + (\lambda_1 + \lambda_2) I_Q(X;Z) - H_Q(Z). \end{aligned} \tag{62}$$

The first term is affine in Q, the second and the third are convex. Thus, overall, the objective is convex in $Q_{Z|X}$ and concave (affine) in λ_2, a fact that allows us to interchange the inner minimization and maximization, and get

$$\begin{aligned} E(R_1,R_2) &= \min_{\lambda_2 \in [0,1]} \max_{\lambda_1 \in [0,1]} \min_{Q_{Z|X}} \big\{D(Q_{Z|X} \| P_{Z|X} | P_X) + \\ &\quad (\lambda_1 + \lambda_2 - 1) I_Q(X;Z) + (1 - \lambda_1) R_1 - \lambda_2 R_2\big\}. \end{aligned} \tag{63}$$

This completes the proof of Theorem 1. □

5 The Correct–Decoding Exponent for the Gaussian Channel

The proof of Theorem 1 relies heavily on the method of types [7] and therefore, strictly speaking, it is applicable to finite alphabets only. Nonetheless, the method of types has analogues for certain families of continuous alphabet sources and channels, most notably, exponential families [14], [16, Section VI] and in particular, the Gaussian channel (see, e.g., [1, Subsection VI.A], [15]). Accordingly, in this section, we provide a brief outline how an analogous derivation of $E(R_1, R_2)$ can be carried out for the additive white Gaussian noise channel. The rigorous derivation can be carried out following the techniques of [15].

Consider the additive white Gaussian noise channel,

$$\boldsymbol{Z} = \boldsymbol{X} + \boldsymbol{W}, \tag{64}$$

where $\boldsymbol{W} \sim \mathcal{N}(0, \sigma^2 I)$ and the random coding distribution is uniform across the surface of a hypersphere of radius $\sqrt{nS}$ ($S > 0$ being a given power constraint), centered at the origin. Here,

$$\begin{aligned} P(\boldsymbol{z}|\boldsymbol{x}) &= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{1}{2\sigma^2}\sum_{t=1}^{n}(z_t - x_t)^2\right\} && (65)\\ &= (2\pi\sigma^2)^{-n/2} \exp\left\{-\frac{n}{2\sigma^2}(\hat{\sigma}_z^2 - 2\hat{\rho}\sqrt{S}\hat{\sigma}_z + S)\right\} && (66)\\ &= \exp\left\{-n\left[\frac{1}{2}\ln(2\pi\sigma^2) + \frac{1}{2\sigma^2}(\hat{\sigma}_z^2 - 2\hat{\rho}\sqrt{S}\hat{\sigma}_z + S)\right]\right\}, && (67) \end{aligned}$$

where $\hat{\sigma}_z^2 = \frac{1}{n}\sum_{i=1}^n z_i^2$ and $\hat{\rho} = \sum_{t=1}^n x_t z_t/(n\sqrt{S}\hat{\sigma}_z)$. A natural definition of conditional type class of $\boldsymbol{x}$ given $\boldsymbol{z}$ is given by a prescribed value (within some infinitesimally small tolerance) of the empirical correlation $\hat{\rho}$. In modifying the proof of Theorem 1 to apply to this case, $I_Q(X; Z)$ should be replaced by $-\frac{1}{2}\ln(1 - \hat{\rho}^2)$, whereas $H_Q(Z)$ should be replaced by $\frac{1}{2}\ln(2\pi e\hat{\sigma}_z^2)$. Thus, referring to eq. (8), we now have

$$\begin{aligned} E(R_1, R_2) &= R_1 + \min_{\hat{\sigma}_z^2, \hat{\rho}}\left\{\frac{1}{2}\ln(2\pi\sigma^2) + \frac{1}{2\sigma^2}(\hat{\sigma}_z^2 - 2\hat{\rho}\sqrt{S}\hat{\sigma}_z + S) - \right.\\ &\qquad \left.\Gamma(\hat{\rho}, R_1, R_2) - \frac{1}{2}\ln(2\pi e\hat{\sigma}_z^2)\right\} && (68)\\ &= R_1 + \min_{\hat{\sigma}_z^2, \hat{\rho}}\left\{\frac{1}{2}\ln\frac{\sigma^2}{\hat{\sigma}_z^2} + \frac{1}{2\sigma^2}(\hat{\sigma}_z^2 - 2\hat{\rho}\sqrt{S}\hat{\sigma}_z + S) - \Gamma(\hat{\rho}, R_1, R_2) - \frac{1}{2}\right\} && (69)\\ &= R_1 + \min_{\hat{\sigma}_z^2, \hat{\rho}}\left\{\frac{1}{2}\left[\frac{(\hat{\rho}\hat{\sigma}_z - \sqrt{S})^2}{\sigma^2} + \frac{\hat{\sigma}_z^2(1 - \hat{\rho}^2)}{\sigma^2} - \ln\frac{\hat{\sigma}_z^2(1 - \hat{\rho}^2)}{\sigma^2} - 1\right] - \right. \end{aligned}$$

$$\frac{1}{2}\ln\frac{1}{1-\hat{\rho}^2} - \Gamma(\hat{\rho}, R_1, R_2)\Bigg\} \tag{70}$$

where

$$\Gamma(\hat{\rho}, R_1, R_2) = \left[R_2 + \frac{1}{2}\ln(1-\hat{\rho}^2)\right]_+ - \left[\frac{1}{2}\ln\frac{1}{1-\hat{\rho}^2} - R_1\right]_+, \tag{71}$$

the term in the square brackets (including the factor 1/2) is the analogue of the divergence term in (8) and the term $\frac{1}{2}\ln\frac{1}{1-\hat{\rho}^2}$ stands for the mutual information term therein. Here we have

$$E_1(R_1, R_2) = R_1 - R_2 + \min_{\hat{\sigma}_z^2} \min_{|\rho| \le \sqrt{1-e^{-2R_2}}} \frac{1}{2}\left[\frac{(\hat{\rho}\hat{\sigma}_z - \sqrt{S})^2}{\sigma^2} + \frac{\hat{\sigma}_z^2(1-\hat{\rho}^2)}{\sigma^2} - \ln\frac{\hat{\sigma}_z^2(1-\hat{\rho}^2)}{\sigma^2} - 1\right] \tag{72}$$

$$\begin{aligned} E_2(R_1, R_2) \quad = \quad & R_1 + \min_{\hat{\sigma}_z^2} \min_{\sqrt{1-e^{-2R_2}} \le |\rho| \le \sqrt{1-e^{-2R_1}}} \left\{\frac{1}{2}\left[\frac{(\hat{\rho}\hat{\sigma}_z - \sqrt{S})^2}{\sigma^2} + \frac{\hat{\sigma}_z^2(1-\hat{\rho}^2)}{\sigma^2} - \right.\right. \\ & \left.\left. \ln\frac{\hat{\sigma}_z^2(1-\hat{\rho}^2)}{\sigma^2} - 1\right] - \frac{1}{2}\ln\frac{1}{1-\hat{\rho}^2}\right\} \end{aligned} \tag{73}$$

and

$$E_3(R_1) = \min_{\hat{\sigma}_z^2} \min_{|\rho| \ge \sqrt{1-e^{-2R_1}}} \frac{1}{2}\left[\frac{(\hat{\rho}\hat{\sigma}_z - \sqrt{S})^2}{\sigma^2} + \frac{\hat{\sigma}_z^2(1-\hat{\rho}^2)}{\sigma^2} - \ln\frac{\hat{\sigma}_z^2(1-\hat{\rho}^2)}{\sigma^2} - 1\right]. \tag{74}$$

The minimization over $\hat{\sigma}_z$, in all three expressions, can be done in closed form (equating the derivative to zero results in a quadratic equation) and the minimizer is

$$\hat{\sigma}_z^* = \frac{1}{2}(\hat{\rho}\sqrt{S} + \sqrt{\hat{\rho}^2 S + 4\sigma^2}). \tag{75}$$

Upon substituting this back into the expressions if $E_1(R_1, R_2)$ $E_2(R_1, R_2)$, and $E_3(R_1)$, it remains to minimize only over $\hat{\rho}$. This minimization in turn is rather complicated to be carried in closed form, but it can always be carried out numerically by a line search, as the range of $\hat{\rho}$ is limited to a finite interval.

References

[1] E. Arikan and N. Merhav, "Guessing subject to distortion," *IEEE Trans. Inform. Theory*, vol. 44, no. 3, pp. 1041–1056, May 1998.

[2] M. Bellare, S. Tessaro, and A. Vardy, "A cryptographic treatment of the wiretap channel," arXiv:1201.2205v2 [cs.IT] 23 Jan 2012.

[3] M. Bellare, S. Tessaro, and A. Vardy, "Semantic security for the wiretap channel," *CRYPTO 2012*, Lecture Notes in Computer Science, vol. 7417, pp. 294–311, 2012.

[4] C. Chan, "Success exponent of wiretapper: a tradeoff between secrecy and reliability," arXiv:0805.3605v4 [cs.IT] 31 May 2008.

[5] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, John Wiley & Sons, Hoboken, New Jersey, U.S.A., 2006.

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

[7] I. Csiszár and J. Körner, *Information Theory – Coding Theorems for Discrete Memoryless Systems*, Academic Press, 1981.

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

[9] Y. Liang, H. V. Poor and S. Shamai (Shitz), "Information theoretic security," *Foundations and Trends in Communications and Information Theory*, vol. 5, issue 4–5, pp. 355–580, June 2009.

[10] H. Mahdavifar and A. Vardy, "Achieving the secrecy capacity of wiretap channels using polar codes," *IEEE Trans. Inform. Theory*, vol. 57, no. 10, pp. 6428–6443, October 2011.

[11] N. Merhav, "A large–deviations notion of perfect secrecy," *IEEE Trans. Inform. Theory*, vol. 49, no. 2, pp. 506–508, February 2003.

[12] N. Merhav, “Statistical physics and information theory,” (invited paper) *Foundations and Trends in Communications and Information Theory*, vol. 6, nos. 1–2, pp. 1–212, 2009.

[13] N. Merhav, “Codeword or noise? Exact random coding exponents for slotted asynchronism,” submitted to *IEEE Trans. Inform. Theory*, August 2013. http://arxiv.org/pdf/1308.4572.pdf

[14] N. Merhav, “On the estimation of the model order in exponential families,” *IEEE Trans. Inform. Theory*, vol. IT-35, no. 5, pp. 1109–1114, September 1989.

[15] N. Merhav, “Universal decoding for memoryless Gaussian channels with a deterministic interference,” *IEEE Trans. Inform. Theory*, vol. 39, no. 4, pp. 1261–1269, July 1993.

[16] N. Merhav, G. Kaplan, A. Lapidoth, and S. Shamai (Shitz), “On information rates for mismatched decoders,” *IEEE Trans. Inform. Theory*, vol. 40, no. 6, pp. 1953–1967, November 1994.

[17] C. Mitrpant, “Information hiding – an application of wiretap channels with side information,” Ph.D. dissertation, der Universitaet Duisburg–Essen, November 2003.

[18] C. Mitrpant, A. J. Han Vinck, and Y. Luo, “An achievable region for the Gaussian wiretap channel with side information,” preprint, 2004.

[19] L. H. Ozarow and A. D. Wyner, “Wire–tap channel II,” *Proc. Eurocrypt 84*, Workshop on Advances in Cryptology: Theory and Applications of Cryptographic Techniques, Paris, France, pp. 33–51, 1985.

[20] N. Shulman, *Communication over an Unknown Channel via Common Broadcasting*, Ph.D. dissertation, Department of Electrical Engineering – Systems, Tel Aviv University, July 2003. http://www.eng.tau.ac.il/∼shulman/papers/Nadav_PhD.pdf

[21] A. Somekh–Baruch and N. Merhav, “Achievable error exponents for the private fingerprinting game,” *IEEE Trans. Inform. Theory*, vol. 53, no. 5, pp. 1827–1838, May 2007.

[22] A. Somekh–Baruch and N. Merhav, “Exact random coding error exponents for erasure decoding,” *IEEE Trans. Inform. Theory*, vol. 57, no. 10, pp. 6444–6454, October 2011.

[23] E. Tekin and A. Yener, “The Gaussian multiple access wire–tap channel,” arXiv:cs.IT/0605028, May 7, 2006.

[24] A. D. Wyner, “The wire-tap channel,” *Bell System Technical Journal*, vol. 54, no. 8, pp. 1355–1387, October 1975.

[25] H. Yamamoto, “Coding theorems for secret sharing communication systems with two noisy channels,” *IEEE Trans. Inform. Theory*, vol. IT–35, no. 3, pp. 572–578, May 1989.

[26] H. Yamamoto, “Rate–distortion theory for the Shannon cipher system,” *IEEE Trans. Inform. Theory*, vol. IT–43, no. 3, pp. 827–835, May 1997.