Universal Optimality of Apollonian Cell Encoders

Fabrizio Biondi, Thomas Given-Wilson, Axel Legay

Universal Optimality of Apollonian Cell Encoders

Fabrizio Biondi[1], Thomas Given-Wilson[2], Axel Legay[2]

[1] CentraleSupélec, France
fabrizio.biondi@inria.fr
[2] Inria, France
{thomas.given-wilson,axel.legay}@inria.fr

Abstract. Preserving privacy of private communication against an attacker is a fundamental concern of computer science security. Unconditional encryption considers the case where an attacker has unlimited computational power, hence no complexity result can be relied upon for encryption. Optimality criteria are defined for the best possible encryption over a general collection of entropy measures. This paper introduces *Apollonian cell encoders*, a class of shared-key cryptosystems that are proven to be universally optimal. In addition to the highest possible security for the message, Apollonian cell encoders prove to have perfect secrecy on their key allowing unlimited key reuse. Conditions for the existence of Apollonian cell encoders are presented, as well as a constructive proof. Further, a compact representation of Apollonian cell encoders is presented, allowing for practical implementation.

1 Introduction

Preserving privacy of private communication with encryption is a fundamental concern of computer science. Such efforts can be divided into two categories, according to whether they assume that the adversary trying to break the encryption has or does not have access to unlimited computational power.

Computational encryption schemes assume that the attacker's computational power is bounded, usually meaning that the attacker cannot solve problems with an superpolynomial complexity, e.g. integer factorization, subset sum, etc. Note however that such complexity results are currently unproven, and may be weakened by technological progress. For instance, Shor's algorithm and quantum computing are effective for integer factorization [**?**], Grover's algorithm and quantum computing can be used against AES [**?**], and memcomputing has been proved to be effective to solve subset sum [**?**].

Unconditional encryption is instead based upon information-theoretic reasoning and proven independently of computational hardness. Thus unconditional encryption results have an elegant mathematical proof of robustness that does not rely on complexity results. However, typically unconditional encryption has much stricter requirements to obtain than computational security [**?**, **?**, **?**, **?**, **?**]. This paper explores only unconditional encryption, focusing on results that are

not conditioned upon any hardness results and so are robust against technological progress.

Fundamental to unconditional encryption is the information measure used to quantify the security of the system being considered, as it is correlated with the attacks against which the cryptosystem is resistant [**?**]. The first formal results on unconditional encryption were published by Shannon [**?**] using results from the formal theory of communication [**?**]. These results exploit Shannon entropy (here denoted as S) as the measure of information, and also introduce *perfect secrecy* as the highest level of security that can be achieved. However, Shannon also proves that perfect secrecy on the message is achievable only by using a key as large as the message, and the key has to be discarded after every use.

Recent work by Khouzani and Malacaria [**?**] has introduced *Khouzani-Malacaria (KM-) entropy*, a very general definition of entropy that generalizes Rényi entropy [**?**], and thus in turn most common and popular entropy measures. This includes generalizing Shannon entropy and min-entropy [**?**] (here denoted as H_∞). The results in this work exploit KM-entropy to yield the greatest generality. In [**?**] Khouzani and Malacaria also recently explored the relation with convex optimization problems and other security scenarios.

Recent work has considered the unconditional encryption of shared-key cryptosystems and presented *max-equivocation* that generalizes Shannon's perfect secrecy and is a measure of the best possible unconditional encryption that is always achievable [**?**, **?**]. The scenario considered is when a sender sends a message m encoded with a key k to a ciphertext c via an encoder enc to a receiver. The receiver uses the same key k and a decoder dec to obtain the message from the ciphertext c. An adversary is able to observe c and, using unlimited computational power, attempts to gain information about m and k using only c and knowledge of enc and dec. Equivocation measures the amount of secrecy of the message or key after the adversary has observed the ciphertext, hence quantifying the difficulty for the adversary to deduce the exact message or key used to produce the ciphertext.

This scenario can be generalized from a specific message, key, and ciphertext to measure the equivocation of the message (and key) based upon the prior distribution of the messages and keys, and the properties of the encoder enc. The prior uncertainty of the adversary on the messages is represented as $H(M)$ and for the key as $H(K)$, where H is a KM-entropy. The equivocation, quantifying the uncertainty of the adversary about the message after observing a ciphertext, can be represented by $H(M|C)$ and similarly $H(K|C)$ for the key.

Biondi et al. [**?**, **?**] show that the upper bound on the message equivocation $H(M|C)$ and key equivocation $H(K|C)$ is the minimum of the prior distributions, i.e $\min(H(M),\ H(K))$. Intuitively, the adversary can exploit their greater prior knowledge on either the message or key to learn more information about the other. For example, if the adversary knows that the key is k_1 (i.e. $H(K) = 0$) then after seeing a ciphertext c the adversary can use the key k_1 to obtain the message m by $dec(c, k_1) = m$. Since only one message m corresponds to $dec(c, k_1) = m$

then $H(M|C) = 0 = H(K)$, showing that no message equivocation can be preserved against an adversary that possesses the encryption/decryption key.

These results prove that the upper bound is $\min(H(M), H(K))$ which translates to $H(K)$ under the (typical) assumption that the message has more entropy than the key (i.e. $H(M) > H(K)$). The best theoretically possible cryptosystem would achieve max-equivocation on the message ($H(M|C) = H(K)$) and perfect secrecy on the key($H(K|C) = H(K)$). Since such cryptosystem does not lose any information on the key, it also allows the key to be reused indefinitely, contrarily to the case studied by Shannon in which the key can be used only once and then has to be discarded to preserve perfect secrecy on the message. However, these results were only proved for Shannon entropy and min-entropy, and indicate the theoretical bound without providing a constructive way to build an encoder that achieves them.

The first contribution of this paper is to generalize these results to KM-entropy. This shows that the same results hold when H is any KM-entropy. From this it is possible to define *message (resp. key) optimality* for a cryptosystem: when the cryptosystem achieves message (resp. key) max-equivocation for a given KM-entropy. This is the best result that can be obtained by a cryptosystem, meaning that no other cryptosystem can be more effective at protecting the secrecy of the message (resp. key) from the adversary. This is generalized to *universal message (resp. key) optimality* of a cryptosystem is message optimal for all KM-entropy measures. Lastly, define *universal optimality* if a cryptosystem is both universally message optimal and universally key optimal for all KM-entropy measures. Note that universal optimality is also discusses in [**?**], however only on the message and only if the sender is allowed to choose the key distribution, while here universal key optimality is also obtained and a uniform prior distribution on the key is sufficient.

The second contribution of the paper is the development of *cell encoders*. Cell encoders generalize the encoders of Biondi et al [**?**] by allowing an encoder function to map to a distribution over the ciphertext space instead of only a single ciphertext. In particular cell encoders map a message and key uniformly to a subset of the ciphertext space, where each ciphertext in the subset has the same probability of being chosen. To maintain the desirable decodability behaviour, cell encoders require *unique decodability*; that given a ciphertext c and key k, the ciphertext can be uniquely decoded to a single message m.

The third contribution is to define *Apollonian cell encoders* and prove that they are universally optimal when the key distribution is uniform. Apollonian cell encoders exploit the mapping of a message-key pair to multiple ciphertexts to obtain a uniform distribution over the ciphertext space. Since both the key and the ciphertext distributions are uniform, this yields a uniform distribution for the messages given the ciphertext, and thus achieves max-equivocation for the message, and perfect secrecy for the key.

One major advantage of this result that Apollonian cell encoders achieved key secrecy is key reuse. Since no information about the key is leaked to the adversary, this means the same key can be used an infinite number of times

without reducing the message or key equivocation. In practice this means that the same key can be reused infinitely and will yield a set of sequences of messages, where each sequence is equally likely to hold from the adversary's perspective.

The fourth contribution of this paper is to prove the necessary and sufficient conditions for an Apollonian cell encoder to exist. The conditions are quite reasonable: (1) that the highest probability message has lower probability than any key, i.e. $H_\infty(M) \geq H_\infty(K)$; and (2) that the probabilities of the messages can all be expressed as rational numbers. Under these conditions it is always possible to define an Apollonian cell encoder that is universally optimal.

The fifth contribution is a compact and practical representation of an Apollonian cell encoder. A universally-optimal Apollonian cell encoder can be represented as a simple combination of the sum, multiplication, and modulo operations over the integers, making it very simple and convenient to implement in practice.

The rest of this paper develops the above in careful detail.

The structure of the paper is as follows. Section **??** introduces notation and concepts for understanding the paper. Section **??** defines optimality for KM-entropy. Section **??** proves optimality of Apollonian cell encoders. Section **??** proves existence and construction of Apollonian cell encoders. Section **??** discusses concerns and limitations with the work here. Section **??** draws conclusions and considers future work.

2 Background

This section recalls notation, definitions, and concepts used throughout this work. After recalling notation, Khouzani-Malacaria entropy is detailed, followed by shared-key cryptosystems, and finally modeling shared-key cryptosystems with Khouzani-Malacria entropy.

The size of a set $\mathcal{S}$ is denoted as $|\mathcal{S}|$. $\mathcal{P}(\mathcal{X})$ denotes the powerset of set $\mathcal{X}$. A function $f : \mathcal{A} \to \mathcal{B}$ is *injective* iff $\forall a_1, a_2 \in \mathcal{A}.\ f(a_1) = f(a_2) \Rightarrow a_1 = a_2$.

Basic concepts from probability and information theory can be found in the literature [**?**], including the definitions of support set $\mathcal{X}$, probability $P(E)$ of an event $E \subseteq \mathcal{X}$, random variable X on $\mathcal{X}$, entropy $H(X)$ of a random variable X, conditional entropy $H(X|Y)$ of a random variable X given another random variable Y, and so on. In this paper $\rho_\mathcal{X}(X)$ denotes a probability distribution on the random variable X on the support set $\mathcal{X} = Supp(\rho_\mathcal{X}(X))$, abbreviated to $\rho(X)$ when the support set is unambiguous, and $\wp(\mathcal{X})$ denotes the space of all probability distributions on $\mathcal{X}$.

2.1 Khouzani-Malacaria Entropy

This section recalls the concept of Khouzani-Malacaria (KM-) entropy from [**?**]. KM-entropy generalizes most commonly-used entropy measures, including all Renyi entropy measures, thus results showing optimality of any KM-entropy

hold for these commonly used entropy measures. An important property of KM-entropy is that any KM-entropy is maximized on a uniform distribution, as will be recalled in Lemma **??**. The rest of this section recalls KM-entropy in detail, including definitions of equivocation and leakage for KM-entropy.

Let a random variable $\boldsymbol{\theta}$ represent a secret from a discrete finite set of secrets $\Theta = \{\theta_1, \ldots, \theta_n\}$ of size n. Consider a defender trying to hide the secret, and an adversary trying to guess it. Let the secret be generated by a probability distribution $\rho(\boldsymbol{\theta})$ which is known to both defender and adversary.

The defender is provided with the *realization* θ of the secret, and must then choose a *cloak*[3] $W \in \mathcal{P}(\Theta)$. The adversary is given the cloak W chosen by the defender and has to try and guess θ. Note that if the defender always chooses the same cloak W for every secret then the adversary gains no information to guess θ by observing W; this corresponds to Shannon's definition of perfect secrecy [**?**].

For each secret θ the defender uses a (randomized) *cloaking strategy* δ to assign the cloak W with a given probability. The probability that cloak W is assigned to secret θ by strategy δ is denoted as $\delta(W;\theta)$. A cloaking strategy δ defines a set of *feasible cloaks* $\mathcal{W}^+$ as $\mathcal{W}^+ = \{W \in \mathcal{P}(\Theta) \mid \exists\theta.\ \delta(W;\theta) > 0\}$. The adversary knows the cloaking strategy δ.

Entropy is a commonly used measure of the uncertainty of a system represented by a probability distribution. Entropy can conveniently represent the information that the adversary has about the secret message before and after an attack. Consequently entropy can be used to quantify the information gained by the adversary due to the attack [**?**, **?**, **?**, **?**, **?**, **?**]. The entropy of a probability distribution ρ over a support set $X = \{x_1, \ldots, x_n\}$ is denoted as $\mathcal{H}(\rho(X))$, or simply by $\mathcal{H}(\rho)$ when the support set is obvious from the context. For the application of entropy to security, many different entropy measures have been proposed, modeling different adversaries and security scenarios. These include *Shannon entropy* $S(\rho(X)) = -\sum_{x\in X} \rho(x) \log \rho(x)$ [**?**] and *min-entropy* $H_\infty(\rho(X)) = -\log \max_{x\in X} \rho(x)$ [**?**]. A generalization of many of these entropy measures (including Shannon and min-entropy) has been proposed by Renyi [**?**], and more recently Khouzani and Malacaria [**?**] proposed a definition of entropy that generalizes Renyi entropy. This *Khouzani-Malacaria (KM-) entropy* is defined as follows:

Definition 1 (Khouzani-Malacaria (KM-) Entropy). *A function $H(\rho)$ from a probability distribution ρ to a real number is a KM-entropy measure if it satisfies the following.*

Symmetry *$H(\rho)$ is invariant under permutation of elements of ρ (i.e. depends on the probabilities but not on their order or labeling).*

Expansibility *Adding elements with probability zero to ρ does not change $H(\rho)$.*

Core-Concavity *$H(\rho)$ can be written as $\eta(F(\rho))$ where $\eta : \mathbb{R} \to \mathbb{R}$ is a non-constant function on real numbers, F is a scalar function on probability distributions, and it holds that either:*

[3] The cloak is referred to as M in [**?**], here W is used to avoid confusion with the message M. Also note that here it is not required that $\theta \in W$.

1. η *is increasing and* F *is concave in* ρ*, or*
2. η *is decreasing and* F *is convex in* ρ*.*

Given a KM-entropy measure H, the *prior entropy* of the secret is $H(\rho(\boldsymbol{\theta}))$. The prior entropy quantifies the uncertainty of the adversary on secret $\boldsymbol{\theta}$ before observing any cloak.

As remarked by Khouzani and Malacaria [**?**], any KM-entropy is maximal on a uniform distribution as a consequence of the symmetry and core-concavity properties:

Lemma 1. *Given a KM-entropy measure* H *and a support set* $\mathcal{X}$*, then the probability distribution* ρ *maximizing* $H(\rho(\mathcal{X}))$ *is the uniform distribution* $\forall x \in \mathcal{X}.\ \rho(x) = \frac{1}{|\mathcal{X}|}$.

Let $\mathbf{W}$ be the random variable on the support set $\mathcal{W}^+$ associated with the observation of the cloak W. *KM-equivocation* (or *conditional KM-entropy*) is used to quantify the uncertainty of the adversary on secret $\boldsymbol{\theta}$ after observing a cloak W, and is defined as $H(\rho(\boldsymbol{\theta})|W)$.

Definition 2 (KM-Equivocation). *The KM-equivocation* $H(\rho(\boldsymbol{\theta})|\mathbf{W})$ *for a KM-entropy* $H(\rho(\boldsymbol{\theta})) = \eta(F(\rho(\boldsymbol{\theta})))$ *is*

$$H(\rho(\boldsymbol{\theta})|\mathbf{W}) = \eta\left(\sum_{W \in \mathcal{W}^+} P(W) F(P(\rho(\boldsymbol{\theta})|W))\right)$$

where $P(W) = \sum_{\theta \in \Theta} \rho(\theta)\delta(W;\theta)$ *is the probability of observing the cloak* W *and* $P(\rho(\boldsymbol{\theta})|W)$ *is the distribution on the secret* $\rho(\boldsymbol{\theta})$ *normalized on the cloak* W*.*

Given the prior entropy of the secret $H(\rho(\boldsymbol{\theta}))$ and the equivocation $H(\rho(\boldsymbol{\theta})|\mathbf{W})$ after the adversary has observed the cloak chosen by the defender according to strategy δ, it is possible to define the secret's *leakage*. That is, the expected amount of information gained by the adversary on the secret by observing the chosen cloak according to δ, as the difference between the secret's prior entropy and equivocation.

Definition 3 (Leakage). *The* leakage $\mathscr{L}(\boldsymbol{\theta})$ *for a given KM-entropy measure* H *on a secret* $\boldsymbol{\theta}$ *is* $\mathscr{L}(\boldsymbol{\theta}) = H(\rho(\boldsymbol{\theta})) - H(\rho(\boldsymbol{\theta})|\mathbf{W})$.

As remarked by Khouzani and Malacaria [**?**], leakage is always non-negative, even when H is not defined positive.

2.2 Shared-Key Cryptosystems

This section recalls the definition of a shared-key cryptosystem, components, and typical use as the basis for the results and scenarios in this paper. Importantly, this paper does *not* address public-key cryptography in any way.

A shared-key cryptosystem can be defined by the following components (adapted from [**?**]).

Definition 4 (Cryptosystem). *A* (shared-key) cryptosystem *is a 4-tuple* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$ *where:*

- *the* message space $\mathcal{M}$ *is a finite set of possible messages;*
- *the* key space $\mathcal{K}$ *is a finite set of possible keys;*
- *the* ciphertext space $\mathcal{C}$ *is a finite set of possible ciphertexts;*
- *the* encoder *enc is a function* $\mathcal{M} \times \mathcal{K} \to \wp(\mathcal{C})$ *from message space and key space to all probability distributions over* $\mathcal{C}$*.*

Note that in [**?**] $\mathcal{C}$ is inductively defined by *enc* since *enc* is defined as a function $\mathcal{M} \times \mathcal{K} \to \mathcal{C}$; however here a more general, probabilistic encoder is considered. Also, in [**?**] it is required that $\forall k \in \mathcal{K}.\ enc(\cdot, k)$ is injective; here this is replaced by the Unique Decodability assumption below.

A given encoder *enc* here maps each message-key pair (m, k) to a probability distribution over $\mathcal{C}$. The probability that a given element of c is chosen is denoted as $P(c|m, k)$, and respects $0 \leq P(c|m, k) \leq 1$ and $\sum_{c \in \mathcal{C}} P(c|m, k) = 1$.

The *cloak* $W(c)$ of a ciphertext $c \in \mathcal{C}$ is the set of all message-key pairs that have a positive probability of producing c via the encoder *enc*, i.e. $W(c) = \{(m \in \mathcal{M}, k \in \mathcal{K}) \mid P(c|m, k) > 0\}$. The *projection* of $W(c)$ on the message (resp. key) space is denoted as $W_m(c)$ (resp. $W_k(c)$). The following is assumed in the rest of the paper and guarantees that a ciphertext c can be uniquely decoded by the receiver using the shared key k.

Assumption 1 (Unique Decodability) *A cryptosystem* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$ *is* uniquely decodable *iff for each key* $k \in \mathcal{K}$ *and ciphertext* $c \in \mathcal{C}$*, a pair* $(\cdot, k)$ *appears at most once in each cloak* $W(c)$*.*

A uniquely decodable cryptosystem guarantees the existence of a decoder function $dec : \mathcal{C} \times \mathcal{K} \to \mathcal{M}$. such that if $c \in Supp(enc(m, k))$ then $dec(c, k) = m$.

The channel model of a cryptosystem was introduced by Shannon [**?**]. In this model, the sender wants to send a message $m \in \mathcal{M}$ to the receiver on a public channel that is eavesdropped by an attacker. Initially, the sender and receiver share a secret *key* $k \in \mathcal{K}$.

The sender encodes the message m with key k into a ciphertext c by choosing such c with probability $P(c|m, k)$. The sender then sends this c to the receiver via a public channel, where c is also eavesdropped by the attacker. Knowing the key k, the receiver decodes the message $m = dec(c, k)$ using the decoder function. Not knowing the key k, the attacker tries to infer m and k from c. The computational power available to the sender, receiver, and attacker is assumed to be unlimited, and the attacker is assumed to know $\mathcal{M}$, $\mathcal{K}$, $\mathcal{C}$, *enc*, and ρ.

The attacker's knowledge about the realizations of the message, key, and ciphertext is modeled by random variables. Let M (resp. K, C) be a random variable on the support set $\mathcal{M}$ (resp. $\mathcal{K}$, $\mathcal{C}$) representing the value of the message m (resp. key k, ciphertext c) according to the attacker.

2.3 Information-theoretical Cryptosystem Modeling

To use the information-theoretical model from Section **??** to model a cryptosystem from Section **??**, consider that instead of having a defender and an adversary, there is a sender, a receiver, and an attacker. The secret is a pair (m, k) of message and key, where the attacker tries to guess m (and k), and k is pre-shared between the sender and the receiver. The prior probability distribution on the message-key pairs is $\rho(M, K)$, and its projections on the message and the key spaces are $\rho(M)$ and $\rho(K)$, respectively. Entropy allows reasoning about the information that the attacker has both on the message and on the key. The entropy of the key is also important since each ciphertext c is uniquely decodable to a given message m given the key k. Hence Definition **??** can be applied separately to the message and the key obtaining the *prior message entropy* $H(\rho(M))$ and the *prior key entropy* $H(\rho(K))$, respectively, where the distribution ρ will be omitted when it is obvious from the context.

The transmission of a ciphertext is a convenient way for the sender to inform the receiver about which cloak has been chosen. Both the receiver and the attacker are assumed to have unlimited computational power and access to the encoder, hence they are immediately able to map a ciphertext $c \in Supp(enc(m, k))$ to its cloak $W(c)$.

The cloaking strategy $\delta(W, \theta)$ is modeled by the choice of the encoder *enc*. As remarked above, (1) the secret θ is a pair of message and key (m, k), (2) choosing a ciphertext c is equivalent to choosing its cloak $W(c)$, and (3) the ciphertext is a function of the message m, key k, and encoder *enc* via $c \in Supp(enc(m, k))$. Hence the probability $\delta(W, \theta)$ of choosing cloak W for secret θ becomes $P(c|m, k)$, i.e. the probability that ciphertext c is chosen for message-key pair (m, k) according to encoder *enc*. This allows the rewriting of Definition **??** to quantify the *message equivocation* and *key equivocation* for a given KM-entropy H.

Definition 5 (Message Equivocation). *Given a cryptosystem* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$*, a probability distribution* ρ *on* $\mathcal{M} \times \mathcal{K}$*, and a KM-entropy measure* $H = \eta(F(\rho))$*, the* message equivocation $H(M|C)$ *of the cryptosystem according to* H *is*

$$H(M|C) = \eta\left(\sum_{c \in \mathcal{C}} P(c) F(P(M|c))\right)$$

where $P(c) = \sum_{(m,k) \in \mathcal{M} \times \mathcal{K}} \rho(m, k) P(c|m, k)$ *is the probability of observing ciphertext* c *and* $P(M|c)$ *is the distribution on the message* $\rho(M)$ *normalized on* $W_m(c)$*.*

Definition 6 (Key Equivocation). *Given a cryptosystem* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$*, a probability distribution* ρ *on* $\mathcal{M} \times \mathcal{K}$*, and a KM-entropy measure* $H = \eta(F(\rho))$*, the* key equivocation $H(K|C)$ *of the cryptosystem according to* H *is*

$$H(K|C) = \eta\left(\sum_{c \in \mathcal{C}} P(c) F(P(K|c))\right)$$

where $P(c) = \sum_{(m,k)\in\mathcal{M}\times\mathcal{K}} \rho(m,k)P(c|m,k)$ *is the probability of observing ciphertext* c *and* $P(K|c)$ *is the distribution on the key* $\rho(K)$ *normalized on* $W_k(c)$.

Following Definition **??**, the *message leakage* and *key leakage* for a given KM-entropy H can also be defined.

Definition 7 (Message Leakage). *Given a cryptosystem* $(\mathcal{M},\mathcal{K},\mathcal{C},\text{enc})$*, a probability distribution* ρ *on* $\mathcal{M}\times\mathcal{K}$*, and a KM-entropy measure* H*, the* message leakage $\mathscr{L}(M)$ *is*

$$\mathscr{L}(M) = H(M) - H(M|C) \quad .$$

Definition 8 (Key Leakage). *Given a cryptosystem* $(\mathcal{M},\mathcal{K},\mathcal{C},\text{enc})$*, a probability distribution* ρ *on* $\mathcal{M}\times\mathcal{K}$*, and a KM-entropy measure* H*, the* key leakage $\mathscr{L}(K)$ *is*

$$\mathscr{L}(K) = H(K) - H(K|C) \quad .$$

3 Universal Optimality

This section explores definitions of optimal for a cryptosystem with respect to KM-entropy measures. In particular, the definition of being optimal for message equivocation, key equivocation, and universally optimal. These results build upon and generalize the work of Biondi et al. [**?**] on max-equivocation, and Khouzani and Malacaria [**?**] for KM-entropy.

Importantly, in this paper the following assumptions also hold:

Assumption 2 (Key Uniformity) *The key distribution* $\rho(K)$ *is assumed to be uniform, i.e.* $\forall k \in \mathcal{K}.\ \rho(k) = \frac{1}{|\mathcal{K}|}$.

The Key Uniformity assumption is very reasonable in a practical setting, since producing valuable uniform randomness is a well-understood problem in cryptography. In general this is not required for all results, in particular for cryptosystems the rôle of message and key can be reversed and this assumption applied to the message distribution instead. For a detailed exploration of this refer to Biondi et al. [**?**].

Assumption 3 (Message-Key Independence) *The message distribution* $\rho(M)$ *and the key distribution* $\rho(K)$ *are assumed to be independent, i.e.* $\forall m \in \mathcal{M}, k \in \mathcal{K}.\ \rho(m,k) = \rho(m)\rho(k)$.

The Message-Key Independence assumption is also reasonable, since the choice of the key should never depend on the choice of the message, otherwise it would likely leak information about the message.

Biondi et al. [**?**] derived maximum message equivocation bounds for Shannon entropy:

Lemma 2 (from Theorem 2 of [?]). *Given a cryptosystem* $(\mathcal{M},\mathcal{K},\mathcal{C},\text{enc})$ *and a probability distribution* ρ *on* $\mathcal{M}\times\mathcal{K}$*, then* $S(M|C) \leq S(K)$.

Note that Lemma **??** also holds for min-entropy by replacing S with H_∞. In fact, this result can be extended to any KM-entropy measure as proved below.

Theorem 1. *Given a cryptosystem* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$*, a probability distribution* ρ *on* $\mathcal{M} \times \mathcal{K}$*, and a KM-entropy measure* H*, then* $H(M|C) \leq H(K)$*.*

Proof. If $H(M) \leq H(K)$ this is trivial since $H(M|C) \leq H(M)$ by non-negativity of leakage, so let $H(M) > H(K)$. Since each key appears at most once in the cloak of each ciphertext by unique decodability (Assumption **??**), no cloak is larger than the key space: $\forall c.\ |W(c)| \leq |\mathcal{K}|$. Since the key is uniformly distributed over $\mathcal{K}$, this is sufficient to show that $\forall c.\ H(M|c) \leq H(\mathcal{K})$ by Lemma **??**. Finish by Definition **??**.

It follows that the message equivocation is bounded from above by $\min(H(M), H(K))$. Biondi et al. [**?**] call an encoder achieving such upper bound as satisfying *max-equivocation*: $H(M|C) = \min(H(M), H(K))$.

More generally, given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$ and KM-entropy measure H, it is possible to derive the message (resp. key) max-equivocation that any cryptosystem $(\mathcal{M}, \mathcal{K}, \mathcal{C}, enc)$ can preserve.

Definition 9 (Message and Key Max-Equivocation). *A cryptosystem* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$ *satisfies* message *(resp.* key*)* max-equivocation *for a given probability distribution* ρ *on* $\mathcal{M} \times \mathcal{K}$ *and KM-entropy measure* H *when* $H(M|C) = \min(H(M), H(K))$ *(resp.* $H(K|C) = \min(H(M), H(K))$*).*

Observe that when the message (resp. key) entropy is lesser, then this corresponds to Shannon's perfect secrecy [**?**] on the message (resp. key) since $H(M|C) = \min(H(M), H(K)) = H(M)$ (resp. $H(K|C) = \min(H(M), H(K)) = HK$)). This property of max-equivocation corresponding to perfect secrecy holds by design of Biondi et al. since max-equivocation generalizes perfect secrecy [**?**].

Once a KM-entropy measure and probability distribution (on messages and keys) is fixed, it is possible to define the properties of an optimal cryptosystem through message and key max-equivocation.

Definition 10 (Message and Key Optimality). *Given a KM-entropy measure* H *and probability distribution* ρ *on* $\mathcal{M} \times \mathcal{K}$*, a cryptosystem* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$ *is* message-optimal *(resp.* key-optimal*)* for H *iff* $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$ *achieves message (resp. key) max-equivocation.*

Observe that in practice such a cryptosystem must achieve perfect secrecy on the message or key since it must achieve both $H(M|C) = \min(H(M), H(K))$ and $H(K|C) = \min(H(M), H(K))$. In the case where $H(M) = H(K)$ then such a cryptosystem achieves perfect secrecy on both the message and key.

When perfect secrecy is achieved on the key then this allows the key to be reused indefinitely, since the system does not leak any information about the key. In practice this yields sequences of messages that correspond to each possible key, however from the attacker's perspective all such sequences of messages are equally likely. (Note that each position in the sequences will correspond to one message being sent, and collectively will have $H(M|C)$ equivocation.)

This definition of optimality can be generalised to account for all KM-entropy measures as follows.

Definition 11 (Universal Message and Key Optimality). *Given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$, a cryptosystem $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$ is* universally message-optimal *(resp.* universally key-optimal*) iff it is message-optimal (resp. key-optimal) for* any *KM-entropy measure H.*

Recall from Lemma **??** that any KM-entropy is maximal on a uniform distribution, and that leakage is minimized when equivocation is maximized; then, an encoder such that the posterior distribution of the message (resp. key) given each ciphertext is always uniform on a space of size $\min(|\mathcal{M}|, |\mathcal{K}|)$ is universally message-optimal (resp. universally key-optimal).

Combining both universal message and key optimality provides a definition for universal optimality.

Definition 12 (Universal Optimality). *Given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$, a cryptosystem $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$ is* universally optimal *if it is both universally message-optimal and universally key-optimal.*

In Section **??** an encoder with universal optimality is presented, i.e. the best possible cryptosystem theoretically achievable for any KM-entropy.

4 Apollonian Cell Encoders are Universally Optimal

This section presents cell encoders and Apollonian cell encoders, concluding with proofs that Apollonian cell encoders are universally optimal. The rest of this section assumes a message space $\mathcal{M}$, key space $\mathcal{K}$, and ciphertext space $\mathcal{C}$.

4.1 Cell encoders

This section presents cell encoders that may map a single message m and key k to a set of possible ciphertexts. Conceptually this is a mapping from a message and key to a cell that contains this set of ciphertexts.

Define a *cell function* $CELL(m, k)$ that given a message $m \in \mathcal{M}$ and key $k \in \mathcal{K}$ returns a non-empty subset of $\mathcal{C}$. The cell function is exploited to define a *cell encoder* as follows.

Definition 13 (Cell Encoder). *A* cell encoder *enc is a function $\mathcal{M} \times \mathcal{K} \to \wp(\mathcal{C})$ that maps a message m and a key k to a probability distribution uniform on $CELL(m, k)$ and zero on $\mathcal{C} \setminus CELL(m, k)$.*

Observe that prior definitions of encoders [**?**, **?**] can also be considered cell encoders by assuming that the $CELL$ function always returns a set of size 1. Thus, the rest of this paper shall assume all encoders are cell encoders and the $CELL$ function is defined implicitly and not shown in the notation.

Let $\Delta(c \in CELL(m,k))$ be the Dirac delta function for the statement $c \in CELL(m,k)$, i.e.

$$\Delta(c \in CELL(m,k)) = \begin{cases} 1 \text{ if } c \in CELL(m,k) \\ 0 \text{ otherwise} \end{cases} .$$

Then the probability of each ciphertext c can be defined as:

$$\rho(c) = \sum_{m \in \mathcal{M}} \sum_{k \in \mathcal{K}} \Delta(c \in CELL(m,k)) \frac{\rho(m,k)}{|CELL(m,k)|} \tag{1}$$

and the conditional probability of each ciphertext c given a message m or a key k as

$$\rho(c|m) = \sum_{k \in \mathcal{K}} \Delta(c \in CELL(m,k)) \frac{\rho(k)}{|CELL(m,k)|} \tag{2}$$

and

$$\rho(c|k) = \sum_{m \in \mathcal{M}} \Delta(c \in CELL(m,k)) \frac{\rho(m)}{|CELL(m,k)|} \tag{3}$$

respectively by Assumption **??**.

A cell encoder is *semi-injective on the message* if $\forall k \in \mathcal{K}.\ enc(\cdot, k)$ is injective, i.e. for all messages m_i and m_j and all keys k then $m_i \neq m_j$ implies $CELL(m_i,k) \cap CELL(m_j,k) = \emptyset$. Note that this is enforced by Assumption **??**.

A cell encoder is *semi-injective on the key* if $\forall m \in \mathcal{M}.\ enc(m, \cdot)$ is injective, i.e. for all keys k_i and k_j and all messages m then $k_i \neq k_j$ implies $CELL(m,k_i) \cap CELL(m,k_j) = \emptyset$.

4.2 On Building a Universally Optimal Encoder

This section provides intuition useful to understand how to build an encoder with universal optimality properties. Recall that in this paper cryptosystems are assumed to be uniquely decodable (Assumption **??**), the probability distribution on the key is assumed to be uniform (Assumption **??**), and the probability distributions of message and key are assumed to be independent (Assumption **??**).

Consider the simple case of uniformly-distributed message and key spaces with three elements each: $\mathcal{M} = \{m_0, m_1, m_2\}$ and $\mathcal{K} = \{k_0, k_1, k_1\}$ and $\rho(m_0) = \rho(m_1) = \rho(m_2) = \rho(k_0) = \rho(k_1) = \rho(k_2) = 1/3$. Lemma **??** states that any KM-entropy measure is maximal on a uniform distribution and Definition **??** remarks that message leakage is minimal when message equivocation is maximal and prior message entropy does not depend on the encoder. Hence, it is simple to see that an encoder that produces a uniform conditional distribution on the message space for any ciphertext maximizes message equivocation and thus minimizes message leakage on any KM-entropy measure, making such encoder universally optimal. (The same holds with message replaced by key.)

Now consider the encoder depicted in Table **??** on the ciphertext space $\mathcal{C} = \{c_0, c_1, c_2\}$. It is easy to see that each ciphertext appears in each row of the

encoder. This means that the cloak of the encoder on the message is the whole message space, hence the attacker is not able to immediately exclude any message from being the one used by the sender just by observing a ciphertext. If the cloak of a ciphertext on the message space did not include all of the messages, then the attacker would gain information about the message sent by observing that ciphertext, and this would cause message leakage. As before, the same holds for the key by considering each column of the encoder. Note that having a ciphertext appearing more than once in the same row would be acceptable, while having a ciphertext appearing more than once in the same column would violate unique decodability (Assumption **??**).

Consider the transmission of multiple ciphertexts again using the encoder in Table **??**. Since the message is redrawn from the message distribution every time, even sequences of ciphertexts will not reveal any information about the key, even though the key is never changed. However, the attacker will be able to map each possible key to a sequence of messages. For instance, if the attacker observes the sequence of ciphertexts (c_0, c_2, c_0) then they can infer that the message sequence is (m_0, m_2, m_0) if the key is k_0, (m_2, m_1, m_2) if the key is k_1, and (m_1, m_0, m_1) if the key is k_2. Since these sequences are equally probable this still does not provide any information about the key to the attacker, so key leakage is zero. However, the attacker will know to exclude 24 out of the 27 possible three-message sequences. Since this is implicitly considering sequences of three messages as the message space, this is unavoidable by Theorem **??** showing that no message (or sequence thereof) can have more equivocation than the entropy of the key used to transmit the sequence. Hence, no better equivocation can be obtained without sharing a new key between the sender and the receiver.

Table 1: An encoder for $|\mathcal{M}| = |\mathcal{K}| = 3$.

		Key		
		k_0	k_1	k_2
Message	m_0	c_0	c_1	c_2
	m_1	c_1	c_2	c_0
	m_2	c_2	c_0	c_1

For the case of Table **??** it is very easy to construct an optimal encoder because (1) the message and key space have the same size, and (2) the message distribution is uniform. However, both these properties are very uncommon in practical cryptography: in general it is preferable to use a key space smaller than the message space, and some messages (e.g. "Hi") will have a very different probability of being transmitted than some others (e.g. "In a right-angle triangle, the sum of the squares of the smaller two sides is equivalent to the square of the hypotenuse"). Note that the assumption of uniform key distribution (Assumption **??**) is instead quite reasonable, since the key can be produced using high-quality randomness and does not have to carry actual information or mean-

ing like the message do. The rest of this section demonstrates what happens to optimal and universal encoding when relaxing some of these properties.

Now consider the encoder depicted in Table **??**, where the message and ciphertext spaces have been extended by one element and the probability of each message is uniform at $1/4$. Each ciphertext still appears in each column, so key leakage is still zero. However, not every ciphertext appears in every row, hence the cloak of each ciphertext on the message is not the whole message space, causing message leakage. For example, if the attacker observes ciphertext c_2 they can immediately infer that the message is not m_3, while gaining no information on the key (even on repeated transmissions, as explained above).

Note that in this case $H(M) > H(K)$ for any KM-entropy measure H, hence again by Theorem **??** it is not possible to have an encoder with zero message leakage even on a single transmission, while it is possible to have an encoder with zero key leakage like the one depicted in Table **??**. In this case the best possible result is for the cryptosystem to have universal message optimality and universal key optimality (both Definition **??**). To have such properties each ciphertext must have a uniformly-distributed cloak of size three on both the message and the key. It can be verified that the encoder in Figure **??** achieves this, and hence has both universal message optimality and universal key optimality, making it the best possible encoder for uniformly-distributed message and key spaces of sizes 4 and 3 respectively. Observe that in this case for all KM-entropy measures, the encoder achieves max-equivocation on the message, and perfect secrecy on the key.

Table 2: An encoder for $|\mathcal{M}| = 4$, $|\mathcal{K}| = 3$.

		Key		
		k_0	k_1	k_2
Message	m_0	c_0	c_1	c_2
	m_1	c_1	c_2	c_3
	m_2	c_2	c_3	c_0
	m_3	c_3	c_0	c_1

Now consider again the case of Table **??** but with the following non-uniform message distribution: $\rho(m_0) = 0.98$, and $\rho(m_1) = \rho(m_2) = 0.01$. In this case the message to be transmitted is m_0 with 98% probability and m_1 or m_2 with 1% probability each, and recall that the attacker knows both the encoder and the message and key distributions. Fix a pre-shared transmission key, for example assume the key is k_1. This means that the attacker will observe ciphertext c_2 98% of the time, which combined with the knowledge of the encoder and message distribution can be used to infer that the key is k_1 with very high probability, causing drastic message and key leakage after the observation of a small number of ciphertexts. In theory an encoder could be constructed that could account for such an uneven message distribution. This turns out to be solvable using cell encoders.

The introduction of cell encoders allows evening the probability of the ciphertexts when the message distribution is non-uniform, allowing for universally

optimal encoding. Consider message and key spaces of size 4 and 3 respectively: $\mathcal{M} = \{m_0, m_1, m_2, m_3\}$ and $\mathcal{K} = \{k_0, k_1, k_1\}$. Let $\rho(m_0) = \rho(m_1) = 0.3$ and $\rho(m_2) = \rho(m_3) = 0.2$, and let the key distribution be uniform. The probability of the message being m_0 or m_1 is hence one and a half times the probability of the message m_2 or m_3, and as explained above this can cause key and message leakage due to ciphertexts having different probabilities and non-uniformly-distributed cloaks. However, if the messages with higher probability had a lower probability of producing some of the ciphertexts compared to the messages with lower probability, this would even out ciphertext and cloak distributions allowing for a universally optimal encoding.

Consider the cell encoder depicted in Table **??**. Recall that each ciphertext can appear at mostly once in each column, otherwise Assumption **??** would be violated. Observe that each cell for messages m_0 and m_1 has one and a half times the ciphertexts of each cell for messages m_2 and m_3. This means that while m_0 has one and a half times the probability of being chosen compared to m_2, the probability that m_2 will produce a particular ciphertext, e.g. c_2, is one and a half times the probability that m_0 will produce the same ciphertext c_2. This evens out the probabilities for the ciphertexts being produced. It can in fact be verified that all the ciphertexts in $\mathcal{C} = \{c_0, \ldots, c_9\}$ have the *same* probability of being produced: the ciphertext distribution $\rho(\mathcal{C})$ is uniform.

Table 3: A cell encoder for $|\mathcal{M}| = 4$, $|\mathcal{K}| = 3$.

		Key		
		k_0	k_1	k_2
Message	m_0	c_0, c_1, c_2	c_3, c_4, c_5	c_6, c_7, c_8
	m_1	c_3, c_4, c_5	c_6, c_7, c_8	c_9, c_0, c_1
	m_2	c_6, c_7	c_9, c_0	c_2, c_3
	m_3	c_8, c_9	c_1, c_2	c_4, c_5

Showing that the cloak on the key of each ciphertext is of size three and uniformly distributed is trivial, since each ciphertext appears in each column exactly once and the key distribution is uniform. More importantly, the cloak on the message of each ciphertext also has size three and uniform distribution.

For example, consider again ciphertext c_2. Note that c_2 does not appear in the row of m_1, hence the probability that the message is m_1 if c_2 is observed is zero. If c_2 is observed then the message is m_0 iff the key is k_0, m_2 iff the key is k_2, and m_3 iff the key is k_1. Hence from the uniform distribution of the key we obtain that $\rho(m_0|c_2) = \rho(m_2|c_2) = \rho(m_3|c_2) = 1/3$, showing that the probability distribution of the cloak of c_2 on the message is also uniform on three elements. This holds for any ciphertext, showing that the cell encoder in Table **??** also has both universal message optimality and universally key optimality for any KM-entropy measure, making it the best possible encoder for the message and key spaces and distributions considered, i.e. universal optimality (Definition **??**).

The main intuition for the construction of such an encoder is that the number of ciphertexts for each message has to counterbalance the relative probability of that message compared to the others. This intuition is formalized in the Apollonian property described in the next section.

4.3 Apollonian Cell Encoders

This section introduces Apollonian cell encoders and their properties. The section concludes by proving that Apollonian cell encoders have universal optimality for all KM-entropy measures.

Assume that given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$ each message probability is rational, i.e. $\forall m_i \in \mathcal{M}.\ \rho(m_i) \in \mathbb{Q}$ and (with no loss of generality) is irreducible. Then each message probability $\rho(m_i)$ can be written as the quotient of a numerator n_i and a denominator d_i: $\rho(m_i) = \frac{n_i}{d_i}$ such that $GCD(n_i, d_i) = 1$. Let μ be the least common multiple of these denominators: $\mu = LCM(d_1, d_2, \ldots, d_{|\mathcal{M}|})$ and so define $n'_i = \frac{n_i \mu}{d_i}$.

By finding a common multiple of all the denominators it is possible to express all the message probabilities with a common denominator. In turn this allows choosing ciphertexts in a cell encoder that can even out the probabilities of the messages.

A cell encoder is *Apollonian* if it is semi-injective on both the key and message, and $|\mathcal{C}| = \mu$ and each cell corresponding to message m_i has size n'_i.

Definition 14 (Apollonian Cell Encoder). *Let ρ be a probability distribution over $\mathcal{M} \times \mathcal{K}$ such that for every $m_i \in \mathcal{M}$ then $\rho(m_i) \in \mathbb{Q}$. Let each $\rho(m_i) = \frac{n_i}{d_i}$ (where $\frac{n_i}{d_i}$ is irreducible) and $\mu = LCM(d_1, d_2, \ldots, d_{|\mathcal{M}|})$ and so define $n'_i = \frac{n_i \mu}{d_i}$. Then a cell encoder enc is* Apollonian *iff*

1. $\forall k \in \mathcal{K}.\ \mathrm{enc}(\cdot, k)$ *is injective and*
2. $\forall m \in \mathcal{M}.\ \mathrm{enc}(m, \cdot)$ *is injective and*
3. $|\mathcal{C}| = \mu$ *and*
4. $\forall m_i \in \mathcal{M}, k_j \in \mathcal{K}.\ |CELL(m_i, k_j)| = n'_i$.

The semi-injectivity properties of Apollonian cell encoders are used to ensure that no ciphertext can map to the same message with two different keys (or the same key with two different messages). This ensures that the cloak of any given ciphertext on the messages always has size $|\mathcal{K}|$ (and similarly that the cloak of any given ciphertext on the key has size $|\mathcal{K}|$). Lastly, the restriction on the size of the cells n'_i matching the numerator (for μ as the denominator) ensures that the probabilities are correctly evened out.

Optimality of Apollonian cell encoders.

This section proves optimality of Apollonian cell encoders. The proofs here rely on: the assumption that the key is uniformly distributed (Assumption **??**), and also that $H_\infty(M) \geq H_\infty(K)$ (this is proved necessary in Section **??**, Theorem **??**). If these assumptions hold, then an Apollonian cell encoder for the

message and key space always exists and achieves universal optimality (Definition **??**). See Section **??** for necessary and sufficient conditions for the existence of Apollonian cell encoders.

The following lemmata hold for Apollonian cell encoders and are used to prove universal optimality.

The following lemma proves that in an Apollonian cell encoder, each ciphertext c appears exactly once in the cells corresponding to each key k_j.

Lemma 3. $\forall k_j \in \mathcal{K}, \forall c \in \mathcal{C}.\ \sum_{m_i \in \mathcal{M}} \Delta(c \in CELL(m_i, k_j)) = 1.$

Proof. By conditions 1 and 3 of Definition **??**.

The following lemma proves that in an Apollonian cell encoder, each ciphertext c appears exactly once in the cells corresponding to each message m_i if and only if m_i is in the cloak of c on the message.

Lemma 4. $\forall m_i \in \mathcal{M}, \forall c \in \mathcal{C}.\ \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) = \Delta(m \in W_m(c)).$

Proof. By conditions 2 and 3 of Definition **??**.

The following lemma proves that in an Apollonian cell encoder, each ciphertext c appears a number of time equal to the size of the key space.

Lemma 5. $\forall c \in \mathcal{C}.\ \sum_{m_i \in \mathcal{M}} \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) = |\mathcal{K}|.$

Proof. By conditions 1 and 3 of Definition **??**.

The following lemma proves that in an Apollonian cell encoder, each ciphertext c has the same probability of appearing, and that probability corresponds to 1 divided by μ.

Lemma 6. $\forall c \in \mathcal{C}.\ \rho(c) = \frac{1}{\mu}.$

Proof.

$$
\begin{aligned}
\rho(c) &= \sum_{m_i \in \mathcal{M}} \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) \frac{\rho(m_i, k_j)}{|CELL(m_i, k_j)|} \quad \text{(by Equation (??))} \\
&= \sum_{m_i \in \mathcal{M}} \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) \frac{\rho(m_i, k_j)}{n'_i} \\
&\qquad \text{(by condition 4 of Definition ??)} \\
&= \sum_{m_i \in \mathcal{M}} \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) \frac{\rho(m_i)\rho(k_j)}{n'_i} \qquad \text{(by Assumption ??)} \\
&= \sum_{m_i \in \mathcal{M}} \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) \frac{\frac{n'_i}{\mu}\frac{1}{|\mathcal{K}|}}{n'_i} \\
&\qquad \text{(by definition of } n'_i \text{ and Assumption ??)} \\
&= \sum_{m_i \in \mathcal{M}} \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) \frac{1}{\mu|\mathcal{K}|} \\
&= \frac{1}{\mu|\mathcal{K}|} \sum_{m_i \in \mathcal{M}} \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) \ \text{(since } \mu \text{ and } |\mathcal{K}| \text{ are constants)} \\
&= \frac{1}{\mu|\mathcal{K}|}|\mathcal{K}| \qquad \text{(by Lemma ??)} \\
&= \frac{1}{\mu} \ .
\end{aligned}
$$

The following lemma proves that in an Apollonian cell encoder, given each key k, each ciphertext c has the same probability of appearing, and that probability corresponds to 1 divided by μ.

Lemma 7. $\forall k \in \mathcal{K}, c \in \mathcal{C}.\ \rho(c|k) = \frac{1}{\mu}$.

Proof.

$$\begin{aligned}
\rho(c|k_j) &= \sum_{m_i \in \mathcal{M}} \Delta(c \in CELL(m_i, k_j)) \frac{\rho(m_i)}{|CELL(m_i, k_j)|} && \text{(by Equation (??))} \\
&= \sum_{m_i \in \mathcal{M}} \Delta(c \in CELL(m_i, k_j)) \frac{\rho(m_i)}{n'_i} && \text{(by condition 4 of Definition ??)} \\
&= \sum_{m_i \in \mathcal{M}} \Delta(c \in CELL(m_i, k_j)) \frac{(\frac{n'_i}{\mu})}{n'_i} && \text{(by definition of } n'_i) \\
&= \sum_{m_i \in \mathcal{M}} \Delta(c \in CELL(m_i, k_j)) \frac{1}{\mu} \\
&= \frac{1}{\mu} \sum_{m_i \in \mathcal{M}} \Delta(c \in CELL(m_i, k_j)) && \text{(since } \mu \text{ is a constant)} \\
&= \frac{1}{\mu} \ . && \text{(by Lemma ??)}
\end{aligned}$$

The following lemma proves that in an Apollonian cell encoder, given a message m_i, each ciphertext c such that m_i is in the cloak $W_m(c)$ of c has the same probability of appearing, and that probability corresponds to 1 divided by n'_i times the size of the key space.

Lemma 8. $\forall c \in \mathcal{C}, m_i \in W_m(c).\ \rho(c|m_i) = \frac{1}{n'_i|\mathcal{K}|}$.

Proof.

$$\begin{aligned}
\rho(c|m_i) &= \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) \frac{\rho(k_j)}{|CELL(m_i, k_j)|} && \text{(by Equation ??)} \\
&= \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) \frac{\rho(k_j)}{n'_i} && \text{(by condition 4 of Definition ??)} \\
&= \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) \frac{1}{n'_i|\mathcal{K}|} && \text{(by Assumption ??)} \\
&= \frac{1}{n'_i|\mathcal{K}|} \sum_{k_j \in \mathcal{K}} \Delta(c \in CELL(m_i, k_j)) \\
& && \text{(since } n'_i \text{ is a constant for a given } m_i \text{ and } |\mathcal{K}| \text{ is a constant)} \\
&= \frac{1}{n'_i|\mathcal{K}|} \ . && \text{(by Lemma ?? since } m_i \in W_m(c))
\end{aligned}$$

The following lemma proves that in an Apollonian cell encoder, given a ciphertext c, the probability that the ciphertext was produced by a message m in the cloak $W_m(c)$ of c always corresponds to 1 divided by the size of the key space.

Lemma 9. $\forall c \in \mathcal{C}, m \in W_m(c).\ \rho(m|c) = \frac{1}{|\mathcal{K}|}$.

Proof.

$$\begin{aligned}
\rho(m|c) &= \frac{\rho(c|m)\rho(m)}{\rho(c)} && \text{(by Bayes' Theorem)} \\
&= \frac{\frac{1}{n_i'|\mathcal{K}|}\rho(m)}{\rho(c)} && \text{(by Lemma ??)} \\
&= \frac{\frac{1}{n_i'|\mathcal{K}|}\frac{n_i'}{\mu}}{\rho(c)} && \text{(by definition of } n_i') \\
&= \frac{\frac{1}{n_i'|\mathcal{K}|}\frac{n_i'}{\mu}}{\frac{1}{\mu}} && \text{(by Lemma ??)} \\
&= \frac{1}{|\mathcal{K}|} \quad .
\end{aligned}$$

The following theorems prove the optimality of Apollonian cell encoders:

Theorem 2. *Given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$ and a cryptosystem $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$ where enc is an Apollonian cell encoder, then the cryptosystem achieves universal message optimality (Definition ??).*

Proof. To prove this result it suffices to show that $\forall c \in \mathcal{C}, m \in W_m(c) \ : \rho(m|c) = \rho(k) = \frac{1}{|\mathcal{K}|}$. The theorem then follows from Lemma ??.

Theorem 3. *Given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$ and a cryptosystem $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$ where enc is an Apollonian cell encoder, then the cryptosystem achieves universal key optimality (Definition ??) with perfect key secrecy $H(K|C) = H(K)$.*

Proof. To prove this result it suffices to show that $\forall c \in \mathcal{C}, k \in \mathcal{K} \ : \ \rho(k|c) = \rho(k)$. Observe that $\rho(k|c) = \frac{\rho(c|k)\rho(k)}{\rho(c)}$ so it suffices to show that $\rho(c) = \rho(c|k)$. The theorem then follows from Lemma ??, Lemma ??, and uniformity of $\rho(K)$.

Observe that this does not require condition 2 of Definition ??.

Theorems ?? & ?? can be combined to show that Apollonian cell encoders are universally optimal.

Corollary 1. *Given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$ and a cryptosystem $(\mathcal{M}, \mathcal{K}, \mathcal{C}, \text{enc})$ where enc is an Apollonian cell encoder, then the cryptosystem achieves universal optimality (Definition ??).*

Proof. By Theorems ?? & ??.

5 Existence and Construction of Apollonian Cell Encoders

This section discusses the necessary and sufficient conditions for an Apollonian cell encoder to exist given a probability distribution on messages and a key

space, and presents a closed form to construct an Apollonian cell encoder when possible. As in the previous section, assume a given message space $\mathcal{M}$, key space $\mathcal{K}$, and ciphertext space $\mathcal{C}$.

Given a distribution ρ over the message space $\mathcal{M}$ and a key space $\mathcal{K}$ such that $\forall m \in \mathcal{M}.\ \rho(m) \in \mathbb{Q}$ and $\forall k \in \mathcal{K}.\ \rho(k) = \frac{1}{|\mathcal{K}|}$, this section proves that an Apollonian cell encoder exists iff $H_\infty(M) \geq H_\infty(K)$.

5.1 Necessity

The following two lemmata are used in the proof of necessary conditions for the existence of an Apollonian cell encoder. The first to clarify the underlying relation between message distribution and key distribution. The second to show necessity of the condition $H_\infty(M) \geq H_\infty(K)$.

Lemma 10. *Given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$, then $H_\infty(M) \geq H_\infty(K) \equiv max_{m_i \in \mathcal{M}}(\rho(m_i)) \leq \frac{1}{|\mathcal{K}|}$*

Proof.

$$H_\infty(M) \geq H_\infty(K) \equiv -\log(max_{m_i \in \mathcal{M}}(\rho(m_i))) \geq \log(|\mathcal{K}|)$$

(by Assumption **??**)

$$\equiv -\log(max_{m_i \in \mathcal{M}}(\rho(m_i))) \geq -\log\left(\frac{1}{|\mathcal{K}|}\right)$$

$$\equiv \log(max_{m_i \in \mathcal{M}}(\rho(m_i))) \leq \log\left(\frac{1}{|\mathcal{K}|}\right)$$

$$\equiv max_{m_i \in \mathcal{M}}(\rho(m_i)) \leq \frac{1}{|\mathcal{K}|}$$

Theorem 4. *Given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$ such that $H_\infty(M) < H_\infty(K)$, then no Apollonian encoder can exist for ρ.*

Proof. As a trivial corollary of Lemma **??**, it holds that $H_\infty(M) < H_\infty(K) \equiv max_{m_i \in \mathcal{M}}(\rho(m_i)) > \frac{1}{|\mathcal{K}|}$. Let message m_j be a message with the highest probability, so $\rho(m_j) = max_{m_i \in \mathcal{M}}(\rho(m_i)) = \frac{n'_j}{\mu}$ (where μ is defined as in Section **??**). Now consider all the cells of the encoder of the form $CELL(m_j, \cdot)$. Observe that to define such cells requires $n'_j|\mathcal{K}|$ distinct ciphertexts. However, since $\frac{n'_j}{\mu} > \frac{1}{|\mathcal{K}|}$ and thus $n'_j|\mathcal{K}| > \mu$ and thus $n'_j|\mathcal{K}| > |\mathcal{C}|$ (by condition 4 of Definition **??**) it is impossible to define such cells without repeating a ciphertext and so contradicting condition 2 of Definition **??**.

Thus, if ρ is such that $H_\infty(M) < H_\infty(K)$ holds no Apollonian cell encoder can exist for ρ.

Observe that the condition $H_\infty(M) \geq H_\infty(K)$ induces a limit on the key space size based upon the message distribution. If $\frac{n}{d} = \max_{m \in \mathcal{M}} \rho(m)$ then the size of the key space is bounded by w where $\frac{1}{w} \geq \frac{n}{d} > \frac{1}{w+1}$ holds. So for instance if the highest message probability is 1/8 no Apollonian encoder can exist for a key space of size larger than 8.

5.2 Sufficiency

This section formalises how to construct an Apollonian encoder given ρ.

First assign to each message m a set of n' sequential base numbers $\mathcal{B} = b_0, b_1, \ldots, b_{n'-1}$ such that $\forall i, j.\ m_i \neq m_j \Rightarrow \mathcal{B}_i \cap \mathcal{B}_j = \emptyset$. For simplicity this can be done by ordering the messages and then assigning to m_0 the base numbers $0, 1, \ldots, n'_0 - 1$, then to m_1 the base numbers $n'_0, n'_0 + 1, \ldots, n'_0 + n'_1 - 1$, etc. Observe that all of these base numbers can be assigned to exactly range over 0 to $\mu - 1$.

For simplicity let $x = \mu \cdot max_{m_i \in \mathcal{M}}(\rho(m_i)) = max(n'_i)$, i.e. the numerator of the representation of $\rho(m_i)$ in the form $\frac{n'_i}{\mu}$. Now an Apollonian cell encoder $enc(k, m_i)$ can be defined by

$$enc(k, m_i) = (b_i + kx)\%\mu$$

where the keys are represented as their index $k_j = j$ (respecting $0 \leq j \leq |\mathcal{K}|-1$), b_i is chosen uniformly at random from the set $\mathcal{B}_i$ of base numbers assigned to m_i, and $\%$ is the modulo operation. Note that here the natural numbers $0, 1, \ldots, \mu-1$ are used as the ciphertexts $\mathcal{C}$.

Theorem 5. *The encoder $enc(k, m_i) = (b_i + kx)\%\mu$ is Apollonian.*

Proof. The proof relies on proving that the encoder has the four conditions in the definition of an Apollonian cell encoder (Definition **??**).

For condition 4, observe that each cell corresponding to the message m_i has size n'_i by construction.

For condition 3 observe that the ciphertexts are the numbers $0, 1, ..., \mu - 1$ and so $|\mathcal{C}| = \mu$ by construction.

For condition 1 consider the union of all the cells of the encoder of the form $CELL(\cdot, k_j)$ for any $k_j \in \mathcal{K}$. This union contains exactly one instance of each ciphertext since when $k_j = 0$ this is exactly the numbers $0, 1, \ldots, \mu - 1$, and this is stable under addition-modulo for any $k_j \cdot x$.

The only non-trivial condition is 2: that no ciphertext is repeated in the union of all cells of the form $CELL(m_i, \cdot)$ for any $m_i \in \mathcal{M}$. Fix some message $m_i \in \mathcal{M}$ and let $b_\alpha = \min_{b \in \mathcal{B}_i}(b)$ and $b_\omega = b_\alpha + n'_i - 1 = \max_{b \in \mathcal{B}_i}(b)$. Observe that when $k = 0$ then the possible ciphertexts are some numbers $b_\alpha, b_\alpha + 1, \ldots, b_\omega$ that are sequential and distinct. Now consider when k_1 and k_2 are both in the range $0, \ldots, |\mathcal{K}| - 1$ inclusive and $k_1 < k_2$. Observe that the maximum ciphertext for k_1 is $b_\omega + k_1 x$ and the minimum for k_2 is $b_\alpha + k_2 x$. It is sufficient to show that $b_\omega + k_1 x < b_\alpha + k_2 x$ (since we can adjust by subtracting $(b_\omega + k_1 x)$ to avoid

issues with $\%\mu$):

$$\begin{aligned}
b_\omega + k_1 x < b_\alpha + k_2 x &\equiv b_\omega + k_1 x < b_\alpha + k_1 x + nx && \text{(for } n \geq 1 \text{ since } k_1 < k_2\text{)} \\
&\equiv b_\omega < b_\alpha + nx \\
&\equiv b_\alpha + n'_i - 1 < b_\alpha + nx && \text{(by definition of } b_\omega\text{)} \\
&\equiv n'_i - 1 < nx \\
&\equiv n'_i < nx + 1 \\
&\equiv n'_i < \min(n)x + 1 \\
&\equiv n'_i < x + 1 && \text{(since } n \geq 1 \text{ by definition)} \\
&\equiv n'_i \leq x
\end{aligned}$$

which must hold since $x = max(n'_i)$. Finally, to conclude requires showing that $(b_\omega - b_\alpha) + (|\mathcal{K}| - 1)x < \mu$:

$$\begin{aligned}
(b_\omega - b_\alpha) + (|\mathcal{K}| - 1)x < \mu &\equiv (b_\alpha + n'_i - 1 - b_\alpha) + (|\mathcal{K}| - 1)x < \mu \\
& && \text{(by definition of } b_\omega\text{)} \\
&\equiv (n'_i - 1) + (|\mathcal{K}| - 1)x < \mu \\
&\equiv (n'_i - 1) + (|\mathcal{K}| - 1)x < |\mathcal{K}|x && \text{(by definition of } \mu\text{)} \\
&\equiv n'_i - 1 < x \\
&\equiv n'_i \leq x
\end{aligned}$$

which always holds since $x = max(n'_i)$ by definition.

This shows that given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$ such that $\forall m \in \mathcal{M}.\ \rho(m) \in \mathbb{Q} \wedge \rho(m) \leq \frac{1}{|\mathcal{K}|}$, then an Apollonian cell encoder can be constructed.

Combining Theorems **??** and **??** yields the necessary and sufficient conditions for the existence of an Apollonian cell encoder given probability distribution ρ over the message space $\mathcal{M}$ and key space $\mathcal{K}$ such that $\rho(\mathcal{K})$ is uniform.

Theorem 6. *Given a probability distribution ρ on $\mathcal{M} \times \mathcal{K}$ such that $\rho(\mathcal{K})$ is uniform, then an Apollonian cell encoder for ρ exists iff $H_\infty(M) \geq H_\infty(K)$.*

Proof. By Theorem **??** and Theorem **??**.

6 Discussion

This section discusses the limitations of the rational valued probabilities, and practical concerns for implementation of Apollonian cell encoders. These include:

- approximating real-valued probabilities;
- size of the ciphertext space for practical transmission;
- compact representation of an Apollonian cell encoder for practical use.

It is advised that the reader familiarizes them self with the rest of the paper before reading this section.

6.1 Approximating Real-Valued Message Probabilities

The assumption that the message probabilities are rational numbers in Section **??** is necessary to compute their least common multiplier μ and build an Apollonian cell encoder. Assuming the message probabilities have to be represented in a physical machine, this is normally not a problem.

In the theoretical sense, any irrational message probability can be approximated to a rational number with arbitrary precision in the usual manner. However, due to the generality of the definition of KM-entropy, it is *not* possible to guarantee in general that such an approximation will not result in a significant increase in leakage for some KM-entropy. It is always possible to construct a pathological KM-entropy measure that happens to be arbitrarily steep exactly on the values of some of the message probabilities, thus magnifying the effect of even very small approximation on the resulting entropy and leakage.

This prevents theoretical guarantees being made about the effect of approximating irrational probabilities to rational ones in general. However, such guarantees can be made if limiting to a subset of sufficiently smooth KM-entropy measures.

However, given a suitably smooth entropy measure and an arbitrary level of precision, a good approximation can be made using rational values for the irrational probabilities. The only practical concern is that this will reduce equivocation some extremely small amount. Thus, both message and key equivocation will be reduced, and unlimited key reuse is no longer safe. This can be mitigated by calculating the maximal leakage (on the key since the key is the limiting variable), and then a safe over-approximation used to ensure a minimum level of message or key equivocation is maintained before a fresh key must be used.

6.2 Ciphertext Space Size

One practical concern is the size of the ciphertext space of an Apollonian cell encoder. Observe that the size of the ciphertext space is μ, and that μ is computed from the probabilities of the messages. Thus, it is straightforward to find (or produce) a message distribution that induces an extremely large μ and so requires an extremely large ciphertext space $\mathcal{C}$.

Fortunately it is straightforward to represent the ciphertexts in a form that is logarithmic w.r.t. μ. In practice this is achieved in the standard manner of representing the natural numbers in base i and thus yielding a ciphertext representation of size $\lceil \log_i \mu \rceil$. For example, if $\mu = 31$ and $i = 2$ then the ciphertexts can be represented by $\lceil \log_2 31 \rceil = 5$ bits. (This assumes that a consistent ciphertext size representation is desirable, otherwise if 0 and 00 are considered distinct, the representation can be even more compact.)

If the above approach alone still yields insufficiently small ciphertext representation, then the probabilities in ρ can be approximated as in Section **??**. Again, this may reduce equivocation and increase leakage for pathological KM-entropy cases, but will hardly be a problem in practice.

6.3 Compact Representation

One practical challenge for encoders is a compact representation that can be exploited to implement the encoder in a physical device. Since in general an encoder is a $|\mathcal{M}| \times |\mathcal{K}|$ matrix that maps from message and key to (a set of) ciphertext(s), finding a more compact representation is significant.

Further to the existence of an optimal Apollonian encoder, the constructive approach (Theorem **??**) provides a highly compact representation. This requires only a $1 \times \mu$ matrix representation, a source of randomness, and the formula $(b+kx)\%\mu$ where b is derived from the matrix and randomness, and x and μ from the matrix. This allows an optimal Apollonian cell encoder for a given message and key space to have a compact representation. Note that for an Apollonian cell encoder it holds that $|\mathcal{C}| = \mu$, hence the reduction of the ciphertext space discussed in Section **??** also implies a reduction of the encoder's representation.

7 Conclusions & Future Work

Finding a good unconditionally secure cryptosystem has been an open problem for some time. Such a cryptosystem must have good equivocation for the message, good equivocation for the key, and a practical implementation. This paper solves this by presenting the necessary and sufficient conditions for a universally optimal unconditional cryptosystem: an Apollonian cell encoder.

The paper presents the definition of universal optimality for a cryptosystem; achieving message and key max-equivocation for a very general class of entropy measures (KM-entropy). For all such entropy measures, and under some very reasonable assumptions, Apollonian cell encoders provide the maximum possible security: max-equivocation for the message, and perfect secrecy for the key. This paper proves that Apollonian cell encoders exist iff $H_\infty(M) \geq H_\infty(K)$, i.e. if the min-entropy of the message is larger than the min-entropy of the key. Further, the proof of existence is constructive and demonstrates a compact representation that allows Apollonian cell encoders to be easily implemented in practice.

7.1 Future Work

In a communication protocol scenario an attacker may execute attacks other than merely attempting to guess the messages and keys by observing ciphertexts. These attacks include: attempting to forge messages, replay messages, and dropping messages, all of which could cause disruption to the receiver. Future work includes finding protections against such attacks by an attacker with unlimited computational power.

Another similar line of research is to consider known-plaintext attacks where the attacker chooses the plaintext, observes the ciphertext, and then attempts to guess the key. This is trivial for an attacker with unlimited computational power at the moment, but perhaps new techniques can be developed to hide the key against such attacks.

A more practical future direction is to provide a reference implementation of Apollonian cell encoders. This could also include the other future works to build a practical implementation that has all the expected protections of a secure communication protocol against attackers with unlimited computational power.

More theoretically curious would be consider the construction of encoders that are universally optimal when both the message and key distributions are non-uniform. Unfortunately it appears unlikely that such a (cell) encoder would be able to achieve perfect secrecy on either message or key in this setting, and so such an encoder would be of limited practical use.