Electronic Colloquium on Computational Complexity, Report No. 149 (2015)

Gambling, Computational Information and Encryption Security

Mohammad Hajiabadi [*] Bruce M. Kapron[*]

September 7, 2015

Abstract

We revisit the question, originally posed by Yao (1982), of whether encryption security may be characterized using computational information. Yao provided an affirmative answer, using a compression-based notion of computational information to give a characterization equivalent to the standard computational notion of semantic security. We give two other equivalent characterizations. The first uses a computational formulation of Kelly's (1957) model for "gambling with inside information", leading to an encryption notion which is similar to Yao's but where encrypted data is used by an adversary to place bets maximizing the rate of growth of total wealth over a sequence of independent, identically distributed events. The difficulty of this gambling task is closely related to Vadhan and Zheng's (2011) notion of KL-hardness, which in certain cases is equivalent to a conditional form of the pseudoentropy introduced by Hastad et. al. (1999). Using techniques introduced to prove this equivalence, we are also able to give a characterization of encryption security in terms of conditional pseudoentropy. Finally, we reconsider the gambling model with respect to "risk-neutral" adversaries in an attempt to understand whether assumptions about the rationality of adversaries may impact the level of security achieved by an encryption scheme.

1 Introduction

The first rigorous characterization of encryption security was given by Shannon in [21], using a formulation based on probability and information theory. The space of plaintexts is equipped with a probability distribution, which along with a distribution on keys induces a distribution on ciphertexts. An encryption scheme is said to be secure if the mutual information between plaintexts and ciphertexts is zero, capturing the intuition that ciphertexts should not "leak information" about plaintexts. A fundamental aspect of Shannon's approach is that it does not account for the computational difficulty of extracting information, impacting its practical significance. We note that this critique is not merely based on hindsight. The drawbacks of a purely information-theoretic approach were noted, by Turing, as early as 1950: "In many types of investigation, e..g., in the theory of information, it is a useful assumption that 'computation costs nothing'. It is important however not to let this assumption become a belief. In particular when one is considering brains and computers the assumption, and the theories based on it, are not applicable" [23]. Despite

[*] Department of Computer Science, University of Victoria,Victoria, BC, CANADA V8W 3P6, {mhaji,bmkapron}@uvic.ca. Work supported in part by NSERC Discovery Grant, "Foundational studies in privacy and security", and Simons Institute for the Theory of Computing program "Real analysis in computer science". Work of the second author was partly completed while a Visiting Member of the School of Mathematics at the Institute for Advanced Study. A preliminary version of this paper appeared in [11].

ISSN 1433-8092

Turing's warning, and the obvious practical drawbacks of Shannon's approach, it took more than thirty years before a rigorous definition of encryption security which accounts for computation was given by Goldwasser and Micali in [10], which proposes as a definition of security that "whatever is efficiently computable about the cleartext given the ciphertext, is also efficiently computable without the ciphertext." This notion, dubbed *semantic security*, may be viewed as a computational version of Shannon's definition, at least in an intuitive sense where *computational information* is identified with "whatever is efficiently computable about the plaintext". However, the information-theoretic approach uses a specific entropy-based notion of *mutual information*, and it is natural to ask whether one could formulate a notion of encryption security by first formulating a computational version of mutual information. In fact, such a question was posed, and answered by Yao [26, 27], who gave definitions of computational entropy and computational mutual information, based on the relationship between entropy and compressibility, and used this to characterize encryption security. In particular, an encryption scheme is secure if no efficient compression scheme in which the decoding function has access to an encryption of the corresponding message can achieve an expected length more than negligibly better than the optimal that can be achieved by any efficient scheme without such access. Yao [27] and Micali, Rackoff and Sloan [18] show that this notion is equivalent to semantic security.

This paper presents a new approach to characterizing encryption security via computational information. Rather than relying on the machinery of data compression, our approach uses a characterization of mutual information given by Kelly in [14], which considers the optimal rate of return for a gambler who has noisy inside information on the outcome of an event. When we take the very natural steps of replacing "noisy" with "encrypted" and considering a computationally bounded gambler, we are led immediately to a definition of encryption security which we dub *gambling security*. We then show (Theorem 4.9)

An encryption scheme is semantically secure iff it is gambling secure

While Yao's characterization of encryption security using computational information is not widely used, his introduction of a notion of computational entropy based on compression is one foundation of computational information theory. Another important contribution in this direction was made by Håstad et. al. in [12], which introduces the notion of *pseudoentropy*. Unfortunately, the relationship between various forms of computational entropy is not well understood. In some cases (e.g. with respect to conditional distributions), Yao's entropy and HILL entropy are not equivalent [13].[1] We may wonder what this means in the setting of encryption security. Do these different notions of computational information lead to different forms of encryption security? We show that this is not the case (Theorem 5.7)

An encryption scheme is semantically secure iff access to the ciphertext does not reduce the pseudoentropy of the plaintext.

The gambling framework allows us to consider gamblers with different utilities. We also consider security against gamblers who are trying to maximize a one-shot payout. Perhaps not surprisingly, in this setting we obtain a much tighter equivalence with semantic security. But this does raise an interesting question. Namely, what can we say about the relationship between security and assumptions about an adversary's rationality?

[1] The cited result considers versions based on min-entropy rather than Shannon entropy.

1.1 Related Work

As noted above, in [26], Yao introduced a notion of effective conditional entropy and effective mutual information. Among other applications, he used these notions to define computational security for encryption. Subsequently, in [27] and [18], it was shown that this definition of coincide with those (*semantic security* and *message indistinguishability*) introduced by Goldwasser and Micali [10]. An alternate approach to computational entropy is introduced in [12], but the notion introduced there is not directly applied to defining encryption security. The question of how these two notions of computational entropy are related was first considered in [1], but with respect to min-entropy, rather than Shannon entropy. A new notion related to HILL entropy, dubbed *metric entropy*, is also introduced here. A series of subsequent works, including [25, 13, 19, 22, 15] further explore the relationship between these notions of computational entropy. There are separations between these notions with respect to certain properties. For example, Krenn et al. [15, 16] show that, under some computational assumptions, a special case of the chain rule (which holds for, e.g., computational Yao and computational *unpredictability* entropy) fails with respect to HILL entropy. While there are numerous applications of these notions in cryptography (see, e.g., [20] for a survey) it does not seem that the relationship between computational entropy and encryption security has been considered other than in the initial works cited above. Finally, we note that our work may be viewed as complementary to recent work by Bellare et. al. [2], which considers a version of message indistinguishability and its relationship with entropy-based definitions of security in the information-theoretic setting. Dodis ([6], Lemma 2) states upper and lower bounds on mutual-information-based security in terms of message indistinguishability which are implicit in [2]. Our Lemmas 4.7 and 4.8 may be viewed as computational versions of Dodis' bounds.

2 Preliminaries

We assume standard facts and definitions about discrete probability spaces, but we will begin by clarifying our notational conventions and use of terminology. A *probability distribution on* a finite set $\mathcal{X}$ is specified by a *probability mass function* $X : \mathcal{X} \to [0,1]$ which satisfies $\sum_{x\in\mathcal{X}} X(x) = 1$. We will abuse terminology and use the term random variable as a synonym for distribution, so that for a distribution with mass function X we have $\Pr[X = x] = X(x)$. In general, we follow the convention of [9] regarding random variables, i.e., multiple occurrences of a variable in a probability expression denote multiple occurrences of a single sampled value. If X is distributed jointly with Y, we write $X|Y$ to denote the corresponding conditional distribution, and X, Y to denote the joint distribution. We write log and ln, respectively, for logarithm base 2 and base e.

Definition 2.1. Suppose that X and Y are jointly distributed random variables on $\mathcal{X}$ and $\mathcal{Y}$, respectively. The *entropy* $\mathrm{H}(X)$ of X is defined by

$$\mathrm{H}(X) = -\sum_{x} \Pr[X = x] \log \Pr[X = x].$$

The *conditional entropy* $\mathrm{H}(X|Y)$ of X *given* Y is defined by

$$\mathrm{H}(X|Y) = -\sum_{y,x} \Pr[Y = y \wedge X = x] \log \Pr[X = x|Y = y]$$

The *mutual information* $\mathrm{I}(X;Y)$ *between* X *and* Y is defined by

$$\mathrm{I}(X;Y) = \sum_{y,x} \Pr[Y = y \wedge X = x] \log \frac{\Pr[Y = y \wedge X = x]}{\Pr[Y = y] \Pr[X = x]}$$

Note that $\mathrm{I}(X;Y) = \mathrm{H}(X) - \mathrm{H}(X|Y)$. For a random variable X', the *KL divergence from* X *to* X' is defined by

$$\mathrm{KL}(X \| X') = \sum_{m} \Pr[X = x] \log \frac{\Pr[X = x]}{\Pr[X' = x]}$$

If X' jointly distributed with Y'. The *conditional KL divergence from* $X|Y$ *to* $X'|Y'$ is defined by

$$\mathrm{KL}(X|Y \| X'|Y') = \sum_{y,x} \Pr[Y = y \wedge X = x] \log \frac{\Pr[X = x|Y = y]}{\Pr[X' = x|Y' = y]}$$

We recall the following facts:

Proposition 2.2 (Chain rule for KL-divergence)**.**

$$\mathrm{KL}(Y, X \| Y', X') = \mathrm{KL}(Y \| Y') + \mathrm{KL}(X|Y \| X'|Y').$$

In particular $\mathrm{KL}(X|Y \| X'|Y) = \mathrm{KL}(Y, X \| Y, X')$.

Proposition 2.3 (Gibb's inequality)**.** $\mathrm{KL}(X \| X') \geq 0$*, with equality when* X' *is distributed identically to* X*.*

3 Proportional Betting with Noisy Inside Information

We begin by recalling the model of Kelly, proposed in his seminal paper [14]. Kelly considers the problem of maximizing the expected *rate* at which a gambler can accumulate wealth over repeated independent, identically-distributed plays of a game such as a coin flip, or horse race. In this scenario each play of the game results in an outcome from a fixed set of outcomes, according to an *a priori* fixed probability distribution known to the gambler. In particular, Kelly considers the advantage that an *eavesdropping gambler* who is given access to a channel providing a "noisy" version of the outcome of each event has over an *honest gambler*, who may only use *a priori* probabilities when placing bets. Kelly shows that the eavesdropping gambler's optimal strategy is proportional betting conditioned on the outcome observed on the noisy channel, and that the advantage of the best eavesdropping gambler over the best honest gambler is equal to the mutual information between the event and its noisy version.

Definition 3.1. Let X be a distribution over a set $\mathcal{X}$ of *outcomes*, i.e., the outcome of each play of the game is independently determined according to X. An *honest gambler* is given by a *betting function* $b : \mathcal{X} \to [0, 1]$ which satisfies $\sum_{x \in \mathcal{X}} b(x) = 1$.

The value $b(x)$ is the fraction of total wealth that the gambler bets on outcome x. Note that we are assuming that the gambler distributes all his wealth over the possible outcomes. The amount paid on a given outcome is determined by an odds function o. In particular, with an odds function o, and a betting function b, after outcome x the gambler's new wealth is $o(x)b(x)$ times his current wealth. In this paper, we will only consider odds functions which satisfy $\sum_{x \in \mathcal{X}} \frac{1}{o(x)} = 1$, in which case the assumption that the gambler bets all his wealth in each race is without loss of generality, because withholding can be simulated by spreading the withheld amount across outcomes in inverse proportion to o. While is it natural to consider more general odds functions, the analysis in this setting is much less tractable, and moreover it is not clear how to interpret such a setting from a security perspective.

Notation. A betting function $b : \mathcal{X} \to [0,1]$ may be viewed as the mass function of a distribution on $\mathcal{X}$. We will abuse notation somewhat and write b to also denote the random variable corresponding to the distribution with this mass function and also writing $b(X)$ for the value of b on x chosen randomly according to X.

Kelly considers a gambler who is trying to maximize the expected rate at which his wealth grows over a sequence of identically distributed independent random events. Asymptotically, this is equivalent to maximizing $\mathrm{E}[\log o(X)b(X)]$ (see [5], Theorem 6.1.1.) In this setting, for an honest gambler we have:

Proposition 3.2. *The maximum over all betting functions b of $\mathrm{E}[\log o(X)b(X)]$ is $\mathrm{E}[\log o(X)] - \mathrm{H}(X)$, and is achieved by b^* where $b^*(x) = \Pr[X = x]$.*

Note that $\mathrm{E}[\log o(X)]$ is the theoretical maximum, achieved by a "clairvoyant" gambler who always has all wealth placed on the winning outcome.

Proof. For any b we have

$$\mathrm{E}[\log o(X)b(X)] = \mathrm{E}[\log o(X)] + \mathrm{E}[\log b(X)] = \mathrm{E}[\log o(X)] - \mathrm{H}(X) - \mathrm{KL}(X\|b)$$

Now recall by Proposition 2.3 that $\mathrm{KL}(X\|b) \geq 0$, with equality when $b = X$. □

An *eavesdropping gambler* has access to the outcome of each race before betting takes place, but the access is noisy. This "noisy inside knowledge" is modeled by a random variable Y, jointly distributed with X.

Definition 3.3. An *eavesdropping gambler* is given by a *conditional betting function*, where $b(x|y)$ is the fraction of wealth bet on outcome x when y is observed.

We will write Y, b for the joint random variable induced by the conditional betting function b and distribution Y on observations, and $b|Y$ for the corresponding conditional random variable.

Definition 3.4. For any honest gambler b' the *advantage over b'* of an eavesdropping gambler b is equal to

$$\mathrm{E}[\log o(X)b(X|Y)] - \mathrm{E}[\log o(X)b'(X)],$$

The eavesdropper's *advantage* is its advantage over the best honest gambler b^*.

By the preceding Proposition, an eavesdropping gambler's advantage is equal to

$$\mathrm{E}[\log o(X)] + \mathrm{E}[\log b(X|Y)] - (\mathrm{E}[\log o(X)] - \mathrm{H}(X)) = \mathrm{H}(X) + \mathrm{E}[\log b(X|Y)]$$

Proposition 3.5. *An eavesdropping gambler's maximum advantage is $\mathrm{I}(X;Y)$ and is achieved by $\hat{b}$ where $\hat{b}(x|y) = \Pr[X = x|Y = y]$*

Proof. For any eavesdropping b we have

$$\begin{aligned}\mathrm{H}(X) + \mathrm{E}[\log b(X|Y)] &= \mathrm{H}(X) - \mathrm{H}(X|Y) - \mathrm{KL}(X|Y \| b|Y) \\ &= \mathrm{I}(X;Y) - \mathrm{KL}(X|Y \| b|Y)\end{aligned}$$

We now use the properties of KL-divergence to obtain the optimal strategy and value. □

Relation to information-theoretic security. While Kelly did not explicitly consider an encrypted channel, it is clear that with this approach, we obtain an alternate characterization of perfect secrecy. In particular suppose that Y is just $\mathcal{E}(K, X)$ where $(\mathcal{E}, \mathcal{D})$ is an encryption scheme (and K denotes a distribution over keys.) According to Shannon's definition, this encryption scheme has perfect secrecy exactly when $\mathrm{I}(X;Y) = 0$. In other words, any eavesdropping gambler using inside information encrypted by $\mathcal{E}$ has at best zero advantage. One advantage of this characterization of encryption security, as opposed to Shannon's characterization in [21] is the explicit introduction of an adversary (i.e. the eavesdropping gambler.) By considering resource bounded adversaries, we are led naturally to a version of Shannon security in the computational setting.

4 Computational Setting

We have seen that Kelly's model of gambling with inside information may be used to give a characterization of information-theoretic encryption security which is equivalent to Shannon's. Our goal now is to use Kelly's model to give a computational defintion of encryption security by considering computationally limited gamblers. We begin by reviewing some basic definitions regarding private-key encryption and security in the computational setting.

Definition 4.1. A *private-key encryption scheme* $\langle \mathcal{E}, \mathcal{D} \rangle$ is a probabilistic poly-time function ensemble $\langle \mathcal{E}_n, \mathcal{D}_n \rangle$ satisfying the following properties, for every n:

1. $\mathcal{E}_n : \{0,1\}^n \times \{0,1\}^{\ell(n)} \to \{0,1\}^{q(n)}$
2. $\mathcal{D}_n : \{0,1\}^n \times \{0,1\}^{q(n)} \to \{0,1\}^{\ell(n)}$
3. For any $k \in \{0,1\}^n$ and $m \in \{0,1\}^{\ell(n)}$, $\mathcal{D}_n(k, \mathcal{E}_n(k,m)) = m$,

where ℓ, q are poly-bounded functions such that $q(n) \geq \ell(n) \geq n$. The value n is the *security parameter* of the scheme.

Without loss of generality, we have dispensed with the specification of a *key generation* function. We may assume that keys are just uniformly generated random strings, as such strings could indeed be viewed as the randomness used in key generation. In what follows, we will write U_n to denote the uniform distribution over keys of length n. We will typically write M_n for an arbitrary distribution over messages of length $\ell(n)$.

In this paper, we limit our attention to *single message* security, that is, definitions of security in which an attacker has only has access to a single ciphertext c drawn from $\mathcal{E}_n(U_n, M_n)$. While it is possible to adapt some of our results to more comprehensive notions of security (e.g. CPA security) we will focus on the conceptual foundations of the notion of attacker success rather than attack models. We will consider the possibility of more "intrinsic" notions of multiple message security in Section 7. We will also only consider *non-uniform* definitions of security. That is, efficient adversaries will be modeled as poly-bounded families of circuits. While most of our results may be transferred to the uniform setting we will retain a non-uniform approach for the sake of conceptual clarity.

We now introduce a notion of encryption security using a resource-bounded formulation of Kelly's model. In this setting, we consider betting functions which are computed by poly-size families of circuits (which are defined in the obvious way.)

Definition 4.2 (Gambling Security). An encryption scheme $(\mathcal{E}, \mathcal{D})$ is *gambling secure* if for every k, every distribution ensemble $\{M_n\}$, where M_n is a distribution over $\{0,1\}^{\ell(n)}$, every poly-size family of betting circuits $\{b_n\}$, and all sufficiently large n

$$\mathrm{H}(M_n) + \mathrm{E}[\log b_n(M_n|\mathcal{E}_n(U_n, M_n))] \leq \frac{1}{n^k}.$$

Note: Our definition of gambling security measures the advantage of a computationally bounded eavesdropping gambler versus the *best* honest gambler. This is without loss of generality for our results, unless we do not admit honest gamblers whose complexity is polynomial in that of the eavesdropping gambler.

We would like to compare gambling security to the more familiar notion of *message indistinguishability*. To do so, we first recall some definitions.

Definition 4.3. A *distinguisher* is a function $D : \{0,1\}^\ell \to \{0,1\}$. If X, X' are distributions defined on $\{0,1\}^\ell$, the *advantage of D in distinguishing between X and X'*, denoted $\mathrm{Adv}_D(X, X')$, is defined by

$$\mathrm{Adv}_D(X, X') = \Pr[D(X) = 1] - \Pr[D(X') = 1]$$

We will also consider *generalized distinguishers* which take values in $[0,1]$.[2] For such a D, Adv_D is defined by

$$\mathrm{Adv}_D(X, X') = \mathrm{E}[D(X)] - \mathrm{E}[D(X')]$$

We say that a (generalized) distinguisher has *size t* if it is computed by a circuit of size t.

In the case of size-bounded generalized distinguishers, the circuit D outputs the binary representation of a (rational) value in $[0,1]$. Note that the size restriction means that the actual range of D's output is contained in $\{0\} \cup [2^{-t}, 1]$. Using standard techniques, with polynomial overhead and at most d bits of randomness, we may transform a generalized distinguisher D into a distinguisher D' such that $\mathrm{E}[D(X)] = \mathrm{E}[D'(X)] = \Pr[D'(X) = 1]$, so that $\mathrm{Adv}_{D'}(X, X') = \mathrm{Adv}_D(X, X')$ for any X and X'. In particular, on input x, D' flips t coins and interprets the result as a value δ in $\{0\} \cup [2^{-t}, 1)$. If $\delta < D(x)$ then $D'(x)$ returns 1. Otherwise it returns 0. We also recall the following

Proposition 4.4. *For any distinguisher D and distributions X_0, X_1,*

$$\mathrm{Adv}_D(X_0, X_1) \leq \epsilon \quad \textit{iff} \quad \Pr[D(X_z) = z] \leq \frac{1}{2} + \frac{\epsilon}{2}$$

where z is selected uniformly at random from $\{0,1\}$.

[2] We use the same terminology here as [24], but note that in their setting, generalized distinguishers take values in $\mathbb{R}^+$.

Definition 4.5. Let ℓ be a poly-bounded function. A *distribution ensemble* is a sequence $\mathbf{X} = \{X_n\}$ of distributions, where X_n is a distribution on $\{0,1\}^{\ell(n)}$. A poly-*size family of distinguishing circuits* is a sequence $\{D_n\}$ of circuits, where $D_n : \{0,1\}^{\ell(n)} \to \{0,1\}$ has size $n^{O(1)}$. Distribution ensembles $\{X_n\}$ and $\{Y_n\}$ are *computationally indistinguishable* if for every k, poly-size family $\{D_n\}$ of distinguishing circuits, and sufficiently large n,

$$\mathrm{Adv}_{D_n}(X_n, Y_n) \le \frac{1}{n^k}$$

We recall the following standard definition of encryption security (equivalent to semantic security [10]):

Definition 4.6 (Message Indistinguishability). An encryption scheme $(\mathcal{E}, \mathcal{D})$ has *indistinguishable messages* if for any k and poly-size family of distinguishing circuits $\{D_n\}$, for all sufficiently large n and pair of messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$

$$\mathrm{Adv}_{D_n}(\mathcal{E}_n(U_n, m_1)), \mathcal{E}_n(U_n, m_0)) \le \frac{1}{n^k}$$

Let $\langle \mathcal{E}, \mathcal{D} \rangle$ be an encryption scheme. The equivalence of message indistinguishability and gambling security for $\langle \mathcal{E}, \mathcal{D} \rangle$ is established using the following two lemmas

Lemma 4.7. *For any n, $0 < \delta < \frac{1}{2}$, size t distinguishing circuit D, and messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$, such that $\mathrm{Adv}_D(\mathcal{E}_n(U_n, m_0), \mathcal{E}_n(U_n, m_1)) > 2\delta$, there is a size poly$(t, \log(1/\delta), \ell(n), q(n))$ betting circuit b and a distribution M on $\{0,1\}^{\ell(n)}$ such that*

$$H(M) + \mathrm{E}[\log b(M|\mathcal{E}_n(U_n, M))] > \frac{2}{\ln 2}\delta^2$$

Proof. By assumption and Proposition 4.4, for uniformly chosen $z \in \{0,1\}$,

$$\Pr[D(\mathcal{E}_n(U_n, m_z)) = z] > \tfrac{1}{2} + \delta$$

Define b with size $t + \mathrm{poly}(\log(1/\delta), \ell(n), q(n))$ as follows

$$b(m|c) = \begin{cases} 0 & \text{if } m \notin \{m_0, m_1\}; \\ \frac{1}{2} + \delta & \text{if } m = m_z \text{ and } D(c) = z; \\ \frac{1}{2} - \delta & \text{otherwise.} \end{cases}$$

Let M be the distribution which assigns m_0 and m_1 probability $\frac{1}{2}$ and all other messages probability 0. Then $\mathrm{H}(M) = 1$, while

$$\mathrm{E}[\log b(M|\mathcal{E}_n(U_n, M))] > (\tfrac{1}{2} + \delta)\log(\tfrac{1}{2} + \delta) + (\tfrac{1}{2} - \delta)\log(\tfrac{1}{2} - \delta)$$

which is just $-\mathrm{h}(\frac{1}{2} + \delta)$, where h is the binary entropy function. So it suffices to show that for $0 \le \delta < \frac{1}{2}$

$$1 - \mathrm{h}(\tfrac{1}{2} + \delta) \ge \tfrac{2}{\ln 2}\delta^2$$

Using a Taylor series expansion we have

$$\begin{aligned} 1 - \mathrm{h}(\tfrac{1}{2} + \delta) &= 1 - \left(1 - \frac{1}{2\ln 2}\sum_{t=1}^{\infty} \frac{(2\delta)^{2t}}{t(2t-1)}\right) \\ &= \frac{1}{2\ln 2}\sum_{t=1}^{\infty} \frac{(2\delta)^{2t}}{t(2t-1)} \ge \frac{1}{2\ln 2}4\delta^2 = \frac{2}{\ln 2}\delta^2. \end{aligned}$$

□

Lemma 4.8. *For any n, any $\delta \geq 0$, any size t betting circuit b and distribution M on $\{0,1\}^{\ell(n)}$, such that $H(M) + \mathrm{E}[\log b(M|\mathcal{E}_n(U_n, M))] > \delta$, there is a size $poly(t, \log(1/\delta), \ell(n), q(n))$ distinguishing circuit D and and messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$ such that*

$$\mathrm{Adv}_D(\mathcal{E}_n(U_n, m_0), \mathcal{E}_n(U_n, m_1)) > \frac{\delta}{2t}$$

Proof. By Proposition 3.2, we have $\mathrm{E}[\log b'(M)] \leq -\mathrm{H}(M)$ for any betting function b', so that

$$\mathrm{E}[\log b(M|\mathcal{E}_n(U_n, M))] - \max_{b'} \mathrm{E}[\log b'(M)] > \delta$$

In particular, define b' by $b'(m) = b(m|\mathcal{E}_n(U_n, m_0))$ for some fixed message m_0 (note that b''s complexity is polynomial in b's, assuming $\mathcal{E}$ is poly-time.) Then we have $\mathrm{E}[\log b(M|\mathcal{E}_n(U_n, M))] - \mathrm{E}[\log b(M|\mathcal{E}_n(U_n, m_0))] > \delta$. By averaging, we conclude that that there must be some fixed m_1 for which $\mathrm{E}[\log b(m_1|\mathcal{E}_n(U_n, m_1))] - \mathrm{E}[\log b(m_1|\mathcal{E}_n(U_n, m_0))] > \delta$. Define D' as follows

$$D'(c) = \begin{cases} \frac{1}{t}(\log b(m_1, c) + t) & \text{if } b(m_1|c) > 0; \\ 0 & \text{otherwise.} \end{cases}$$

Then $D'(c) \in [0, 1]$, and

$$\mathrm{Adv}_{D'}(\mathcal{E}_n(U_n, m_1), \mathcal{E}_n(U_n, m_0)) > \frac{\delta}{t}.$$

As shown in [24] (Theorem 3.22), D' may be approximated using a Taylor series to precision $\frac{\delta}{2}$ by a circuit D of size $\mathrm{poly}(t, \log(1/\delta), \ell(n), q(n))$ such that

$$\mathrm{Adv}_D(\mathcal{E}_n(U_n, m_1), \mathcal{E}_n(U_n, m_0)) > \frac{\delta - \delta/2}{t} = \frac{\delta}{2t}$$

As discussed previously, we may assume that, with polynomial overhead, instead of outputting a value $p \in [0, 1]$, D outputs 1 with probability p and 0 with probability $1 - p$, so that D is a distinguisher. □

Theorem 4.9. *$(\mathcal{E}, \mathcal{D})$ is gambling secure iff it has indistinguishable messages.*

Proof. Supose that $(\mathcal{E}, \mathcal{D})$ does not have indistinguishable messages. So, for some k, and poly-size family $\{D_n\}$ of distinguishing circuits, it is the case that for infinitely many n there is a pair of messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$ such that

$$\mathrm{Adv}_{D_n}(\mathcal{E}_n(U_n, m_1)), \mathcal{E}_n(U_n, m_0)) > \frac{1}{n^k}$$

Let $n_1, n_2, \ldots$ be the values of n for which this holds. Define the distribution ensemble $\{M_n\}$ where, for $n = n_i$, M_n is the distribution guaranteed by Lemma 4.7, and otherwise is arbitrary, say the uniform distribution, and let $k' = 2k + 1$. Then, by Lemma 4.7, there is a poly-size family of betting circuits $\{b_n\}$, such that for infinitely many n

$$\mathrm{H}(M_n) + \mathrm{E}[\log b_n(M_n|\mathcal{E}_n(U_n, M_n))] > \frac{1}{(2\ln 2)n^{2k}} \geq \frac{1}{n^{k'}}$$

So $(\mathcal{E}, \mathcal{D})$ is not gambling secure.

Now suppose that $(\mathcal{E}, \mathcal{D})$ is not gambling secure. So there exists a k, a distribution ensemble $\{M_n\}$, where M_n is a distribution over $\{0,1\}^{\ell(n)}$, and a poly-size family of betting circuits $\{b_n\}$, such that for infinitely many n

$$\mathrm{H}(M_n) + \mathrm{E}[\log b_n(M_n|\mathcal{E}_n(U_n, M_n))] > \frac{1}{n^k}.$$

Suppose b_n has size n^j. Let $k' = k + j + 1$. Then by Lemma 4.8, there is a poly-size family of distinguishing circuits $\{D_n\}$, such that for infinitely many n there is a pair of messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$ such that

$$\mathrm{Adv}_{D_n}(\mathcal{E}_n(U_n, m_1)), \mathcal{E}_n(U_n, m_0)) > \frac{1}{2n^{k+j}} \geq \frac{1}{n^{k'}}$$

So, $(\mathcal{E}, \mathcal{D})$ does not have indistinguishable messages. □

5 A Characterization Based on Pseudoentropy

The notion of *KL-hardness*, introduced by Vadhan and Zheng in [24], characterizes the difficulty of approximating a distribution with respect to KL-divergence. In the nonuniform setting, they show that a conditional distribution is KL-hard if and only if it has high conditional pseudoentropy.[3] KL-hardness is closely related to gambling security: as we have already seen in the information-theoretic setting, an eavesdropping gambler b maximizes its advantage by minimizing $\mathrm{KL}(M|C\|b|C)$. Moving to the computational setting, things are less straightforward. In particular, the definition of KL-hardness given in [24] depends on the notion of a *KL-predictor*, which does not correspond exactly to a betting function. A KL-predictor is obtained by normalizing a *measure*, which is a function from the space of outcomes to $(0, +\infty)$. Nevertheless, the results of [24] and the preceding section suggest that we should be able to give a characterization of encryption security based on conditional pseudoentropy. We will do this directly, relying heavily on techniques introduced in [24]. We could also first establish an equivalence between KL-hardness for gambling functions and KL-hardness for normalized measures and then appeal directly to the main result of Vadhan and Zheng; this approach is discussed in Appendix A.

Definition 5.1. Suppose $\mathbf{X} = \{X_n\}$, $\mathbf{Y} = \{Y_n\}$ are distribution ensembles. $\mathbf{X}$ has *conditional pseudoentropy at least k given* $\mathbf{Y}$, written $\tilde{H}(\mathbf{X}|\mathbf{Y}) \geq k$, if there is a distribution ensemble $\{X'_n\}$ such that $\{(Y_n, X_n)\}\}$ and $\{(Y_n, X'_n)\}$ are computationally indistinguishable and for all c and sufficiently large n $\mathrm{H}(X'_n|Y_n) \geq k - \frac{1}{n^c}$.

We recall that according to Shannon [21], an encryption scheme is perfectly secure if $\mathrm{I}(M; C) = \mathrm{H}(M) - \mathrm{H}(M|C) = 0$. In the computational setting, it seems natural to quantify the degree of security in terms of $\mathrm{H}(M) - \tilde{\mathrm{H}}(M|C)$. Before showing that the corresponding security notion corresponds to message indistinguishability, we must consider a general relationship between conditional betting functions and distinguishers, required in the proof of Lemma 5.5 below. We will limit our attention to betting functions which are nonzero, taking values in $[2^{-t}, 1]$ (any nonzero size t betting function will take values in this range.) Any such function b determines a generalized distinguisher D_b, which we now define.

[3] Results are obtained in the uniform setting as well, but only for joint distributions of the form X, B over $\{0,1\}^n \times [q]$, where q is $\mathrm{poly}(n)$.

Definition 5.2. Suppose b is a conditional betting function taking values in $[2^{-t}, 1]$. Define the generalized distinguisher D_b, taking values in $[0, 1]$, as follows:

$$D_b(y, x) = \frac{1}{t}(\log b(x|y) + t)$$

Note that $\log b(x|y) = tD_b(y, x) - t$. We have the following (information theoretic) relationships between b and D_b.

Lemma 5.3. *Suppose b is a conditional betting function taking values in $[2^{-t}, 1]$. Then for any distribution X on $\mathcal{X}$ and Y on $\mathcal{Y}$,*

$$\mathrm{KL}((Y, X)\|(Y, b)) = \mathrm{H}(b|Y) - \mathrm{H}(X|Y) - t\mathrm{Adv}_{D_b}((Y, X), (Y, b))$$

Proof. This is just a reformulation of [24], Lemma 3.13 in our setting.

$$\begin{aligned}
&\mathrm{KL}((Y, X)\|(Y, b))\\
&= \mathrm{E}_Y\left[\sum_x X(x|Y)\log\frac{X(x|Y)}{b(x|Y)}\right]\\
&= \mathrm{H}(b|Y) - \mathrm{H}(X|Y) + \mathrm{E}_Y\left[\sum_x (X(x|Y) - b(x|Y))\log\frac{1}{b(x|y)}\right]\\
&= \mathrm{H}(b|Y) - \mathrm{H}(X|Y) + \mathrm{E}_Y\left[\sum_x (X(x|Y) - b(x|Y)(t - tD_b(Y, x))\right]\\
&= \mathrm{H}(b|Y) - \mathrm{H}(X|Y) + t\mathrm{E}_Y\left[(1-1) - \sum_x D_b(Y, x)(X(x|Y) - b(x|Y))\right]\\
&= \mathrm{H}(b|Y) - \mathrm{H}(X|Y) - t\mathrm{Adv}_{D_b}((Y, X), (Y, b))
\end{aligned}$$

□

Lemma 5.4. *Suppose b is a conditional betting function taking values in $[2^{-t}, 1]$. Then for all distributions X, X' on $\mathcal{X}$ and Y on $\mathcal{Y}$,*

$$\mathrm{Adv}_{D_b}((Y, X), (Y, X')) \geq (\mathrm{H}(X'|Y) + \mathrm{E}[\log b(X|Y)])/t$$

Proof. By Lemma 5.3 we have

$$\mathrm{H}(b|Y) - \mathrm{H}(X'|Y) - t\mathrm{Adv}_{D_b}((Y, X'), (Y, b)) = \mathrm{KL}((Y, X')\|(Y, b))$$

since $\mathrm{KL}((Y, X')\|(Y, b)) \geq 0$, it follows that

$$\mathrm{Adv}_{D_b}((Y, X'), (Y, b)) \leq \frac{\mathrm{H}(b|Y) - \mathrm{H}(X'|Y)}{t} \tag{†}$$

Applying Lemma 5.3 again, we obtain

$$\mathrm{H}(b|Y) - \mathrm{H}(X|Y) - t\mathrm{Adv}_{D_b}((Y, X), (Y, b)) = \mathrm{KL}((Y, X)\|(Y, b))$$

But

$$\begin{aligned}\mathrm{KL}((Y,X)\|(Y,b)) &= \mathrm{I}(X;Y) - \mathrm{H}(X) - \mathrm{E}[\log b(X|Y)] \\ &= -\mathrm{H}(X|Y) - \mathrm{E}[\log b(X|Y)]\end{aligned}$$

so that

$$\mathrm{Adv}_{D_b}((Y,X),(Y,b)) = \frac{\mathrm{H}(b|Y) + \mathrm{E}[\log b(X|Y)]}{t} \tag{††}$$

But then

$$\begin{aligned}\mathrm{Adv}_{D_b}((Y,X),(Y,X')) &= \mathrm{Adv}_{D_b}((Y,X),(Y,b)) - \mathrm{Adv}_{D_b}((Y,X'),(Y,b)) \\ &\geq (\mathrm{H}(X'|Y) + \mathrm{E}[\log b(X|Y)])/t \qquad \text{by (†) and (††)}\end{aligned}$$

□

Lemma 5.5. *For any n, $0 \leq \delta < \frac{1}{4}$, size t distinguishing circuit D, and messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$, such that $\mathrm{Adv}_D(\mathcal{E}_n(U_n,m_0),\mathcal{E}_n(U_n,m_1)) > 2\delta$, there is a distinguishing circuit D_b of size $poly(t, \log(1/\delta), \ell(n), q(n))$ and distribution M on $\{0,1\}^{\ell(n)}$ such that for any $\gamma \geq 0$ and any M' such that $H(M'|\mathcal{E}_n(U_n,M)) \geq H(M) - \gamma$*

$$\mathrm{Adv}_{D_b}((\mathcal{E}_n(U_n,M),M),(\mathcal{E}_n(U_n,M),M')) > \frac{\delta^2}{t\ln 2} - \frac{\gamma}{t}$$

Proof. By Proposition 4.4, for uniformly chosen $z \in \{0,1\}$,

$$\Pr[D(\mathcal{E}_n(U_n,m_z)) = z] > \tfrac{1}{2} + \delta$$

Let M be the distribution which assigns m_0 and m_1 probability $\frac{1}{2}$ and all other messages probability 0. We now define a betting function b such that D_b is a distinguisher between M and any M' such that $\mathrm{H}(M'|\mathcal{E}_n(U_n,M)) > \mathrm{H}(M) - \gamma$. For any $\sigma > 0$, we define b as follows:

$$b(m|c) = \begin{cases} \frac{\sigma}{2^{\ell(n)}-2} & \text{if } m \notin \{m_0,m_1\}; \\ 1/2 + \delta & \text{if } m = m_z \text{ and } D(c) = z; \\ 1/2 - \delta - \sigma & \text{otherwise.} \end{cases}$$

Let $D_b(y,x) = \frac{1}{t}(\log b(x|y) + t)$ be the associated distinguisher given in Definition 5.2. Let C denote $\mathcal{E}_n(U_n,M)$. Consider any M' for which $\mathrm{H}(M'|C) \geq \mathrm{H}(M) - \gamma$. By Lemma 5.4

$$\begin{aligned}\mathrm{Adv}_{D_b}((C,M),(C,M')) &\geq (\mathrm{H}(M'|C) + \mathrm{E}[\log b(M|C)])/t \\ &\geq (\mathrm{H}(M) + \mathrm{E}[\log b(M|C)] - \gamma)/t\end{aligned}$$

Now $\mathrm{H}(M) = 1$, while we have, assuming $\delta < \frac{1}{4}$ and $\sigma < \frac{2}{9}$,

$$\begin{aligned}\mathrm{E}[\log b(M|\mathcal{E}_n(U_n,M))] &> (\tfrac{1}{2}+\delta)\log(\tfrac{1}{2}+\delta) + (\tfrac{1}{2}-\delta)\log(\tfrac{1}{2}-\delta-\sigma) \\ &= -\mathrm{h}(\tfrac{1}{2}+\delta) - \tfrac{1}{\ln 2}\textstyle\sum_{j=1}^{\infty} \tfrac{2^{j-1}\delta^j}{j(1-2\gamma)^{j-1}} \\ &\geq -\mathrm{h}(\tfrac{1}{2}+\delta) - \tfrac{1}{4\ln 2}\textstyle\sum_{s=1}^{\infty} 4^j\sigma^j \\ &= -\mathrm{h}(\tfrac{1}{2}+\delta) - \tfrac{4\sigma}{(4\ln 2)(1-4\sigma)} \\ &\geq -\mathrm{h}(\tfrac{1}{2}+\delta) - \tfrac{2}{\ln 2}\sigma\end{aligned}$$

As demonstrated previously, $1 - \mathrm{h}(\frac{1}{2} + \delta) \geq \frac{2}{\ln 2}\delta^2$, so that, setting $\sigma = \delta^2/2$ we have

$$\mathrm{H}(M) + \mathrm{E}[\log b(M|\mathcal{E}_n(U_n, M))] \geq \tfrac{\delta^2}{\ln 2}$$

and

$$\mathrm{Adv}_{D_b}((\mathcal{E}_n(U_n, M), M), (\mathcal{E}_n(U_n, M), M')) > \tfrac{\delta^2}{t \ln 2} - \tfrac{\gamma}{t}$$

Finally, we note that D_b has size $poly(t, \log(1/\delta), \ell(n), q(n))$ □

Lemma 5.6. *For any n, $\delta \geq 0$, and distribution M on $\{0,1\}^{\ell(n)}$ which has the property that for any distribution M' on $\{0,1\}^{\ell(n)}$ with $\mathrm{H}(M'|\mathcal{E}_n(U_n, M)) \geq \mathrm{H}(M)$, there is a size t distinguishing circuit D such that*

$$\Pr[D(\mathcal{E}_n(U_n, M), M')) = 1] - \Pr[D(\mathcal{E}_n(U_n, M), M) = 1] > \delta,$$

there is a size t distinguisher D' and messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$ such that

$$\mathrm{Adv}_{D'}(\mathcal{E}_n(U_n, m_0), \mathcal{E}_n(U_n, m_1)) > \delta.$$

Proof. Take M' which is independent of M but identically distributed. Then

$$\mathrm{H}(M'|\mathcal{E}_n(U_n, M)) = \mathrm{H}(M') = \mathrm{H}(M)$$

and so by assumption there is a size t distinguisher D such that

$$\Pr[D(\mathcal{E}_n(U_n, M), M')) = 1] - \Pr[D(\mathcal{E}_n(U_n, M), M) = 1] > \delta$$

But then there are messages m_0, m_1 such that

$$\Pr[D(\mathcal{E}_n(U_n, m_0), m_1) = 1] - \Pr[D(\mathcal{E}_n(U_n, m_1), m_1) = 1] > \delta \qquad (\dagger)$$

Defining $D'(c) = D(c, m_1)$ completes the proof. To obtain the required m_0, m_1, we apply an averaging argument twice. The first application allows us to conclude that

$$\Pr[D(\mathcal{E}_n(U_n, M), m_1) = 1] - \Pr[D(\mathcal{E}_n(U_n, m_1), m_1) = 1] > \delta.$$

We then average again to obtain (†). □

Using the preceding lemmas, we are now able to conclude

Theorem 5.7. *$(\mathcal{E}, \mathcal{D})$ has indistinguishable messages iff for any message distribution ensemble $\mathbf{M} = \{M_n\}$, $\tilde{\mathrm{H}}(\mathbf{M}|\mathbf{C}) \geq \mathrm{H}(\mathbf{M})$, where $\mathbf{C} = \{C_n\}$ is the distribution ensemble such that for each n, $C_n = \mathcal{E}_n(U_n, M_n)$.*

Proof. Supose that $(\mathcal{E}, \mathcal{D})$ does not have indistinguishable messages. So, for some k, and poly-size family $\{D_n\}$ of distinguishing circuits, it is the case that for infinitely many n there is a pair of messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$ such that

$$\mathrm{Adv}_{D_n}(\mathcal{E}_n(U_n, m_1)), \mathcal{E}_n(U_n, m_0)) > \frac{1}{n^k}$$

Let $n_1, n_2, \ldots$ be the values of n for which this holds. Define the distribution ensemble $\mathbf{M} = \{M_n\}$ where, for $n = n_i$, M_n is the distribution guaranteed by Lemma 5.5, and otherwise is the uniform distribution, and let $\{D'_n\}$ be the poly-sized circuit family, where, for $n = n_i$, D'_n is the ciruit D_b provided by Lemma 5.5 and otherwise is some default circuit, say a circuit for the constant zero function. Suppose D_n has size n^j, and consider any distribution ensemble $\{M'_n\}$ such that for all sufficiently large n, $\mathrm{H}(M'_n|\mathcal{E}_n(U_n, M_n)) \geq \mathrm{H}(M_n) - \frac{1}{n^c}$. Take $c = 2k+1$, and define $k' = 2k+1+j$. By Lemma 5.5 it follows that, for infinitely many n,

$$\mathrm{Adv}_{D'_n}((\mathcal{E}_n(U_n, M), M), (\mathcal{E}_n(U_n, M), M')) > \frac{1}{n^{k'}},$$

and so it follows that $\tilde{\mathrm{H}}(\mathbf{M}|\mathbf{C}) < \mathrm{H}(\mathbf{M})$.

Now suppose that for some message ensemble $\mathbf{M}$, $\tilde{\mathrm{H}}(\mathbf{M}|\mathbf{C}) < \mathrm{H}(\mathbf{M})$. It is then the case that for any distribution ensemble $\{M'_n\}$ such that $\mathrm{H}(M'_n|\mathcal{E}_n(U_n, M_n)) > \mathrm{H}(M_n)$, there is a family $\{D_n\}$ of distinguishing circuits and a constant k such that for infinitely many n,

$$\Pr[D_n(\mathcal{E}_n(U_n, M_n), M'_n) = 1] - \Pr[D_n(\mathcal{E}_n(U_n, M), M_n) = 1] > \frac{1}{n^k}.$$

Then by Lemma 5.6, there is a poly-size family of distinguishing circuits $\{D'_n\}$, such that for infinitely many n there is a pair of messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$ such that

$$\mathrm{Adv}_{D'_n}(\mathcal{E}_n(U_n, m_1)), \mathcal{E}_n(U_n, m_0)) > \frac{1}{n^k}$$

So, $(\mathcal{E}, \mathcal{D})$ does not have indistinguishable messages. □

6 Risk-Neutral Adversaries

We now revisit the model proposed by Kelly. Kelly's gambler may be viewed as trying to maximize the rate of return over repeated plays or, alternately, as just having a logarithmic utility for total wealth. What happens if we consider gamblers with different utility functions? We will now consider the case of linear utility. Such gamblers are typically referred to as being *risk-neutral*. Clearly, in this case, if no inside information is available the optimal strategy is to bet everything on the outcome which gives the maximum expected payout. An important difference in this setting is that the odds now make a difference, and the advantage of an eavesdropping gambler over the best honest gambler will be

$$\mathrm{E}[o(X)b(X|Y)] - \max_{x \in \mathcal{X}}(o(x)\Pr[X = x])$$

While in a more realistic setting we would want to take the odds function into account, as a first step we just assume constant odds, say $o(x) = \frac{1}{|\mathcal{X}|}$ for all x. In this way we remove consideration of odds from the gambler's strategy, leading to the following

Definition 6.1 (Risk-Neutral Gambling Security). An encryption scheme $(\mathcal{E}, \mathcal{D})$ is *gambling secure against risk-neutral adversaries* if for every k, every distribution ensemble $\{M_n\}$, where M_n is a distribution over $\{0,1\}^{\ell(n)}$, and every poly-size family of betting circuits $\{b_n\}$, for all sufficiently large n

$$\mathrm{E}[b_n(M_n|\mathcal{E}_n(U_n, M_n))] - \max_{m \in \{0,1\}^{\ell(n)}} \Pr[M_n = m] \leq \frac{1}{n^k}.$$

Lemma 6.2. *For any n, $0 \leq \delta < \frac{1}{2}$, size t distinguishing circuit D, and messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$, such that $\mathrm{Adv}_D(\mathcal{E}_n(U_n, m_0), \mathcal{E}_n(U_n, m_1)) > 2\delta$, there is a size poly$(t, \log(1/\delta), \ell(n))$ betting circuit b and a distribution M on $\{0,1\}^{\ell(n)}$ such that*

$$\mathrm{E}[\log b(M|\mathcal{E}_n(U_n, M))] - \max_{m \in \{0,1\}^{\ell(n)}} \Pr[M = m] > \delta$$

Proof. By Proposition 4.4, for uniformly chosen $z \in \{0,1\}$

$$\Pr[D(\mathcal{E}_n(U_n, m_z)) = z] > \tfrac{1}{2} + \delta$$

Define b with size $t + \mathrm{poly}(\ell(n), q(n))$ as follows

$$b(m|c) = \begin{cases} 1 & \text{if } m = m_z \text{ and } D(c) = z; \\ 0 & \text{otherwise.} \end{cases}$$

Let M be the distribution which assigns m_0 and m_1 probability $\frac{1}{2}$ and all other messages probability 0. Clearly,

$$\mathrm{E}[b(M|\mathcal{E}_n(U_n, M))] > (\tfrac{1}{2} + \delta) \cdot 1 + (\tfrac{1}{2} - \delta) \cdot 0 = \tfrac{1}{2} + \delta$$

The result follows by observing that $\max_{m \in \{0,1\}^{\ell(n)}} \Pr[M_n = m] = \frac{1}{2}$. □

Lemma 6.3. *For any n, any $\delta \geq 0$, any size t betting circuit b and distribution M on $\{0,1\}^{\ell(n)}$, such that $\mathrm{E}[b(M|\mathcal{E}_n(U_n, M))] - \max_{m \in \{0,1\}^{\ell(n)}} \Pr[M_n = m] > \delta$, there is a size poly$(t, \log(1/\delta), \ell(n), q(n))$ circuit D and messages $m_0, m_1 \in \{0,1\}^{\ell(n)}$ such that $\mathrm{Adv}_D(\mathcal{E}_n(U_n, m_0), \mathcal{E}_n(U_n, m_1)) > \delta$.*

Proof. As we have argued above, for any honest b', $\mathrm{E}[b(M_n|\mathcal{E}_n(U_n, M_n))] - \mathrm{E}[b'(M_n)] > \delta$. In particular, define b' by $b'(m) = b(m, \mathcal{E}_n(U_n, m_0))$ for some fixed message m_0. Then $\mathrm{E}[b(M_n|\mathcal{E}_n(U_n, M))] - \mathrm{E}[b(M_n|\mathcal{E}_n(U_n, m_0))] > \delta$, and so there is some m_1 such that

$$\mathrm{E}[b(m_1|\mathcal{E}_n(U_n, m_1))] - \mathrm{E}[b(m_1|\mathcal{E}_n(U_n, m_0))] > \delta$$

Define the generalized distinguisher D' by $D'(c) = b(m_1|c)$ and use D' to obtain a distinguisher D for which

$$\mathrm{Adv}_D(\mathcal{E}_n(U_n, m_0), \mathcal{E}_n(U_n, m_1)) > \delta$$ □

Theorem 6.4. *$(\mathcal{E}, \mathcal{D})$ is gambling secure against risk-neutral adversaries iff it has indistinguishable messages.*

We omit the proof, which is similar to that for Theorem 4.9. We also note that combining Lemmas 4.7 and 6.3 we obtain the following

Corollary 6.5. *For any n, any $\delta \geq 0$, any size t betting circuit b and distribution M on $\{0,1\}^{\ell(n)}$, such that $\mathrm{E}[b(M|\mathcal{E}_n(U_n, M))] - \max_{m \in \{0,1\}^{\ell(n)}} \Pr[M_n = m] > \delta$, there is a betting circuit b' of size poly$(t, \log(1/\delta), \ell(n), q(n))$ and a distribution M on $\{0,1\}^{\ell(n)}$ such that*

$$H(M) + \mathrm{E}[\log b'(M|\mathcal{E}_n(U_n, M))] > \frac{1}{2 \ln 2} \delta^2$$

In other words, an encryption scheme which achieves ϵ-gambling security also achieves $\sqrt{\epsilon}$-gambling security against risk-neutral gamblers.[4]

[4] The fact that the loss is quadratic is not surprising, as a risk-neutral gambler is optimizing the arithmetic mean of her returns, while a Kelly-style gambler is doing the same for the geometric mean. We note that the idea of optimizing the geometric mean of returns has a long history, dating at least to the 18th Century work of D. Bernoulli [3].

7 Conclusions and Future Work

We have revisited Yao'a program of characterizing encryption security via computational information, providing two new equivalent characterizations based on different approaches to computational entropy. In some sense this is more of a contribution to computational information theory than to cryptography, as we have shown that, at least in the setting of encryption security, various notions coincide. This of course raises the question of how these notions are related in more general settings. We now have another notion of computational entropy, based on Kelly's model, although this is closely related to Vadhan and Zheng's notion of KL-hardness, which in turn is closely related to pseudoentropy. Indeed, in a general setting, KL-hardness and pseudoentropy coincide for nonuniform adversaries ([24], Corollary 3.10.) There are still numerous open questions regarding the relationship Yao and HILL entropy; a broader view involving notions such as KL-hardness and gambling entropy may be useful here.

In the information-theoretic setting, gambling and data compression are equivalent. In [5], Section 6.5, a reduction from compression to gambling is given, using the gambling function to construct a cumulative distribution function which is then used in an arithmetic coding scheme, but this reduction is not efficient. We conjecture that under an appropriate complexity-theoretic assumption, no such efficient reduction is possible.

Our results only concern single-message security. We could easily give a version of CPA security, by considering the usual CPA-game, but replacing the challenge phase with one in which the adversary is a gambler rather than a distinguisher. On the other hand, Kelly's model suggests forms of multiple-message security which are ostensibly weaker than CPA, but stronger than standard multiple-message security. In particular, we could consider a situation in which the same key is used to encrypt the results of multiple races in an on-line fashion, and where the gambler is able to use information about his success in each round to place future bets. This is very similar to the setting of on-line prediction (see, e.g. [4]). We would like to consider adversaries performing this sort of on-line prediction task, or in the setting of on-line game playing as introduced in [7].

Our results imply that ϵ-gambling security implies $\sqrt{\epsilon}$-gambling security against risk neutral adversaries (or, equivalently, $\sqrt{\epsilon}$-message indistinguishability.) We may ask whether there is an inherent loss of security entailed by assuming adversaries have logarithmic utility, i.e., are there encryption schemes which are ϵ-gambling secure, but not ϵ'-message indistinguishable for some $\epsilon' \geq \epsilon$? In general, we would like to understand how assumptions about an adversary's utility impact security. This has the potential to contribute to a decision-theoretic approach to security (cf. *rational protocol design* as presented in [8].)

Acknowledgements We would like to thank Salil Vadhan and Colin Jia Zheng for their suggestions, and John Mitchell for asking the questions that led to this work.

References

[1] Boaz Barak, Ronen Shaltiel, and Avi Wigderson. Computational analogues of entropy. In Sanjeev Arora, Klaus Jansen, José D. P. Rolim, and Amit Sahai, editors, *6th International Workshop on Approximation Algorithms for Combinatorial Optimization Problems, APPROX 2003 and 7th International Workshop on Randomization and Approximation Techniques in*

Computer Science, RANDOM 2003, Proceedings, volume 2764 of *Lecture Notes in Computer Science*, pages 200–215. Springer, 2003.

[2] Mihir Bellare, Stefano Tessaro, and Alexander Vardy. Semantic security for the wiretap channel. In Reihaneh Safavi-Naini and Ran Canetti, editors, *Advances in Cryptology - CRYPTO 2012 - 32nd Annual Cryptology Conference, Santa Barbara, CA, USA, August 19-23, 2012. Proceedings*, volume 7417 of *Lecture Notes in Computer Science*, pages 294–311. Springer, 2012.

[3] Daniel Bernouilli. Exposition of a new theory on the measurement of risk. *Econometrica*, 22(1):23–36, 1954.

[4] Nicolò Cesa-Bianchi and Gábor Lugosi. *Prediction, Learning, and Games.* Cambridge University Press, 2006.

[5] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory (2nd Ed.).* Wiley, 2006.

[6] Yevgeniy Dodis. Shannon impossibility, revisited. In Adam Smith, editor, *Information Theoretic Security - 6th International Conference, ICITS 2012*, Lecture Notes in Computer Science, pages 100–110. Springer, 2012.

[7] Yoav Freund and Robert E. Schapire. Adaptive game playing using multiplicative weights. *Games and Economic Behavior*, 29(12):79 – 103, 1999.

[8] Juan A. Garay, Jonathan Katz, Ueli Maurer, Björn Tackmann, and Vassilis Zikas. Rational protocol design: Cryptography against incentive-driven adversaries. In *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013*, pages 648–657, 2013.

[9] Oded Goldreich. *The Foundations of Cryptography - Volume 1, Basic Techniques.* Cambridge University Press, 2001.

[10] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *J. Comput. Syst. Sci.*, 28(2):270–299, 1984.

[11] Mohammad Hajiabadi and Bruce M. Kapron. Gambling, computational information and encryption security. In Lehmann and Wolf [17], pages 141–158.

[12] Johan Håstad, Russell Impagliazzo, Leonid A. Levin, and Michael Luby. A pseudorandom generator from any one-way function. *SIAM J. Comput.*, 28(4):1364–1396, 1999.

[13] Chun-Yuan Hsiao, Chi-Jen Lu, and Leonid Reyzin. Conditional computational entropy, or toward separating pseudoentropy from compressibility. In Moni Naor, editor, *26th Annual International Conference on Advances in Cryptology - EUROCRYPT 2007,*, volume 4515 of *Lecture Notes in Computer Science*, pages 169–186. Springer, 2007.

[14] John L. Kelly, Jr. A new interpretation of information rate. *Bell System Technical Journal*, 35(4):917–926, 1956.

[15] Stephan Krenn, Krzysztof Pietrzak, and Akshay Wadia. A counterexample to the chain rule for conditional HILL entropy - and what deniable encryption has to do with it. In *TCC*, pages 23–39, 2013.

[16] Stephan Krenn, Krzysztof Pietrzak, Akshay Wadia, and Daniel Wichs. A counterexample to the chain rule for conditional HILL entropy. *IACR Cryptology ePrint Archive*, 2014:678, 2014.

[17] Anja Lehmann and Stefan Wolf, editors. *Information Theoretic Security - 8th International Conference, ICITS 2015, Lugano, Switzerland, May 2-5, 2015. Proceedings*, volume 9063 of *Lecture Notes in Computer Science*. Springer, 2015.

[18] Silvio Micali, Charles Rackoff, and Bob Sloan. The notion of security for probabilistic cryptosystems. *SIAM J. Comput.*, 17(2):412–426, 1988.

[19] Alexandre Pinto. Comparing notions of computational entropy. *Theory Comput. Syst.*, 45(4):944–962, 2009.

[20] Leonid Reyzin. Some notions of entropy for cryptography - (invited talk). In Serge Fehr, editor, *Information Theoretic Security - 5th International Conference, ICITS 2011, Amsterdam, The Netherlands, May 21-24, 2011. Proceedings*, volume 6673 of *Lecture Notes in Computer Science*, pages 138–142. Springer, 2011.

[21] Claude E. Shannon. Communication theory of secrecy systems. *Bell System Technical Journal*, 28(4):656–715, 1949.

[22] Maciej Skorski. Metric pseudoentropy: Characterizations, transformations and applications. In Lehmann and Wolf [17], pages 105–122.

[23] Alan M. Turing. Response to A. M. Uttley's 'Information, Machines and Brains'. In *Conference on Information Theory, London*, 1950.

[24] Salil P. Vadhan and Colin Jia Zheng. Characterizing pseudoentropy and simplifying pseudorandom generator constructions. *Electronic Colloquium on Computational Complexity (ECCC)*, 18:141, 2011.

[25] Hoeteck Wee. On pseudoentropy versus compressibility. In *19th Annual IEEE Conference on Computational Complexity (CCC 2004), 21-24 June 2004, Amherst, MA, USA*, pages 29–41. IEEE Computer Society, 2004.

[26] Andrew Chi-Chih Yao. Theory and applications of trapdoor functions (Extended abstract). In *23rd Annual Symposium on Foundations of Computer Science, Chicago, Illinois, USA, 3-5 November 1982*, pages 80–91, 1982.

[27] Andrew Chi-Chih Yao. Computational information theory. In Y.B. Abu-Mostafa, editor, *Complexity in Information Theory*, pages 1–15. Springer, 1988.

A An Alternate Proof of Lemma 5.5

Here we outline an alternate approach to proving a version of Lemma 5.5, using one of the main results of [24].

Suppose $\mathcal{X}, \mathcal{Y}$ are finite sets. A *measure* is a mapping $P : \mathcal{Y} \times \mathcal{X} \to (0, +\infty)$. Associated with P is a conditional mass function C_P defined by $C_P(x|y) = P(y,x)/\sum_{z \in \mathcal{X}} P(y,z)$. Suppose X, Y respectively are distributions on $\mathcal{X}, \mathcal{Y}$. X is (t, δ) *KL-hard given* Y if for any measure P computed by a circuit of size t, $\mathrm{KL}(Y, X \| Y, C_P) > \delta$.

Theorem A.1 ([24], Theorem 3.8(2))**.** *Let (Y, X) be a $\mathcal{Y} \times \mathcal{X}$-valued random variable, $\delta, \epsilon > 0$. If X has nonuniform (t, ϵ) conditional pseudoentropy at least $H(X|Y) + \delta$ given Y, then for every $\sigma > 0$, X is $(t', \delta - \sigma)$ KL-hard given Y, for $t' = \min\{t^{\Omega(1)}/\mathrm{polylog}(1/\sigma), \Omega(\sigma/\epsilon)\}$*

We now note that the betting function b defined in the proof of Lemma 5.5 is strictly positive, so it is a measure. Moreover $C_b = b$. We have

$$\mathrm{H}(M) + \mathrm{E}[\log b(M|\mathcal{E}_n(U_n, M))] \geq \tfrac{\delta^2}{\ln 2}$$

Now

$$\mathrm{H}(M) + \mathrm{E}[\log b(M|\mathcal{E}_n(U_n, M))] = \\ \mathrm{H}(M) - \mathrm{H}(M|\mathcal{E}_n(U_n, M)) - \mathrm{KL}(M|\mathcal{E}_n(U_n, M) \| b|\mathcal{E}_n(U_n, M)))$$

so that

$$\mathrm{KL}(M|\mathcal{E}_n(U_n, M) \| b|\mathcal{E}_n(U_n, M))) \leq \mathrm{H}(M) - \mathrm{H}(M|\mathcal{E}_n(U_n, M)) - \tfrac{\delta^2}{\ln 2}$$

Applying the above-cited result, we can conclude that given $\mathcal{E}_n(U_n, M)$, M does not have conditional pseudoentropy at least $\mathrm{H}(M)$, as required for the "if" direction of Theorem 5.7.

By appealing to the result of [24], we have made Lemmas 5.3 and 5.4 redundant. On the other hand, the work of Vadhan and Zheng involves considerable machinery beyond what is needed for Lemma 5.5.