Shannon Impossibility, Revisited

Yevgeniy Dodis
New York University
Email: `dodis@cs.nyu.edu`

In this note we revisit the famous result of Shannon [Sha49] stating that any encryption scheme with perfect security against computationally unbounded attackers must have a secret key as long as the message. This result motivated the introduction of modern encryption schemes, which are secure only against a computationally bounded attacker, and allow some small (negligible) advantage to such an attacker. It is a well known folklore that both such relaxations — limiting the power of the attacker and allowing for some small advantage — are necessary to overcome Shannon's result. To our surprise, we could not find a clean and well documented proof of this folklore belief. (In fact, two proofs are required, each showing that only one of the two relaxations above is not sufficient.) Most proofs we saw either made some limiting assumptions (e.g., encryption is deterministic), or proved a much more complicated statement (e.g., beating Shannon's bound implies the existence of one-way functions [IL89].)

In this note we rectify this situation, by presenting two clean, elementary extensions of Shannon's impossibility result, showing that, in order to beat the famous Shannon lower bound [Sha49] on key length for one-time-secure encryption, one must *simultaneously* restrict the attacker to be efficient, and also allow the attacker to break the system with some non-zero (i.e., negligible) probability. Unlike most prior proofs we have seen, our proof seamlessly handles *probabilistic* encryption, small decryption error, and can be taught without any extra background (e.g., notions of entropy, etc.) in a first lecture of an introductory cryptography class.

For intellectual curiosity, we also discuss some "entropy extensions" of our proof, and the relation between our "indistinguishability-based" proof and Shannon's original "mutual-information-based" proof.

Organization. The main results are presented in Sections 1 and 2, giving the main definitions and impossibility results. These are presented in a completely elementary way (e.g., no notion of entropy is used). In Section 3 we give some simple "entropy-based" extensions of our "indistinguishability-based" definition, and in Section 4 we also present the "mutual-information-based" definitions, and discuss their relation to "indistinguishability-based" notions.

1 Definitions

Some Notation. In general, we use capital letters for random variables, and lower case letters for specific values; e.g., M, C, S denote appropriately defined random messages, ciphertexts and keys, while m, c, s denote some specific value of those. When A is a probabilistic algorithm taking input x, we write $Y \leftarrow A(x)$ to denote the random variable $A(x; R)$ for uniformly random R. When X itself it a random variable, we write $Y \leftarrow A(X)$. Finally, we use calligraphic letters for message spaces; e.g., key space $\mathcal{S}$ and message space $\mathcal{M}$.

Encryption. Let $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ be any encryption scheme with key space $\mathcal{S}$ and message space $\mathcal{M}$. The key generation algorithm Gen outputs a secret key s chosen according to some key distribution S over $\mathcal{S}$. In most common schemes S is simply uniform over $\mathcal{S}$, but our results hold for any key distribution S, so we will not assume that S must be uniform.

The encryption algorithm Enc takes a key $s \in \mathcal{S}$, a message $m \in \mathcal{M}$, and outputs ciphertext $C \leftarrow \mathsf{Enc}_s(m)$. We stress that we allow the encryption algorithm Enc to be *probabilistic*, so C is really $\mathsf{Enc}_s(m; R)$ for random coins R. Luckily, we structure our proofs in a way which will easily handle this case, without explicitly talking about the random coins R. In particular, to simplify the notation, when some encryption is computed inside some probability, we do not explicitly put the choice or R under Pr; for example, $\Pr_S[\mathsf{Enc}_S(m) = c]$ really means $\Pr_{S,R}[\mathsf{Enc}_S(m; R) = c]$. We will assume that the message m is chosen from some distribution M over $\mathcal{M}$ which is independent of the key distribution $S \leftarrow \mathsf{Gen}()$.

The (possibly probabilistic) decryption algorithm Dec takes a ciphertext c and a key s and outputs the decryption $\tilde{M} \leftarrow \mathsf{Dec}_s(c)$. Ordinarily, we require *perfect correctness* stating that for any $m \in \mathcal{M}$ and $s \in \mathcal{S}$ we have $\mathsf{Dec}_s(\mathsf{Enc}_s(m)) = m$. However, since we are proving a lower bound, we relax the correctness guarantee to allow for some small decryption error γ.

Definition 1 An encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is called $(1-\gamma)$-*correct* on M if

$$\Pr_{S,M}[\mathsf{Dec}_S(\mathsf{Enc}_S(M)) = M] \geq 1 - \gamma \tag{1}$$

We say that $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is $(1-\gamma)$-*correct* (in general) if it is $(1-\gamma)$-correct on every message distribution M; equivalently, for any $m \in \mathcal{M}$, $\Pr_S[\mathsf{Dec}_S(\mathsf{Enc}_S(m)) = m] \geq 1 - \gamma$. ♢

Security. There are many equivalent formulations of "perfect" Shannon's security, when the attacker *Eve* is allowed to be computationally unbounded, and the "advantage" of any such *Eve* must be 0. Roughly, these definitions can be partitioned into two types. Some, including Shannon's original notion [Sha49], use the notions of Shannon's entropy and mutual information (see Section 4). While elegant and easy to state, it is not obvious how to relax such notions to *computationally bounded* attackers.[1] Other definitions, inspired by the Goldwasser-Micali [GM84] notions of semantic security and indistinguishability, are based on statistical distance. Such definitions have a clean and natural extensions to both computationally bounded attackers and non-zero advantage. Therefore, our definition below will be of this type. Since we are proving a lower bound, we will state what we feel is the *weakest* such definition. Of course, since our lower bound will be so strong even for such "weak-looking" definition, it will imply lower bounds for other, stronger definitions.

Definition 2 An encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is called (t, ε)-*secure on message distribution M* if for there exists a random variable Y (independent of M) such that for any (possibly

[1] However, in Section 4 we will propose a natural relaxation to small non-zero advantage.

probabilistic) adversary Eve running in time at most t, it holds

$$| \Pr_{S,M}[Eve(M, \mathsf{Enc}_S(M)) = 1] - \Pr_{S,Y}[Eve(M, Y) = 1] \,| \le \varepsilon \tag{2}$$

An encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is called (t, ε)-*secure* if it is (t, ε)-secure on all message distributions M. When Eve is allowed to be computationally unbounded (e.g., $t = \infty$), we say that $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is ε-*secure*. ♢

1.1 Few Remarks on the Definition

We make a few remarks on our definition. These remarks can be skipped by readers who already find the definition to be natural (and such readers can directly move to Section 2).

Intuitively, our definition states that whatever bit of information about M Eve could derive from the actual ciphertext C, she could have also derived from some random variable Y which is independent of M. Thus, Eve did not learn any new information from the ciphertext which she could not have learned from simply knowing the a-priori message distribution M (and some side information Y *independent* of M). However, while restricting Eve to run in time at most t, we do not make any restrictions on the complexity of sampling this independent distribution Y, and do not "charge" Eve for sampling Y. In particular, we do not insist on setting $Y \leftarrow \mathsf{Enc}_S(M')$, where M' is a fresh independent sample of M. Similarly, for general (t, ε)-security, we allow different Y's for different M's. Once again, such relaxations are done to make our lower bound stronger.

Also notice that the above definition is trivially true for any "singleton" distribution $M \leftarrow m$, for any $m \in \mathcal{M}$, and seems getting harder and harder as M becomes more and more "well-spread" (see Theorem 2 how this intuition translates to our lower bound). Still, even for the most "well-spread" uniform distribution M over $\mathcal{M}$, although we will see that our definition implies a strong bound on the size of the key space (Theorem 1), the definition is still noticeably weaker than general (t, ε)-security for *all* message distributions. For example, modifying a secure encryption (such as one-time pad) to be identity on some fixed $m \in \mathcal{M}$, still leaves the encryption very secure on the uniform distribution, while making the encryption of m easily distinguishable from encryptions of all other messages m'. In contrast, the general definition of security against all distributions is easily seen to be equivalent (ignoring factor of 2 in ε) to security against all distributions $M_{m,m'}$, for all $m, m' \in \mathcal{M}$, where each $M_{m,m'}$ is uniform over a pair of messages $\{m, m'\}$. In turn, the latter definition is simply the classical definition of (t, ε)-*indistinguishability* of Goldwasser-Micali [GM84], which states that for any messages $m, m' \in \mathcal{M}$, and any adversary Eve running in time at most t, it holds

$$|\Pr_S[Eve(\mathsf{Enc}_S(m)) = 1] - \Pr_S[Eve(\mathsf{Enc}_S(m')) = 1]| \le \varepsilon \tag{3}$$

We refer to [IO11] for discussions of several other nearly equivalent forms of "indistinguishability-based" security (such as semantic security) for one-time symmetric-key encryption, and stress that our lower bound easily holds for all such notions. We also discuss a natural "mutual-information-based" definition in Section 4.

2 Main Result

Recall the classical Shannon lower bound [Sha49] states that $(\infty, 0)$-security implies $|\mathcal{S}| \ge |\mathcal{M}|$. In fact, this conclusion holds even if M is restricted to be the uniform distribution over $\mathcal{M}$. Here we show an elegant extension of this result confirming that, in order to beat the Shannon bound in a non-trivial way, one must *simultaneously* restrict *Eve* to be efficient, as well as allow for some non-zero (but possibly negligible) probability ε of security failure. Just like the Shannon's original bound, our bounds will already follow by restricting M to be the uniform distribution. Our proof also handles decryption error γ.

Theorem 1 *Let M be the uniform distribution over $\mathcal{M}$, and assume* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* $(1-\gamma)$-*correct on M. Then:*

- **Small error needed.** *Let v denote maximum bit length of a plaintext plus ciphertext. If* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* $(v, 0)$-*secure on M, then* $|\mathcal{S}| \geq |\mathcal{M}|(1-\gamma)$.

- **Small time needed.** *Let d denote maximum decryption time. If* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* $(|\mathcal{S}|d, \varepsilon)$-*secure on M, then* $|\mathcal{S}| \geq |\mathcal{M}|(1-\varepsilon-\gamma)$.

Proof of First Part. Let Y be the distribution on ciphertexts guaranteed by Definition 2, so that Equation (2) holds with $\varepsilon = 0$ for any *Eve* running in time at most v. We claim that this implies that the joint distribution $(M, \mathsf{Enc}_S(M))$ is *identical* to (M, Y), where Y is independent from M:

$$(M, \mathsf{Enc}_S(M)) \equiv (M, Y) \tag{4}$$

To show this formally, for any fixed message $m \in \mathcal{M}$ and ciphertext c, consider the following $Eve_{m,c}(m', c')$ running in time $t = v$:

$Eve_{m,c}(m', c')$: *output* 1 *if and only if* $m' = m$ *and* $c' = c$.

Applying Equation (2) with $\varepsilon = 0$ to $Eve_{m,c}$, we get

$$\Pr_{S,M}[M = m \text{ and } \mathsf{Enc}_S(M)) = c] = \Pr_{M,Y}[M = m \text{ and } Y = c]$$

Using the fact that M is uniform and independent from Y, the above is equivalent to

$$\Pr_S[\mathsf{Enc}_S(m)) = c] = \Pr_Y[Y = c]$$

Since the above holds for all m and c, the distribution $\mathsf{Enc}_S(m) \equiv Y$ for all $m \in \mathcal{M}$, which means that the ciphertext distribution is the same for all messages. In particular, going back to the uniform distribution M, we have $(M, \mathsf{Enc}_S(M)) \equiv (M, Y)$, as claimed in Equation (4).[2]

Now, pick a fresh uniformly random key S' and look at[3]

$$\Delta \overset{\text{def}}{=} \Pr_{M,S',Y}[\mathsf{Dec}_{S'}(Y) = M] \tag{5}$$

On the one hand, it is clear that

$$\Delta \leq \frac{1}{|\mathcal{M}|} \tag{6}$$

[2] In essence, we showed a more general fact: to conclude that two distributions A and B are identical, it is sufficient to show that they are $(t, 0)$-indistinguishable, for t equal to the maximum description length of any element in the support of A and B.

[3] Note, if $S \leftarrow \mathsf{Gen}()$ is not uniform, S' has a different distribution than S.

since M is uniform and $\mathsf{Dec}_{S'}(Y)$ is independent of M. On the other hand, we know that the distribution (M, Y) is identical to $(M, \mathsf{Enc}_S(M))$. Hence, we can rewrite Equation (5) as

$$\begin{aligned}\Delta &= \Pr_{S,M,S'}[\mathsf{Dec}_{S'}(\mathsf{Enc}_S(M)) = M] \\ &\geq \Pr[S = S'] \cdot \Pr_{M,S}[\mathsf{Dec}_S(\mathsf{Enc}_S(M)) = M] \qquad (7) \\ &\geq \frac{1}{|\mathcal{S}|} \cdot (1-\gamma) \qquad (8)\end{aligned}$$

Here Equation (7) followed from the fact that the distribution of S conditioned on the event $S = S'$ is *the same* as the original distribution S, since S' is uniform. On the other hand, Equation (8) followed from Equation (1) and, again, the fact that S' is uniform, so $\Pr[S = S'] = 1/|\mathcal{S}|$.

Comparing the resulting inequality above with Equation (6), we get $\frac{1}{|\mathcal{S}|} \cdot (1-\gamma) \leq \Delta \leq \frac{1}{|\mathcal{M}|}$, which implies $|\mathcal{S}| \geq (1-\gamma)|\mathcal{M}|$. □

Proof of Second Part. We show that $(|\mathcal{S}|d, \varepsilon)$-security implies $|\mathcal{S}| \geq |\mathcal{M}|(1-\varepsilon-\gamma)$. As before, let Y be the ciphertext distribution guaranteed by Definition 2. Consider the following attacker *Eve* of complexity $t = |\mathcal{S}|d$:

Eve(m, c): *Run* $\mathsf{Dec}_s(c)$ *for all* $s \in \mathcal{S}$. *Output* 1 *if and only if at least one answer was* m.

Now, let us compute both probabilities when we apply Equation (2) to this *Eve*. First,

$$\begin{aligned}\Pr_{S,M}[Eve(M, \mathsf{Enc}_S(M)) = 1] &= \Pr_{S,M}[\exists s \text{ s.t. } \mathsf{Dec}_s(\mathsf{Enc}_S(M)) = M] \\ &\geq \Pr_{S,M}[\mathsf{Dec}_S(\mathsf{Enc}_S(M)) = M] \\ &\geq 1-\gamma\end{aligned}$$

where the last inequality used Equation (1). By Equation (2), we get

$$\Pr_{M,Y}[Eve(M, Y) = 1] \geq \Pr_{S,M}[Eve(M, \mathsf{Enc}_S(M)) = 1] - \varepsilon \geq 1 - \varepsilon - \gamma \qquad (9)$$

On the other hand,

$$\begin{aligned}\Pr_{M,Y}[Eve(M, Y) = 1] &= \Pr_{M,Y}[\exists s \text{ s.t. } \mathsf{Dec}_s(Y) = M] \\ &\leq \sum_s \Pr_{M,Y}[\mathsf{Dec}_s(Y) = M]\end{aligned}$$

However, M is uniform over $\mathcal{M}$ and, for any $s \in \mathcal{S}$, $\mathsf{Dec}_s(Y)$ is independent of M. Thus, $\Pr[M = \mathsf{Dec}_s(Y)] \leq \frac{1}{|\mathcal{M}|}$, which means that

$$\Pr_{M,Y}[Eve(M, Y)) = 1] \leq \sum_s \frac{1}{|\mathcal{M}|} = \frac{|\mathcal{S}|}{|\mathcal{M}|} \qquad (10)$$

Combining Equation (9) and Equation (10), we get $1 - \varepsilon - \gamma \leq \frac{|\mathcal{S}|}{|\mathcal{M}|}$ or $|\mathcal{S}| \geq |\mathcal{M}|(1-\varepsilon-\gamma)$. □

TIGHTNESS. Both bounds are nearly tight, which can be shown by tweaking the generalization of the one-time pad (OTP) encryption for general cardinality N message spaces (not just the power of 2, which can be accomplished by addition modulo N). For simplicity, we only do it for two special cases $\varepsilon = 0$ and $\gamma = 0$, leaving the common generalization as a (tedious) exercise. For both cases we will actually satisfy the stronger (t, ε)-indistinguishability given by Equation (3).

First, assume $\varepsilon = 0$. Take any $|\mathcal{M}|$ of cardinality N, and any subset $\mathcal{M}_0 \subseteq \mathcal{M}$ of cardinality $N(1-\gamma)$. Start with the OTP scheme over $\mathcal{M}_0$ (so that $|\mathcal{S}| = N(1-\gamma)$ as well), and enlarge it to all of $\mathcal{M}$ by taking any fixed $m_0 \in \mathcal{M}_0$ and defining $\mathsf{Enc}_s(m_1) = \mathsf{Enc}_s(m_0)$, for $m_1 \in \mathcal{M} \backslash \mathcal{M}_0$. The addition of these γN messages (which decrypt incorrectly) to our OTP does not affect the security of the scheme (since $\mathsf{Enc}(m_0)$ is perfectly secure), but creates a decryption error with probability γ, and with $|\mathcal{S}| = |\mathcal{M}|(1-\gamma)$.

Second, assume $\gamma = 0$. Now, for any $\mathcal{M}$ of cardinality N, take the OTP for $\mathcal{M}$ (so that $|\mathcal{S}| = N$), and simply remove $\varepsilon N/2$ keys from $\mathcal{S}$, defining the actual set $\mathcal{S}_0$ of $N(1-\varepsilon/2)$ keys, and sampling a random key s from $\mathcal{S}_0$. To argue the $\Omega(\varepsilon)$-security of this scheme, one can imagine sampling a key $s \leftarrow \mathcal{S}_0$ by first sampling the key $s \leftarrow \mathcal{S}$ and claiming that Eve unconditionally won the game if $s \in \mathcal{S} \backslash \mathcal{S}_0$. Equivalently, we can always actually run Eve on a fully uniform key s from $\mathcal{S}$, but then declare Eve victorious anyway if $s \in \mathcal{S} \backslash \mathcal{S}_0$. Clearly, when s is fully uniform, Eve has probability exactly $1/2$ telling apart encryptions of m_0 from m_1, so now her probability is at most $1/2 + \varepsilon/2$, creating distinguishing advantage at most ε with $|\mathcal{S}_0| = |\mathcal{M}|(1-\varepsilon/2)$.

3 Some Extensions

The result of the previous section was completely elementary, did not explicitly use any technical notions such as entropy, statistical distance, etc., and could be easily taught in the first lecture of an undergraduate class (especially for the case of perfect correctness $\gamma = 0$). In this section we make several elementary "entropy-extensions" of our main result.

3.1 Extension to general M

We observe that Theorem 1 easily generalizes to arbitrary message distributions M (as opposed to the uniform distribution), as follows. We define the *min-entropy* of M to be $\mathbf{H}_\infty(M) \overset{\text{def}}{=} -\log(\max_m \Pr[M = m])$. In particular, for any random variable M' independent of M, we have $\Pr[M' = M] \leq 2^{-\mathbf{H}_\infty(M)}$. Examining now the proofs of both parts of Theorem 1, we see that the only places where the uniformity of M was used were Equation (6) and Equation (10). In both cases, we needed to upped bound $\Pr[M' = M]$ for some probability distribution M' which was independent of M (e.g., $M' = \mathsf{Dec}_{S'}(Y)$ for Equation (6) and $M' = \mathsf{Dec}_s(Y)$ for Equation (10)). Hence, we get the following analog of Theorem 1 where $|\mathcal{M}|$ is replaced by $2^{\mathbf{H}_\infty(M)}$.

Theorem 2 *Let M be the any distribution over $\mathcal{M}$, and assume* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* $(1-\gamma)$*-correct on M. Then:*

- **Small error needed.** *Let v denote maximum bit length of a plaintext plus ciphertext.*
  If $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* $(v, 0)$*-secure on M, then* $|\mathcal{S}| \geq 2^{\mathbf{H}_\infty(M)} \cdot (1-\gamma)$.

- **Small time needed.** *Let d denote maximum decryption time.*
  If $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* $(|\mathcal{S}|d, \varepsilon)$*-secure on M, then* $|\mathcal{S}| \geq 2^{\mathbf{H}_\infty(M)} \cdot (1-\varepsilon-\gamma)$.

Notice, this bound is tight, in general, by taking M to be uniform over some subset $\mathcal{M}'$ of $\mathcal{M}$ of cardinality $2^{\mathbf{H}_\infty(M)}$, and then doing the OTP scheme over $\mathcal{M}'$.

3.2 Slightly Stronger Bound for Perfect Completeness and Perfect Security

Recall, the bounds of Theorem 1 (and more general Theorem 2) held for any key distribution $S \leftarrow \mathsf{Gen}()$, but only gave lower bounds of the cardinality of $\mathcal{S}$ (or, more generally, on cardinality of the support set of S). In contrast, as we recap in Section 4 below, Shannon's original bound [Sha49] gave the lower bound on the Shannon entropy $\mathbf{H}_1(S)$ of S, which could be stronger for sufficiently non-uniform S. Here we observe that our proof for the first part of Theorem 1 can be strengthened to give the lower bound on the min-entropy $\mathbf{H}_\infty(S)$ for the case of perfect correctness $\gamma = 0$. For elegance, we right away state the improved bound for general message distribution M as well.

Theorem 3 *Let M be the any distribution over $\mathcal{M}$, and assume* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* 1*-correct on M. Let v denote maximum bit length of a plaintext plus ciphertext. Then, if* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* $(v, 0)$*-secure on M, then* $\mathbf{H}_\infty(S) \geq \mathbf{H}_\infty(M)$*. In particular, if* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is* $(v, 0)$*-secure on uniform M, we have* $\mathbf{H}_\infty(S) \geq \log |\mathcal{M}|$.

Proof: We follow the same proof as in Theorem 1 (and its extension to general M in Theorem 2), except we define the value S' to be the most likely value s of the key S, instead of being uniform. Namely, we set $S' = s$ satisfying $\Pr[S = s] = 2^{-\mathbf{H}_\infty(S)}$. Then, the value Δ becomes

$$\Delta \stackrel{\text{def}}{=} \Pr_{M,Y}[\mathsf{Dec}_s(Y) = M]$$

We can argue, as before, that $\Delta \leq 2^{-\mathbf{H}_\infty(M)}$, since M is independent of $\mathsf{Dec}_s(Y)$. On the other hand, since the distribution (M, Y) is identical to $(M, \mathsf{Enc}_S(M))$ and we have perfect completeness, we get

$$\begin{aligned} \Delta &= \Pr_{S,M}[\mathsf{Dec}_s(\mathsf{Enc}_S(M)) = M] \\ &\geq \Pr[S = s] \cdot \Pr_M[\mathsf{Dec}_s(\mathsf{Enc}_s(M)) = M] \\ &\geq 2^{-\mathbf{H}_\infty(S)} \cdot 1 \qquad (11) \end{aligned}$$

where Equation (11) used the definition of s and the perfect correctness of the encryption. Combining the two bounds on Δ, we get $2^{-\mathbf{H}_\infty(S)} \leq \Delta \leq 2^{-\mathbf{H}_\infty(M)}$, which implies $\mathbf{H}_\infty(S) \geq \mathbf{H}_\infty(M)$. □

As we recap in Section 4 below, when $\varepsilon = 0$ our definition is equivalent to the original definition of Shannon [Sha49], who showed the bound $\mathbf{H}_1(S) \geq \log |\mathcal{M}|$, where $\mathbf{H}_1$ is Shannon's entropy. Since $\log |\mathcal{S}| \geq \mathbf{H}_1(S) \geq \mathbf{H}_\infty(S)$, we can view the last bound of Theorem 3 as a nice strengthening of Shannon's original bound for perfect security (and perfect correctness):[4] not only $\mathbf{H}_1(S) \geq \log |\mathcal{M}|$, but also $\mathbf{H}_\infty(S) \geq \log |\mathcal{M}|$.

[4] Actually, our proof above extends to imperfect correctness, as long as we require that $\Pr_M[\mathsf{Dec}_s(\mathsf{Enc}_s(M)) = M] \geq 1 - \gamma$, for all $s \in \mathcal{S}$, instead of only on average over $s \leftarrow S$.

4 Bounds for Mutual Information Based Definition

The *Shannon entropy* of a random variable X is defined as $\mathbf{H}_1(X) \overset{\text{def}}{=} \mathbb{E}_{x \leftarrow X}[-\log \Pr[X = x]]$. We also define *conditional Shannon entropy* of a random variable X conditioned on another random variable Z by

$$\mathbf{H}_1(X|Z) \quad \overset{\text{def}}{=} \quad \mathbb{E}_{(x,z)\leftarrow(X,Z)}\left[-\log \Pr[X = x|Z = z]\right]$$

where $\mathbb{E}_{z \leftarrow Z}$ denotes the expected value over $z \leftarrow Z$. It is well known that $\mathbf{H}_1(X) \geq \mathbf{H}_1(X|Z) \geq 0$. The *mutual information* between X and Y is $\mathbf{I}(X;Y) \overset{\text{def}}{=} \mathbf{H}_1(X) - \mathbf{H}_1(X|Y)$. It is well known that $\mathbf{I}(X;Y) = \mathbf{I}(Y;X) \geq 0$. The conditional mutual information of X and Y given Z is defined analogously. We assume the reader is familiar with other elementary facts about Shannon entropy and mutual information (such as the chain rule used below); see [CT06].

Let $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ be encryption scheme, S be it key distribution $\mathsf{Gen}()$, M be some message distribution and $C \leftarrow \mathsf{Enc}_S(M)$. We now give the following natural definitions generalizing the original definitions of [Sha49] to imperfect correctness and security.

Definition 3 An encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is called $(1-\gamma')$-*Shannon correct on* M if

$$\mathbf{H}_1(M|\mathsf{Dec}_S(C)) \leq \gamma'$$

$(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is $(1-\gamma')$-*Shannon correct* (in general) if it is $(1-\gamma')$-Shannon correct on all message distributions M.

An encryption scheme $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is called ε'-*Shannon secure on* M if

$$\mathbf{I}(M;C) \leq \varepsilon'$$

$(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ is ε'-*Shannon secure* (in general) if it is ε'-Shannon secure on all message distributions M. ♢

We start with the following Lemma, translating the elegant "proof-by-picture" exposition of Shannon's bound by Wolf [Wol98] (for $\varepsilon' = 0$) into a concrete inequality. (We suspect the Lemma is well-known, but we could not locate an explicit reference.)

Lemma 4 *For any (possibly correlated) distributions M, S, C, we have*

$$\mathbf{H}_1(S) \geq \mathbf{H}_1(M) - \mathbf{H}_1(M|(S,C)) - \mathbf{I}(M;C) \tag{12}$$

Proof:

$$\begin{aligned} \mathbf{H}_1(M) &= \mathbf{H}_1(M|(S,C)) + \mathbf{I}(M;(S,C)) \\ &= \mathbf{H}_1(M|(S,C)) + \mathbf{I}(M;C) + \mathbf{I}(M;S|C) \\ &= \mathbf{H}_1(M|(S,C)) + \mathbf{I}(M;C) + \mathbf{H}_1(S|C) - \mathbf{H}_1(S|(M,C)) \\ &\leq \mathbf{H}_1(M|(S,C)) + \mathbf{I}(M;C) + \mathbf{H}_1(S) \end{aligned}$$

where the equalities used the definitions and the chain rule, and the last inequality used the facts that $\mathbf{H}_1(S|C) \leq \mathbf{H}_1(S)$ and $\mathbf{H}_1(S|(M,C)) \geq 0$. □

As a corollary, we get the following straightforward extension of Shannon's result:

Theorem 5 *If* $(\mathsf{Gen}, \mathsf{Enc}, \mathsf{Dec})$ *is ε'-Shannon secure and $(1-\gamma')$-Shannon correct on M, then*

$$\mathbf{H}_1(S) \geq \mathbf{H}_1(M) - \gamma' - \varepsilon' \tag{13}$$

In particular, if M is the uniform distribution over $\mathcal{M}$, then $\mathbf{H}_1(S) \geq \log|\mathcal{M}| - \gamma' - \varepsilon'$.

Proof: Follows from Equation (12) and $\mathbf{H}_1(M|(S,C)) \leq \mathbf{H}_1(M|\mathsf{Dec}_S(C)) \leq \gamma'$. □

Relation to Indistinguishability Notions. Here we relate the $(1-\gamma')$-Shannon correctness and ε'-Shannon security to the "indistinguishability-based" notions of $(1-\gamma)$-correctness and ε-security[5] from Section 1. Using Fano's inequality (see [CT06]), we can relate γ' to γ as follows:

$$\gamma' \leq h(\gamma) + \gamma \cdot (\log|\mathcal{M}| - 1) \tag{14}$$

where h is binary entropy function $h(x) = -x\log_2 x - (1-x)\log_2(1-x)$. Unfortunately, no meaningful converse relation can be made, since changing $\mathsf{Dec}_S(C)$ to return $M+1$ instead of M has 0-correctness and 1-Shannon correctness.[6]

More interestingly, to relate Shannon security on M with indistinguishability security on M, we use the following result (implicitly) proven by Bellare et al. [BTV12] using Pinsker's inequality (see [CT06]).

Lemma 6 ([BTV12]) *For any (possibly correlated) distributions M, C over some spaces $\mathcal{M}$ and $\mathcal{C}$, let*[7]

$$\varepsilon = \mathbf{SD}((M,C); M \times C)$$

where $M \times C$ is the product distribution of the independent marginal distributions M and C. Then,

$$2\varepsilon^2 \leq \mathbf{I}(M;C) \leq 2\varepsilon \cdot \log(|\mathcal{M}|/\varepsilon) \tag{15}$$

In particular, notice that our notion of ε-security on M from Definition 2 is essentially equivalent to $\mathbf{SD}((M,C); M \times C) \leq \varepsilon$.[8] Thus, ε'-Shannon security on M implies $\sqrt{2\varepsilon'}$-security on M. Conversely, ε-security on M implies $(2\varepsilon \cdot \log(|\mathcal{M}|/\varepsilon))$-Shannon security on M. Hence, ignoring efficiency issues for Eve and the square root degradation on ε', Shannon starts with stronger security assumption than we do, but also gets slightly stronger conclusion: bound on $\mathbf{H}_1(S)$, not just $|\mathcal{S}|$. However, for perfect security $\varepsilon = 0$ we are still slightly stronger by Theorem 3, getting a bound on $\mathbf{H}_\infty(S)$, and not just $\mathbf{H}_1(S)$.

Acknowledgments: The author would like to thank Dario Fiore, Stefano Tessaro and Daniel Wichs for useful discussions.

[5] Here we set $t = \infty$, as it is not clear what is the analog of time for Shannon's security.

[6] Because of this, in contrast to standard correctness, the notion of "Shannon-correctness" is not a very useful notion, and we defined it only because the quantity $\mathbf{H}_1(M|\mathsf{Dec}_S(C))$ naturally came up in the proof. Luckily, Equation (14) shows that $(1-\gamma)$-correctness implies a decent bound on γ' as well.

[7] The *statistical distance* $\mathbf{SD}(X;Y)$ between two random variables X, Y is defined by:

$$\mathbf{SD}(X,Y) \overset{\text{def}}{=} \frac{1}{2}\sum_x |\Pr[X=x] - \Pr[Y=x]| = \max_{Eve} |\Pr[Eve(X)=1] - \Pr[Eve(Y)=1]|$$

[8] Up to a factor of 2 in ε since Y might not be equal to C. I.e., ε-security implies 2ε bound on the statistical distance above, and is implied by the ε bound on that distance.

References

[BTV12] Mihir Bellare, Stefano Tessaro and Alexander Vardy. Semantic Security for the Wiretap Channel. *CRYPTO 2012*. Earlier version available at *Cryptology ePrint Archive: Report 2012/015*, 2012.

[CT06] Thomas M. Cover, Joy A. Thomas. Elements of information theory (2. ed.). *Wiley*, 2006.

[GM84] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *JCSS*, **28**(2), pp. 270–299, April 1984.

[IL89] Russell Impagliazzo and Michael Luby. One-way Functions are Essential for Complexity Based Cryptography. *FOCS 1989*, pp. 230–235.

[IO11] Mitsugu Iwamoto and Kazuo Ohta. Security Notions for Information Theoretically Secure Encryptions. *ISIT 2011*. Available at `http://arxiv.org/abs/1106.1731`.

[Sha49] Claude Shannon. Communication Theory of Secrecy systems. In *Bell Systems Technical J.*, 28:656–715, 1949. Note: The material in this paper appeared originally in a confidential report 'A Mathematical Theory of Cryptography', dated Sept. 1, 1945, which has now been declassified.

[Wol98] Stefan Wolf. Unconditional Security in Cryptography. *Lectures on Data Security*, pp. 217–250, 1998.