The Shannon Cipher System with a Guessing Wiretapper

Neri Merhav, *Fellow, IEEE*, and Erdal Arikan, *Senior Member, IEEE*

***Abstract*—The Shannon theory of cipher systems is combined with recent work on guessing values of random variables. The security of encryption systems is measured in terms of moments of the number of guesses needed for the wiretapper to uncover the plaintext given the cryptogram. While the encrypter aims at maximizing the guessing effort, the wiretapper strives to minimize it, e.g., by ordering guesses according to descending order of posterior probabilities of plaintexts given the cryptogram. For a memoryless plaintext source and a given key rate, a single-letter characterization is given for the highest achievable guessing exponent function, that is, the exponential rate of the ρth moment of the number of guesses as a function of the plaintext message length. Moreover, we demonstrate asymptotically optimal strategies for both encryption and guessing, which are universal in the sense of being independent of the statistics of the source. The guessing exponent is then investigated as a function of the key rate and related to the large-deviations guessing performance.**

***Index Terms*—Cryptanalysis, cryptography, guessing, Shannon cipher system.**

I. Introduction

In the classical Shannon-theoretic approach to cryptology [10], the security of cipher systems is traditionally measured in terms of the equivocation, that is, the conditional entropy of the plaintext (or the key) given the cryptogram. As is well known (see, e.g., [8]), this conditional entropy can be at most as large as the rate of the purely random key stream. Thus perfect *theoretical secrecy* is attainable if and only if the key rate is at least as large as the message rate. This pessimistic result stimulated Shannon to also establish the notion of *practical secrecy*, which is measured by the average amount of work required to break the key given a certain amount of ciphertext. Diffie and Hellman [5] were the first to show that practical secrecy (or *computational security* in their terminology) is possible without any transfer of secret key between the sender and the legitimate receiver. The notion of computational security relies on the fact that certain computational tasks (such as factoring, or taking discrete logarithms of very large numbers) are considered difficult because there are no known procedures of performing them within reasonable amount of computation time.

Ever since these two pioneering papers of Shannon [10] and Diffie and Hellman [5] have been published, there has been a vast amount of research work on both theoretical and practical aspects of cryptography, which has been summarized in several excellent tutorial papers (see, e.g., [7], [8], and [11]). The universal assumption in most of these works is that, regardless of the computational resources that the enemy may have, s/he has exactly one chance to estimate the plaintext message or the key based on cryptogram (and perhaps also other side information that might be available). Success or failure are then determined by some measure of quality of this estimator, such as the probability of error or the distortion. The rationale behind this assumption is that in certain instances of the secure communications problem, the enemy may not have the chance to verify whether the estimated message is correct and to improve it if not.

But in other instances of the problem, the enemy eavesdropper might have a testing mechanism by which s/he can know whether the estimate was correct, and then more chances to guess the message in case of failure. For example, the enemy may wish to break an encrypted version of a secret personal verification information and/or an encrypted password into a computer account, or a bank account contacted via the Internet, or any other classified database that consists of sensitive information. Here it is clear that upon the first successful estimate, or guess, the system becomes accessible and hence the above mentioned testing mechanism naturally exists. In such cases, the enemy has the option to sequentially submit multiple estimates, or *guesses*, where at each trial, the fact that all previous guesses have failed, serves as an additional side information for the next guess. The work of Hellman [6] can be considered as one step in this direction of multiple guessing. Hellman proposed to measure the degree of security of a cryptosystem in terms of the expected number of *spurious messages*, i.e., the expected number of plaintext-key combinations that may explain the given cryptogram. The assumption in [6] is that the number of meaningful messages of a given length N within the language of the source, is very small compared to the total number of possible N-vectors.

In this paper, we aim at characterizing more directly the best attainable moments of the number of guesses that the eavesdropper may have to submit before success. To this end, the Shannon theory of cipher systems is combined with

Manuscript received January 1, 1998; revised February 16, 1999. The work of N. Merhav was supported by the Israel Science Foundation administered by the Israeli Academy of Sciences and Humanities.

N. Merhav is with the Department of Electrical Engineering, Technion–Israel Institute of Technology, Haifa 32000, Israel.

E. Arikan is with the Electrical-Electronics Engineering Department, Bilkent University, 06533 Ankara, Turkey.

Communicated by I. Csiszár, Associate Editor for Shannon Theory.

Publisher Item Identifier S 0018-9448(99)05878-2.

recent work on guessing values of random variables [1], [2]. Assuming that the generation of each guess demands a certain amount of computational burden on the wiretapper's part, this gives an alternative notion of computational security.

We consider Shannon's model of a secrecy system [10], where a message $\boldsymbol{X} = (X_1, \cdots, X_N)$ is to be communicated as securely as possible from a transmitter to a legitimate receiver. The transmitter and receiver have access to a common key string of K purely random bits $\boldsymbol{U} = (U_1, \cdots, U_K)$ that is independent of $\boldsymbol{X}$. The transmitter generates a cryptogram $\boldsymbol{Y} = f_N(\boldsymbol{X}, \boldsymbol{U})$ and sends it over a public channel to the receiver. The cryptogram $\boldsymbol{Y}$ is a string (possibly, of variable length) over an alphabet that is not necessarily the same as the source alphabet. The encryption function is invertible given the key in the sense that there exists an inverse, decryption function $\boldsymbol{X} = f_N^{-1}(\boldsymbol{Y}, \boldsymbol{U})$ to be used by the legitimate receiver who observes both $\boldsymbol{Y}$ and $\boldsymbol{U}$. An enemy wiretapper, who knows the encryption function f_N (and hence also the decryption function f_N^{-1}) and the statistics of the plaintext source, but not the key itself, aims at decrypting $\boldsymbol{X}$ from the observed cryptogram $\boldsymbol{Y}$ only. The wiretapper has a test mechanism by which s/he can identify whether any given candidate message $\hat{\boldsymbol{X}}$ is the true message. Given the encryption function f_N and the probability mass function of the plaintext messages $P(\boldsymbol{X})$, the posterior probabilities of all hypothesized plaintexts given the cryptogram $P(\boldsymbol{X}|\boldsymbol{Y})$ are all completely determined. Then, it is clear that the best guessing strategy (in any reasonable sense) is to first guess the most likely $\boldsymbol{X}$ given $\boldsymbol{Y}$, then try the second most likely guess, and so on, until eventually, the correct message is found. For a given sequential *guessing strategy*, i.e., an ordered list of guesses $\mathcal{G}_N = \{\hat{\boldsymbol{x}}_1(\boldsymbol{y}), \hat{\boldsymbol{x}}_2(\boldsymbol{y}), \cdots\}$ for any given $\boldsymbol{y}$, let the random variable $G_N(\boldsymbol{X}|\boldsymbol{Y})$ denote the number of guesses of the wiretapper until identification of the true message $\boldsymbol{X}$. In other words, $G_N(\boldsymbol{X}|\boldsymbol{Y})$ is the smallest integer i such that $\hat{\boldsymbol{x}}_i(\boldsymbol{Y}) = \boldsymbol{X}$. The degree of security can now be measured by the expected number of guesses $\boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})\}$, or more generally, by arbitrary positive moments $\boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})^\rho\}$, $\rho > 0$.

The goal is to investigate performance limits of such sequentially guessing wiretappers. For a memoryless plaintext source, we study the highest asymptotic exponential growth rate of the moment $\boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})^\rho\}$, as $N \to \infty$, attainable by the encrypter for a given *key rate* $K/N \to R$. This exponential growth rate of $\boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})^\rho\}$, as a function of R and ρ, is henceforth referred to as the *guessing exponent* function.

More precisely, let $f_1, f_2, \cdots$, denote a sequence of encryption functions, for $N = 1, 2, \cdots$, to be chosen by the encrypter. Since the wiretapper is assumed to know the encryption function f_N for every N and the plaintext message source P, we assume that the guessing wiretapper would always employ the best guessing strategy for f_N and P, that is, order guesses according to descending posterior probabilities as explained above. Under this assumption, we define

$$E^-(R,\rho) = \liminf_{N\to\infty} \sup_{f_N} \frac{1}{N} \log \boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})^\rho\} \qquad (1)$$

and

$$E^+(R,\rho) = \limsup_{N\to\infty} \sup_{f_N} \frac{1}{N} \log \boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})^\rho\} \qquad (2)$$

where both limits are taken under the regime $\lim_{N\to\infty} K/N = R$. Our main result is that $E^+(R,\rho)$ and $E^-(R,\rho)$ are equal (i.e., the liminf and limsup are in fact limits) and both are given by the single-letter expression

$$E(R,\rho) \stackrel{\Delta}{=} \max_Q [\rho h(Q,R) - D(Q\|P)] \qquad (3)$$

where $h(Q,R) \stackrel{\Delta}{=} \min\{H(Q), R\}$, P is the memoryless source that governs the plaintext message, $H(Q)$ is the entropy associated with a memoryless source Q, and $D(Q\|P)$ is the information divergence between Q and P. Moreover, $E(R,\rho)$ is attainable by encryption and guessing strategies that are universal in the sense of being independent of P and ρ.

We also investigate the guessing exponent function $E(R,\rho)$ and examine its behavior as a function of R for fixed ρ. This study reveals that $E(R,\rho)$ exhibits different behavior in three different regions. For rates smaller than the entropy of the source $H(P)$, the guessing exponent grows linearly as $E(R,\rho) = \rho R$, which means that the key space is sufficiently small that exhaustive search over all $2^K = 2^{NR}$ possible key strings is the best thing to do, regardless of the statistics of the message source. On the other extreme, for key rates beyond a certain threshold that is larger than $H(P)$, the amount of randomness introduced by the key is so large that the cryptogram becomes virtually useless for the purpose of guessing. In this case, the wiretapper may ignore the cryptogram altogether and submit "blind" guesses that are based only upon prior knowledge of P. The value of $E(R,\rho)$ coincides, in this range, with the guessing exponent without side information [1]. The threshold rate beyond which $E(R,\rho)$ exhibits this plateau behavior is given by the entropy $H(P_\rho)$ of an auxiliary memoryless source P_ρ whose letter probabilities are proportional to those of the original source P, raised to the power of $1/(1+\rho)$. Since $H(P_\rho)$ is never smaller, and normally strictly larger, than $H(P)$, this is a rather unexpected result. The reason is that, as mentioned earlier, $R = H(P)$ is well known to suffice for perfect secrecy in the traditional Shannon-theoretic sense. The explanation for this more demanding requirement on the key rate, lies in the fact that guessing performance is determined by the large deviations (atypical) behavior of the source, whereas the more familiar equivocation criterion has to do with the typical behavior. For key rates in the intermediate range $H(P) < R < H(P_\rho)$, it turns out that optimal guessing should target both the key and message statistics simultaneously. We describe such a guessing strategy and give an explicit expression for $E(R,\rho)$ for this range of key rates as well.

Finally, we relate the guessing exponent $E(R,\rho)$ to the best attainable large deviations performance defined as the probability of the event $G_N(\boldsymbol{X}|\boldsymbol{Y}) \geq 2^{NL}$ (L positive constant) as a function of L and R. It is shown that the exponential rate of this probability as a function of L for fixed R is the Fenchel–Legendre transform of $E(R,\rho)$ as function of ρ.

The outline of the paper is as follows. In the next section, we define the notation and give some definitions. In Section III, we

give a single-letter characterization of the guessing exponent function and in Section IV, we investigate this function. In Section V, we characterize the attainable large deviations performance of the guessing wiretapper, and show that the corresponding rate function is related to the guessing exponent function via the Fenchel–Legendre transform. Finally, in Section VI, we summarize the results and state some open problems.

II. Definitions and Notation Conventions

Throughout the paper, scalar random variables will be denoted by capital letters while their sample values will be denoted by the respective lower case letters. A similar convention will apply to random vectors and their sample values, which will be denoted by boldface letters. Thus for example, if $\boldsymbol{X}$ denotes a random vector $(X_1, \cdots, X_N)$, then $\boldsymbol{x} = (x_1, \cdots, x_N)$ would designate a specific realization of $\boldsymbol{X}$.

The plaintext message will be assumed to be drawn from a discrete memoryless source (DMS) with a finite alphabet $\mathcal{X}$ and probability mass function (PMF) $P = \{P(x), x \in \mathcal{X}\}$. The probability of a vector $\boldsymbol{x}$ will be denoted $P(\boldsymbol{x})$, which is given by $\Pi_{i=1}^{N} P(x_i)$. The Nth-order Cartesian power of $\mathcal{X}$, that is, the space of all N-vectors over $\mathcal{X}$, will be denoted by $\mathcal{X}^N$. The probability of an event $A \subseteq \mathcal{X}^N$ will be denoted by $P(A)$ or $\Pr\{A\}$. We shall use the letter Q to denote a generic DMS over the alphabet $\mathcal{X}$, and use the same notational conventions as for P.

For a DMS Q, we recall that the Shannon entropy is given by

$$H(Q) = -\sum_{x \in \mathcal{X}} Q(x) \log Q(x) \tag{4}$$

where logarithms throughout the sequel are taken to the base 2. The relative entropy between Q and P is defined as pt

$$D(Q\|P) = \sum_{x \in \mathcal{X}} Q(x) \log \frac{Q(x)}{P(x)}. \tag{5}$$

The Rényi entropy [9] of order α ($\alpha > 0$, $\alpha \neq 1$) associated with Q is defined as

$$H_\alpha(Q) = \frac{1}{1-\alpha} \log \sum_{x \in \mathcal{X}} Q(x)^\alpha \tag{6}$$

with $H_1(Q)$ being interpreted as the Shannon entropy $H(Q)$.

For a given source vector $\boldsymbol{x} \in \mathcal{X}^N$, the empirical probability mass function (EPMF) is the vector $Q_{\boldsymbol{x}} = \{Q_{\boldsymbol{x}}(a), a \in \mathcal{X}\}$, where $Q_{\boldsymbol{x}}(a) = N_{\boldsymbol{x}}(a)/N$, $N_{\boldsymbol{x}}(a)$ being the number of occurrences of the letter a in the vector $\boldsymbol{x}$. The set of all EPMF's of vectors in $\mathcal{X}^N$, that is, rational PMF's with denominator N, will be denoted by $\mathcal{Q}_N$. The type class $T_{\boldsymbol{x}}$ of a vector $\boldsymbol{x}$ is the set of all vectors $\boldsymbol{x}' \in \mathcal{X}^N$ such that $Q_{\boldsymbol{x}'} = Q_{\boldsymbol{x}}$. When we need to attribute a type class to a certain rational PMF $Q \in \mathcal{Q}_N$ rather than to a sequence in $\mathcal{X}^N$, we shall use the notation T_Q. It is well known [4] that the number of type classes of N-vectors is bounded by $(N+1)^{|\mathcal{X}|-1}$, where $|\mathcal{X}|$ denotes the cardinality of $\mathcal{X}$. The standard reference about the method of types is the book by Csiszár and Körner [4]. Finally, throughout the sequel, $o(N)$ designates a quantity that grows sublinearly with N, i.e., $o(N)/N \to 0$ as $N \to \infty$.

III. The Guessing Exponent Function

Our main result in this section is the following.

Theorem 1: For every DMS P and every $\rho > 0$

$$E^+(R, \rho) = E^-(R, \rho) = E(R, \rho) \tag{7}$$

where $E(R, \rho)$ is defined as in (3).

The remaining part of this section is devoted to the proof of Theorem 1 along with a description of optimum strategies for both parties.

Proof: Since $E^-(R, \rho)$ clearly cannot be strictly larger than $E^+(R, \rho)$, it is sufficient to prove that

$$E^+(R, \rho) \leq E(R, \rho) \leq E^-(R, \rho). \tag{8}$$

The left inequality is a converse theorem from the viewpoint of cryptography and a direct theorem from the viewpoint of cryptanalysis, whereas the right inequality is the other way around.

We start from the proof of the left inequality. For the sake of simplicity, we will present a suboptimal (but asymptotically optimal) guessing strategy that is easy to analyze. Consider first a guessing strategy that ignores the cryptogram altogether: Let $\boldsymbol{x}_1, \boldsymbol{x}_2, \cdots$ consist of an enumeration of all vectors of $\mathcal{X}^N$ in ascending order of empirical entropies, i.e., $H(Q_{\boldsymbol{x}_1}) \leq H(Q_{\boldsymbol{x}_2}) \leq \cdots$. More precisely, suppose one first lists all elements of the type class T_Q with the minimum entropy $H(Q)$, then those of the type class with the second smallest entropy, and so on. (The ordering within each type class is immaterial.) Now, if the message $\boldsymbol{x}$ belongs to T_Q, then the number of guesses is clearly upper-bounded by $\Sigma_{Q': H(Q') \leq H(Q)} |T_{Q'}|$. Since $|T_{Q'}| \leq 2^{NH(Q')}$ [4, p. 30] and the number of type classes is bounded polynomially in N, the total number of guesses is further upper-bounded by $2^{NH(Q)+o(N)}$.

Consider next, an exhaustive key-search attack defined by using the following guessing list: $f_N^{-1}(\boldsymbol{y}, \boldsymbol{u}_1), f_N^{-1}(\boldsymbol{y}, \boldsymbol{u}_2), \cdots$, where $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots$ is an arbitrary ordering of all possible key streams of length $K = NR$. Clearly, this guessing list finds any message $\boldsymbol{x}$ using no more than 2^{NR} guesses. Finally, to gain the benefits of both lists, let us examine the interlaced list

$$\mathcal{G}_N^* = \{\boldsymbol{x}_1, f_N^{-1}(\boldsymbol{y}, \boldsymbol{u}_1), \boldsymbol{x}_2, f_N^{-1}(\boldsymbol{y}, \boldsymbol{u}_2), \cdots\}$$

which needs no more than twice the number of guesses of the better of the two original lists for any given message $\boldsymbol{x}$. Thus for any $\boldsymbol{x} \in T_Q$, the corresponding number of guesses is upper-bounded by

$$\begin{aligned} G_N^*(\boldsymbol{x}|\boldsymbol{y}) &\leq 2 \cdot \min\{2^{NR}, 2^{NH(Q)+o(N)}\} \\ &= 2^{N \min\{R, H(Q)\}+o(N)} = 2^{Nh(Q,R)+o(N)}. \end{aligned} \tag{9}$$

Since $P(T_Q) \leq 2^{-ND(Q\|P)}$ [4, p. 32], we obtain

$$\boldsymbol{E}\{G_N^*(\boldsymbol{X}|\boldsymbol{Y})^\rho\} \leq \sum_{Q \in \mathcal{Q}_N} 2^{-ND(Q\|P)} 2^{\rho N h(Q,R)+o(N)} \tag{10}$$

$$\leq 2^{N \max_Q [\rho h(Q,R) - D(Q\|P)]+o(N)} \tag{11}$$

$$= 2^{NE(R,\rho)+o(N)}. \tag{12}$$

Since the last inequality holds for every encryption function f_N, then by the definition of $E^+(R,\rho)$, we get

$$E^+(R,\rho) \le \limsup_{N\to\infty} \frac{1}{N}\log \boldsymbol{E}\{G_N^*(\boldsymbol{X}|\boldsymbol{Y})^\rho\} \le E(R,\rho) \tag{13}$$

completing the proof of the left inequality in (8).

To prove the right inequality in (8), consider the following encryption function f_N^*. Given a source vector $\boldsymbol{x}\in T_Q$, we first compress it losslessly into a codeword $c(\boldsymbol{x})$ of the following structure. The first field of $l_1(\boldsymbol{x}) = \lceil \log|\mathcal{Q}_N|\rceil$ bits describes the index of the type class $T_Q = T_{\boldsymbol{x}}$. The second field of $l_2(\boldsymbol{x}) = \lceil \log|T_Q|\rceil$ bits gives the index of $\boldsymbol{x}$ within T_Q. Now, assume that NR is an integer and consider the two cases $NR < \log|T_Q|$ and $NR \ge \log|T_Q|$. If $NR < \log|T_Q|$ then the second field of the code is in turn implemented in two parts. We partition T_Q into $n = \lfloor |T_Q|/2^{NR}\rfloor$ disjoint subsets $T_Q^1, T_Q^2, \cdots, T_Q^n$, each of size 2^{NR}, and perhaps an additional remainder subset T_Q^{n+1} of size at most $2^{NR}-1$. Now, the first part of the second field encodes the index i of the subset T_Q^i that contains $\boldsymbol{x}$, whereas the second part, of NR bits, encodes the index of $\boldsymbol{x}$ within T_Q^i. Having compressed $\boldsymbol{x}$ in the above described manner, encryption is carried out as follows. If $NR \ge \log|T_Q|$, then the cryptogram $\boldsymbol{y}$ is the codeword $c(\boldsymbol{x})$ with the last $l_2(\boldsymbol{x})$ bits encrypted using simple bit-by-bit XOR with the bits of $\boldsymbol{U}$. (Note, that since NR is assumed integer, $NR \ge \log|T_Q|$ actually implies $NR \ge \lceil \log|T_Q|\rceil = l_2(\boldsymbol{x})$.) Otherwise, only the last NR bits of the codeword (that is, the second part of the second field) are encrypted in the above manner.

For the purpose of obtaining a lower bound on $\boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})^\rho\}$, we may assume that the guesser is informed of the type T_Q of the message $\boldsymbol{x}$. Obviously, any lower bound on $\boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})^\rho\}$ for such an informed guesser is also a lower bound for the original, uninformed guesser, because the class of guessing strategies with side information is a superset of the class of guessing strategies without it. Since P is assumed memoryless, then for any given Q, the conditional PMF $P(\boldsymbol{x}|\boldsymbol{x}\in T_Q)$ is uniform within T_Q independently of P. Due to the above described encryption mechanism, the conditional probability of $\boldsymbol{y}$ given $\boldsymbol{x}$ in T_Q, is given by $P(\boldsymbol{y}|\boldsymbol{x}) = 2^{-m(\boldsymbol{x})}$ for $\boldsymbol{y}\in B(\boldsymbol{x})$ and zero elsewhere, where $m(\boldsymbol{x}) = \min\{NR, \lceil \log|T_Q|\rceil\}$ and $B(\boldsymbol{x})$ is the set of $\boldsymbol{y}$-vectors that can be obtained as cryptograms of $\boldsymbol{x}$, i.e., all $\boldsymbol{y}$-vectors of the same length as $c(\boldsymbol{x})$, which agree with $c(\boldsymbol{x})$ except perhaps for the last $m(\boldsymbol{x})$ bits. By the Bayes rule, it now follows that for $\boldsymbol{x}\in T_Q$ and $\boldsymbol{y}\in B(\boldsymbol{x})$

$$\begin{aligned} P(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{x}\in T_Q) &= \frac{P(\boldsymbol{x}|\boldsymbol{x}\in T_Q)P(\boldsymbol{y}|\boldsymbol{x})}{\sum_{\boldsymbol{x}'\in T_Q} P(\boldsymbol{x}'|\boldsymbol{x}'\in T_Q)P(\boldsymbol{y}|\boldsymbol{x}')} \\ &= \frac{|T_Q|^{-1}2^{-m(\boldsymbol{x})}}{\sum_{\boldsymbol{x}'\in T_Q\cap B^{-1}(\boldsymbol{y})} |T_Q|^{-1}2^{-m(\boldsymbol{x}')}} \\ &\overset{a)}{=} \frac{1}{|T_Q\cap B^{-1}(\boldsymbol{y})|} \triangleq \frac{1}{M(\boldsymbol{y})} \end{aligned} \tag{14}$$

where $B^{-1}(\boldsymbol{y}) = \{\boldsymbol{x}\colon \boldsymbol{y}\in B(\boldsymbol{x})\}$ and equality a) follows from the fact that $m(\boldsymbol{x})$ is constant within a type class.

Now, since $P(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{x}\in T_Q)$ is a uniform PMF over a set of $M(\boldsymbol{y})$ elements, then for any guesser that is informed of the type of $\boldsymbol{X}$, we have

$$\begin{aligned} &\boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})^\rho|\boldsymbol{X}\in T_Q,\boldsymbol{Y}=\boldsymbol{y}\} \\ &\quad= \sum_{\boldsymbol{x}\in T_Q\cap B^{-1}(\boldsymbol{y})} P(\boldsymbol{x}|\boldsymbol{y},\boldsymbol{x}\in T_Q)G_N(\boldsymbol{x}|\boldsymbol{y})^\rho \\ &\quad= \frac{1}{M(\boldsymbol{y})}\sum_{i=1}^{M(\boldsymbol{y})} i^\rho \\ &\quad\ge \frac{1}{M(\boldsymbol{y})}\int_0^{M(\boldsymbol{y})} u^\rho\, du \\ &\quad= \frac{M(\boldsymbol{y})^\rho}{1+\rho}. \end{aligned} \tag{15}$$

Now there are three cases: If

$$NR \ge \log|T_Q|$$

then $T_Q\cap B^{-1}(\boldsymbol{y}) = T_Q$ and so, $M(\boldsymbol{y}) = |T_Q|$. Otherwise, if

$$NR < \log|T_Q|$$

and $\boldsymbol{y}$ falls in T_Q^i for some $1\le i\le n$, then $T_Q\cap B^{-1}(\boldsymbol{y}) = T_Q^i$, because any contents of the last NR bits form an existing codeword of some $\boldsymbol{x}\in T_Q$, and so, $M(\boldsymbol{y}) = 2^{NR}$. Finally, if

$$NR < \log|T_Q|$$

and $\boldsymbol{x}\in T_Q^{n+1}$, then $T_Q\cap B^{-1}(\boldsymbol{y}) = T_Q^{n+1}$ which might be small, but this happens with probability

$$|T_Q^{n+1}|/|T_Q| \le (2^{NR}-1)/|T_Q| \le 1/2$$

(even if n is as small as 1). Therefore, to summarize all three cases, we have the following:

$$\begin{aligned} &\boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})^\rho|\boldsymbol{X}\in T_Q\} \\ &\quad\ge \frac{\boldsymbol{E}\{M(\boldsymbol{Y})^\rho|\boldsymbol{X}\in T_Q\}}{1+\rho} \\ &\quad\ge \frac{1}{2}\,\frac{1}{1+\rho}\,[\min\{2^{NR},|T_Q|\}]^\rho \\ &\quad= \frac{1}{2(1+\rho)}2^{N\rho\min\{R,H(Q)-o(N)\}} \\ &\quad\ge \frac{1}{2(1+\rho)}2^{N\rho h(Q,R)-o(N)}. \end{aligned} \tag{16}$$

Finally, by averaging with respect to (w.r.t.) the probabilities of $\{T_Q\}$, taking advantage of the fact that $\Pr\{T_Q\} \ge 2^{-ND(Q\|P)-o(N)}$, and using the method of types, we conclude that for the above described encryption scheme, and for any guessing strategy

$$\liminf_{N\to\infty} \frac{1}{N}\log \boldsymbol{E}\{G_N(\boldsymbol{X}|\boldsymbol{Y})^\rho\} \ge E(R,\rho). \tag{17}$$

Since we have considered a specific encryption scheme, the left-hand side is clearly a lower bound on $E^-(R,\rho)$, and this completes the proof of the right inequality in (8). □

It is interesting to note that both the guessing strategy and the encryption strategy described in the above proof are universally asymptotically optimum in the sense of being independent of the underlying memoryless source P and the moment order ρ. Recall, that the strictly optimum guessing strategy depends on $\{P(\boldsymbol{x}|\boldsymbol{y})\}$ and hence also on $\{P(\boldsymbol{x})\}$.

IV. A More Explicit Expression

In this section, we give a more explicit expression for the guessing exponent function $E(R,\rho)$ and investigate its behavior as a function of R for fixed ρ.

First observe that

$$\begin{aligned}\rho h(Q,R) &= \rho \min\{H(Q),R\} \\ &= \min_{0\le\theta\le\rho}[\theta H(Q)+(\rho-\theta)R]. \end{aligned} \tag{18}$$

Substituting this into (3), we obtain

$$E(R,\rho) = \max_{Q}\min_{0\le\theta\le\rho}[\theta H(Q)+(\rho-\theta)R - D(Q\|P)] \tag{19}$$

$$= \min_{0\le\theta\le\rho}\max_{Q}[\theta H(Q)+(\rho-\theta)R - D(Q\|P)] \tag{20}$$

where the maximization and minimization are interchangeable because the bracketed expression is concave in Q and affine in θ.

Let $P_s = \{P_s(x), x\in\mathcal{X}\}$ denote an auxiliary DMS with letter probabilities given by

$$P_s(x) = \frac{P^{1/(1+s)}(x)}{\sum_{x'\in\mathcal{X}} P^{1/(1+s)}(x')}. \tag{21}$$

It is easy to show (see, e.g., [1]) that for $s>0$

$$\max_{Q}[sH(Q) - D(Q\|P)] = sH_{1/(1+s)}(P) \tag{22}$$

and that the maximum is achieved by $Q=P_s$. Thus we have

$$E(R,\rho) = \min_{0\le\theta\le\rho}[\theta H_{1/(1+\theta)}(P)+(\rho-\theta)R]. \tag{23}$$

It is also easy to check that

$$\frac{d}{d\theta}\left[\theta H_{1/(1+\theta)}(P)\right] = H(P_\theta). \tag{24}$$

Thus the derivative of bracketed term in (23) w.r.t. θ is $H(P_\theta)-R$. Since $H(P_\theta)$ is nondecreasing in $\theta\ge 0$ (as can be easily shown using (22)), the bracketed term in (23) has a nondecreasing slope and hence is convex in $\theta\ge 0$. So, for the minimum in (23) we have three cases: i) $H(P_\theta)-R>0$ for all $0\le\theta\le\rho$, or equivalently, $H(P)>R$, and the minimum is achieved at $\theta=0$; ii) $H(P_\theta)-R<0$ for all $0\le\theta\le\rho$, or equivalently, $H(P_\rho)<R$, and the minimum is achieved at $\theta=\rho$; and iii) there exists a unique solution $0\le\theta_R\le\rho$ to the equation $H(P_\theta)=R$ that achieves the minimum. These may be summarized as follows.

Proposition 1: The guessing exponent for a DMS is given by

$$E(R,\rho) = \begin{cases} \rho R, & R<H(P) \\ (\rho-\theta_R)R+\theta_R H_{1/(1+\theta_R)}(P), & H(P)\le R\le H(P_\rho) \\ \rho H_{1/(1+\rho)}(P), & R>H(P_\rho) \end{cases} \tag{25}$$

where θ_R is the unique solution of the equation $R=H(P_\theta)$ for R in the range $H(P)\le R\le H(P_\rho)$.

Thus for low rates, i.e., $R\le H(P)$, the guessing exponent $E(R,\rho)$ is just ρR, which can be interpreted as a situation where the key rate is so small that it pays off just to make an exhaustive search over all possible key sequences, namely, examine $f_N(\boldsymbol{y},\boldsymbol{u}_i)$, for all $i=1,2,\cdots,2^{NR}$, and essentially all of them will be examined (in the exponential sense).

On the other extreme of high key rates $R>H(P_\rho)$, we have $E(R,\rho)=\rho H_{1/(1+\rho)}(P)$ (a plateau region), which means that the cryptogram $\boldsymbol{Y}$ is so "noisy" that it is effectively useless for guessing $\boldsymbol{X}$ and the wiretapper might as well ignore it and guess at $\boldsymbol{X}$ directly only from knowledge of the prior probabilities $\{P(\boldsymbol{x})\}$. It is not surprising then, that the term $\rho H_{1/(1+\rho)}(P)$ coincides with the guessing exponent without side information studied in [1].

For key rates between $H(P)$ and $H(P_\rho)$, corresponding to the curvy part of the function $E(R,\rho)$, the optimal guessing strategy can be thought of as a combination of exhaustive search for the key and the message (in the spirit of the first part of the proof of Theorem 1).

Next consider the slope of $E(R,\rho)$ as a function of R for a fixed ρ. The partial derivative $\partial E(R,\rho)/\partial R$ equals ρ for $R<H(P)$, and equals zero for $R>H(P_\rho)$. For $H(P)<R<H(P_\rho)$, we have

$$\begin{aligned}\frac{\partial E(R,\rho)}{\partial R} &= \rho-\theta_R - R\frac{d\theta_R}{dR}+\frac{d\theta_R}{dR}\frac{d}{d\theta_R} \\ &\quad\cdot[\theta_R H_{1/(1+\theta_R)}(P)] \qquad (26)\\ &= \rho-\theta_R - R\frac{d\theta_R}{dR}+\frac{d\theta_R}{dR}H(P_{\theta_R}) \qquad (27)\\ &= \rho-\theta_R. \qquad (28)\end{aligned}$$

The function θ_R is increasing in R in the range $H(P)<R<H(P_\rho)$, which starts at $\theta_R=0$ for $R=H(P)$ and monotonically increases to $\theta_R=\rho$ at $R=H(P_\rho)$. Thus $\rho-\theta_R$ is decreasing in R, and hence, $E(R,\rho)$ is concave in $R\ge 0$ for any fixed $\rho\ge 0$. The typical shape of $E(\rho,R)$ as a function of R is shown in Fig. 1.

V. Large Deviations Performance

Moments of the number of guesses are intimately related to the large deviations performance of the guesser (see also [2], [3]), i.e., the best attainable exponential rate of $\Pr\{G_N(\boldsymbol{X}|\boldsymbol{Y})\ge 2^{NL}\}$ for some positive constant L. Analogously to the definitions regarding the guessing exponent

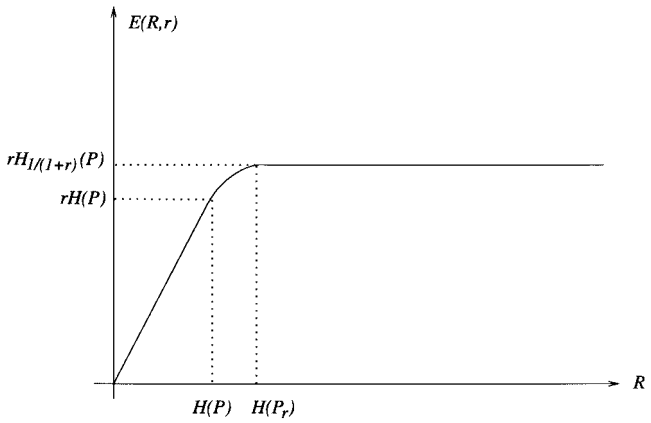

Fig. 1. Guessing exponent function $E(R,r)$ versus R.

function, let us define

$$F^{-}(R,L)=\liminf_{N\to\infty}\inf_{f_N}\left[-\frac{1}{N}\log\Pr\{G_N(\boldsymbol{X}|\boldsymbol{Y})\ge 2^{NL}\}\right] \tag{29}$$

and, similarly,

$$F^{+}(R,L)=\limsup_{N\to\infty}\inf_{f_N}\left[-\frac{1}{N}\log\Pr\{G_N(\boldsymbol{X}|\boldsymbol{Y})\ge 2^{NL}\}\right] \tag{30}$$

where the assumptions on the guessing strategy and on the asymptotic key rate are as above. Our next result is the following.

Theorem 2: For every DMS P and every $L>0$

$$\begin{aligned}F^{+}(R,L)&=F^{-}(R,L)=F(R,L)\\&\triangleq\min_{\{Q:\,h(Q,R)\ge L\}}D(Q\|P).\end{aligned} \tag{31}$$

Note that $F(R,L)$ is infinite for $L>R$, and given by the source-coding exponent [4, p. 45], $\min_{\{Q:\,H(Q)\ge R\}}D(Q\|P)$, for $L\le R$.

Proof: The proof is similar to the proof of Theorem 1. Again, it is sufficient to prove that

$$F^{+}(R,L)\le F(R,L)\le F^{-}(R,L). \tag{32}$$

For the left inequality, consider again the guessing strategy described in the proof of Theorem 1. Since

$$G_N(\boldsymbol{x}|\boldsymbol{y})\le 2\min\{2^{NR},2^{NH(Q_{\boldsymbol{x}})}\}=2\cdot 2^{Nh(Q_{\boldsymbol{x}},R)}$$

the probability that $G_N(\boldsymbol{X}|\boldsymbol{Y})$ would exceed 2^{NL} cannot be larger than the probability of the event $h(Q_{\boldsymbol{X}},R)+1/N>L$, which is easily shown (using the method of types) to decay exponentially at the rate of $F(R,L)$.

To prove the right inequality in (32), consider again the encryption scheme f_N^* described in the proof of Theorem 1. Using the same considerations as in the proof of Theorem 1, we have the following. For type classes whose size $|T_Q|$ is less than 2^{NR}

$$\Pr\{G_N(\boldsymbol{X}|\boldsymbol{Y})\ge 2^{NL}|\boldsymbol{X}\in T_Q\}=\begin{cases}0, & |T_Q|<2^{NL}\\ 1-\dfrac{2^{NL}}{|T_Q|}, & |T_Q|\ge 2^{NL}.\end{cases} \tag{33}$$

On the other hand, for type classes whose size $|T_Q|$ is larger than 2^{NR}

$$\begin{aligned}&\Pr\{G_N(\boldsymbol{X}|\boldsymbol{Y})\ge 2^{NL}|\boldsymbol{X}\in T_Q\}\\&\qquad\ge\begin{cases}0, & 2^{NR}<2^{NL}\\ \dfrac{1}{2}\left(1-\dfrac{2^{NL}}{2^{NR}}\right), & 2^{NR}\ge 2^{NL}.\end{cases}\end{aligned} \tag{34}$$

These two equations can be unified to

$$\begin{aligned}&\Pr\{G_N(\boldsymbol{X}|\boldsymbol{Y})\ge 2^{NL}|\boldsymbol{X}\in T_Q\}\\&\ge\begin{cases}0, & \min\{2^{NR},|T_Q|\}<2^{NL}\\ \dfrac{1}{2}\left(1-\dfrac{2^{NL}}{\min\{2^{NR},|T_Q|\}}\right), & \min\{2^{NR},|T_Q|\}\ge 2^{NL}.\end{cases}\end{aligned} \tag{35}$$

Thus

$$\begin{aligned}&\Pr\{G_N(\boldsymbol{X}|\boldsymbol{Y})\ge 2^{NL}\}\\&\quad\ge\sum_{Q\in\mathcal{Q}_N}P(T_Q)\Pr\{G_N(\boldsymbol{X}|\boldsymbol{Y})\ge 2^{NL}|\boldsymbol{X}\in T_Q\}\\&\quad\ge\sum_{\{Q:\,\min\{2^{NR},|T_Q|\}\ge 2^{NL+1}\}}P(T_Q)\cdot\frac{1}{4}\\&\quad\ge\sum_{\{Q:h(Q,R)\ge L-o(N)\}}2^{-ND(Q\|P)-o(N)}\\&\quad\ge 2^{-NF(R,L)-o(N)}.\end{aligned} \tag{36}$$

This completes the proof of Theorem 2. □

Note that the same encryption and guessing strategies of the proof of Theorem 1, are also asymptotically optimal in the large deviations sense.

We next show that $E(R,\rho)$ and $F(R,L)$ are related via the Fenchel–Legendre transform.

Theorem 3: For a DMS P and every key rate R

$$E(R,\rho)=\sup_{L>0}[\rho L-F(R,L)]=\sup_{0<L\le R}[\rho L-F(R,L)] \tag{37}$$

and

$$F(R,L)=\sup_{\rho>0}[\rho L-E(R,\rho)]. \tag{38}$$

Proof: The first equality of (37) is obtained as follows:

$$\begin{aligned}\sup_{L>0}[\rho L-F(R,L)]&=\sup_{L>0}[\rho L-\min_{\{Q:h(Q,R)\ge L\}}D(Q\|P)]\\&=\sup_{L>0}\max_{\{Q:h(Q,R)\ge L\}}[\rho L-D(Q\|P)]\\&=\max_{Q}\max_{\{L:\,L\le h(Q,R)\}}[\rho L-D(Q\|P)]\\&=\max_{Q}[\rho h(Q,R)-D(Q\|P)]\\&=E(R,\rho).\end{aligned} \tag{39}$$

The second equality of (37) follows from the fact that $F(R,L)=\infty$ for $L>R$. As for (38), we have the following:

$$\begin{aligned}
&\sup_{\rho>0}\{\rho L - E(R,\rho)] \\
&\quad = \sup_{\rho>0}[\rho L - \max_{Q}[\rho h(Q,R) - D(Q||P)]\} \\
&\quad = \sup_{\rho>0}\min_{Q}[\rho L - \rho h(Q,R) + D(Q||P)] \\
&\quad = \min_{Q}\sup_{\rho>0}[\rho L - \rho h(Q,R) + D(Q||P)] \qquad (40)\\
&\quad = \min_{\{Q:\, h(Q,R)\geq L\}} D(Q||P) \\
&\quad = F(R,L) \qquad (41)
\end{aligned}$$

where the interchangeability of minimization and maximization is justified by the fact that the bracketed expression is affine in ρ and concave in Q. This is true because $h(Q,R)$ is the minimum between a constant and a concave function of Q. This completes the proof of Theorem 3. □

VI. Conclusion and Further Research

In this paper, we introduced measures of cryptographic security that are based on the notion of guessing, and gave formulas for computing them. To this end, we have combined earlier works on guessing with Shannon-theoretic cryptography.

One important comment is in order: The Shannon cipher system that we have considered here allows for variable-length cryptograms. Therefore, strictly speaking, our results hold for encryption of a single block or, equivanently, under the assumption that the wiretapper knows (or is able to determine) the boundaries between the encrypted words given their concatenation. The natural question that arises here is what happens if this assumption is relaxed. The lower bound $E(R,\rho)$ to any moment of the number of guesses continues, of course, to hold because without the boundary information, the expected number of guesses can only grow. The upper bound remains valid as well as long as the length $l_{\max}(N)$ of the longest cryptogram $\boldsymbol{y}=f_N(\boldsymbol{x},\boldsymbol{u})$ (over all $\boldsymbol{x}$ and $\boldsymbol{u}$) is a subexponential function of N (normally it is linear). If this is the case, the guesser can synchronize to the encrypted bitstream by scanning $l^2_{\max}(N)$ hypotheses corresponding to $l_{\max}(N)$ consecutive possible locations of the beginning of the next encrypted word, times $l_{\max}(N)$ possible word lengths, and interlace the corresponding guesses according to the scheme described in the proof of Theorem 1. The total number of guesses would thereby increase by a factor of no more than $l^2_{\max}(N)$, which is still subexponential and hence would not affect the exponent $E(R,\rho)$.

We would like to mention some extensions of the present problem setting, which might be interesting to consider for future research. First, it should be stressed that the Shannon cipher system that we have considered here allows for variable-length cryptograms, and our results hold for an encryption of a single block or under the assumption that the wiretapper is synchronized First, it would be of interest to generalize the results to sources with memory, such as Markov sources, that can model natural languages. Secondly, one might consider the case in which the wiretapper is not required to reconstruct the message $\boldsymbol{X}$ exactly, but allowed some reconstruction error. In other words, as soon as the wiretapper provides a guess within distortion level D from the true message [2], we might regard the cipher as broken. The problem then is to determine the guessing and large deviations exponents. This type of reconstruction with some distortion has been studied by Yamamoto [12] in the ordinary paradigm of the Shannon cipher system. Another extension that might be considered is the case where the wiretapper observes a noisy version of the cryptogram, e.g., after $\boldsymbol{Y}$ passes through a noisy channel. It would be of interest to determine how the wiretapper's performance would be degraded in that case.

References

[1] E. Arikan, "An inequality on guessing and its application to sequential decoding," *IEEE Trans. Inform. Theory*, vol. 42, pp. 99–105, Jan. 1996.

[2] E. Arikan and N. Merhav, "Guessing subject to distortion," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1041–1056, May 1998.

[3] ——, "Joint source-channel coding and guessing with application to sequential decoding," *IEEE Trans. Inform. Theory*, vol. 44, pp. 1756–1769, Sept. 1998.

[4] I. Csiszár and J. Körner, *Information Theory: Coding Theorems for Discrete Memoryless Systems.* New York: Academic, 1981.

[5] W. Diffie and M. E. Hellman, "New directions in cryptography," *IEEE Trans. Inform. Theory*, vol. IT-22, pp. 644–654, Nov. 1976.

[6] M. E. Hellman, "An extension of the Shannon theory approach to cryptography," *IEEE Trans. Inform. Theory*, vol. IT-23, pp. 289–294, May 1977.

[7] A. Lempel, "Cryptology in transition," *Comput. Surv.*, vol. 11, no. 4, pp. 285–303, Dec. 1979.

[8] J. L. Massey, "An introduction to contemporary cryptology," *Proc. IEEE*, vol. 76, pp. 533–549, May 1988.

[9] A. Rényi, "On measures of entropy and information," in *Proc. 4th Berkeley Symp. Math. Statist. Probability* (Berkeley, CA), 1961, vol. 1, pp. 547–561.

[10] C. E. Shannon, "Communication theory of secrecy systems," *Bell Syst. Tech. J.*, vol. 28, no. 3, pp. 565–715, Oct. 1949.

[11] H. Yamamoto, "Information theory in cryptology," *IEICE Trans.*, vol. E 74, no. 9, pp. 2456–2464, Sept. 1991.

[12] ——, "Rate-distortion theory for the Shannon cipher system," *IEEE Trans. Inform. Theory*, vol. 43, pp. 827–835, May 1997.