How to Fool an Unbounded Adversary with a Short Key

Alexander Russell and Hong Wang

Department of Computer Science and Engineering,
University of Connecticut
Storrs, Connecticut 06269, USA
acr@engr.uconn.edu, hongmuw@engr.uconn.edu

Abstract. We consider the symmetric encryption problem which manifests when two parties must securely transmit a message m with a short shared secret key. As we permit arbitrarily powerful adversaries, any encryption scheme must leak information about m—the mutual information between m and its ciphertext cannot be zero. Despite this, we present a family of encryption schemes which guarantee that for any message space in $\{0,1\}^n$ with minimum entropy $n-\ell$ and for any Boolean function $h : \{0,1\}^n \to \{0,1\}$, no adversary can predict $h(m)$ from the ciphertext of m with more than $1/n^{\omega(1)}$ advantage; this is achieved with keys of length $\ell+\omega(\log n)$. In general, keys of length $\ell+s$ yield a bound of $2^{-\Theta(s)}$ on the advantage. These encryption schemes rely on no unproven assumptions and can be implemented efficiently.

1 Introduction

One of the simplest and most secure encryption systems is the *one time pad*: two parties who have agreed on a uniformly selected secret key $s \in \{0,1\}^n$ can exchange a single message $m \in \{0,1\}^n$ by transmitting $m \oplus s$, this parity being taken componentwise. If we think of the message m and the secret key s as independent random variables, then it is easy to see that the message m and the ciphertext $m \oplus s$ are uncorrelated: we say that this encryption system offers *perfect secrecy*.

One unfortunate consequence of this absolute security guarantee is that any such system must use a fresh secret key $s \in \{0,1\}^n$ for each new message of length n. Indeed, regardless of the system employed, if a uniformly selected message $m \in \{0,1\}^n$ is encrypted with a key of length $k < n$, then at least $n-k$ bits of "information" about m have leaked into the ciphertext. (See, e.g., [25] for a formal discussion of message equivocation.) Despite this, we construct a family of encryption systems utilizing short keys which guarantee that for any message space with sufficient min-entropy, no adversary can predict any Boolean function of the message m with non-negligible advantage; specifically, if the message space has min-entropy $n-\ell$, and secret keys of length $\ell+s$ are utilized, no Boolean function can be predicted with advantage $2^{-\Theta(s)}$. These systems rely on no unproven assumptions, and encryption (and decryption) can be computed efficiently. The precise notion of security is described below.

Of course, if a pseudorandom generator exists, then it is possible to construct encryption systems with satisfactory security guarantees against resource-bounded adversaries, even when the length of the message m exceeds the length of the key. A traditionally accepted notion of security in this resource-bounded case is that of *semantic security* [11], though a number of stronger (and important) notions exist (see, e.g., [4, 9, 17, 19]). A system with semantic security guarantees that observation of $E(m)$, the encryption of a message m, offers essentially no advantage to a bounded adversary in predicting any Boolean function of the message m. (This Boolean function may be some specific bit of m, or, perhaps, a complicated function capturing some global property of m.) Furthermore, this guarantee is offered *regardless of the a priori distribution of the message m.* In the last section of the paper, we discuss some potential applications of the information-theoretic encryption systems of Sects. 3 and 4 to this complexity-theoretic framework. Specifically, we observe that a hybrid approach can reduce the complexity of the resulting system at the expense of weakening (in a controlled fashion) the notion of semantic security. Finally, we mention that if the adversary is space-limited and the parties have access to a long public random string, strong privacy guarantees can be obtained with short keys [15, 2].

Returning to the case for unbounded adversaries, we say that an encryption system offers *entropically bounded security* if for all message distributions with sufficient min-entropy, and all pieces of partial information $h : \{0,1\}^* \to \{0,1\}$, observation of the ciphertext of m offers no adversary non-negligible advantage in prediction of $h(m)$. If the definition is strengthened so that it applies for all message spaces and the error terms in the advantage are removed, then we exactly recover the definition of perfect secrecy. (See the next section for precise definitions.) Initially, we give a simple encryption system offering entropic security in the case when the adversary has *no* a priori information about the message (i.e., the message distribution is uniform); the scheme can be realized with keys of length $\omega(\log n)$. We then show that for message spaces with min-entropy $n - \ell$, an encryption system offering entropically bounded security can be realized with keys of length $\ell + \omega(\log n)$.

The two main theorems in the article, Theorem 2 and Theorem 3, are both instantiations of common paradigms in cryptography. The first is an information-theoretic variant of the standard practice of encrypting a short seed which is then used for a pseudorandom generator (in our case, this will be an ϵ-biased space). The second is a variant of the "simple embedding schemes" often used in practice, where a message is encrypted by applying a one-way permutation after a suitable (bijective) hash function. The system of Bellare and Rogaway [5] is also theoretical evidence for the quality of such schemes.

In Sect. 2 we give basic definitions, including a brief discussion of ϵ-biased spaces, universal hash functions, and the Fourier transform over $\mathbb{Z}_2^n$, which will be used in the main results, presented in Sects. 3 and 4. In Sect. 5, we discuss some applications of these theorems to resource-bounded encryption systems.

2 Definitions

Definition 1. *A pair* $(\mathcal{E}, \mathcal{D})$ *is a* symmetric encryption system *with parameters* (ℓ_s, ℓ_e) *if*

1. $\ell_s : \mathbb{N} \to \mathbb{N}$ *and* $\ell_e : \mathbb{N} \to \mathbb{N}$ *are functions, determining the length of the secret key and the length of the encryption for messages of length* n*,*
2. $\mathcal{E} = \{E_n : \{0,1\}^n \times \{0,1\}^{\ell_s(n)} \to \{0,1\}^{\ell_e(n)} \mid n \geq 1\}$ *is the family of encryption functions, and*
3. $\mathcal{D} = \{D_n : \{0,1\}^{\ell_e(n)} \times \{0,1\}^{\ell_s(n)} \to \{0,1\}^n \mid n \geq 1\}$ *is the family of decryption functions,*

so that for all $n \geq 1$*,* $m \in \{0,1\}^n$*, and* $s \in \{0,1\}^{\ell_s(n)}$*,* $D_n(E_n(m, s), s) = m$*. When the length* n *of the message can be inferred from context, we write* $E(m, s)$ $(D(m, s))$ *rather than* $E_n(m, s)$ $(D_n(m, s))$*. When an encryption system is clear from context, we let* S_n *denote the random variable uniform on the set* $\{0,1\}^{\ell_s(n)}$*.*

As defined above, encryption and decryption are deterministic; in Sect. 4 we shall consider the case when the encryption algorithm may depend on some private randomness.

Definition 2. *A* message space $\mathcal{M}$ *is a sequence of random variables* $\mathcal{M} = \{M_n \mid n \geq 1\}$ *so that* M_n *takes values in* $\{0,1\}^n$*. (When we couple* $\mathcal{M}$ *with an encryption system, we always assume that* M_n *and* S_n *are independent.)*

A symmetric encryption system $(\mathcal{E}, \mathcal{D})$ with parameters (ℓ_s, ℓ_e) is said to possess *perfect secrecy* if for all message spaces $\mathcal{M}$, all $n > 0$, and all $e \in \mathbf{im}\, E_n$, $\mathsf{Pr}[M_n = m] = \mathsf{Pr}[M_n = m \mid E(M_n, S_n) = e]$. An equivalent definition of perfect secrecy is the following:

Definition 3 (Perfect Secrecy). *A symmetric encryption system* $(\mathcal{E}, \mathcal{D})$ *with parameters* (ℓ_s, ℓ_e) *is said to possess* perfect secrecy *if for all* $\mathcal{M}$*, all* $n \geq 1$*, and all functions* $f : \{0,1\}^{\ell_e(n)} \to \{0,1\}$*, there is a random variable* G_f*, independent of* M_n*, so that for every* $h : \{0,1\}^n \to \{0,1\}$*,*

$$\mathsf{Pr}[f(E(M_n, S_n)) = h(M_n)] = \mathsf{Pr}[G_f = h(M_n)] \ .$$

Intuitively, this asserts that if there is an adversary f which can predict some Boolean function of m based on the ciphertext of m, then there is another adversary G_f which can predict this same Boolean function of m *without* even witnessing the ciphertext. (If one suitably changes this definition so that the function f and the random variable G_f are polynomial time computable and allows for negligible error, then one obtains the notion of semantic security.)

A random variable M_n taking values in $\{0,1\}^n$ has *min-entropy* $n - \ell$ when $\forall m_o \in \{0,1\}^n$, $\mathsf{Pr}[M_n = m_o] \leq 2^{-n+\ell}$. A message space $\mathcal{M}$, is said to have min-entropy $n - \ell(n)$ when the random variable M_n possesses this property for each n.

Definition 4. *We say that an encryption system possesses* $\ell(n)$-entropic security *if for every message space* $\mathcal{M}$ *with min-entropy* $n - \ell(n)$*, every* $n > 0$*, and all functions* $f : \{0,1\}^{\ell_e(n)} \to \{0,1\}$*, there is a random variable* G_f*, independent of* M_n*, so that for every* $h : \{0,1\}^n \to \{0,1\}$*,*

$$|\mathsf{Pr}[f(E(M_n)) = h(M_n)] - \mathsf{Pr}[G_f = h(M_n)]| = n^{-\omega(1)} \ .$$

Observe that if no constraint is placed on the min-entropy in $\mathcal{M}$ and the $n^{-\omega(1)}$ error term is removed, we recover the definition of perfect secrecy. We will construct two encryption systems, $(\mathcal{E}^u, \mathcal{D}^u)$ and $(\mathcal{E}^k, \mathcal{D}^k)$, so that

- E^u possesses 0-entropic security (i.e., provides security when the message space is uniform) and uses keys of length $w(n) \log n$, where $w(n)$ is any function tending to infinity. E^u (and D^u) can be computed in time $O(w(n) n \log^{1+c} n)$ for any $c > 0$.
- E^k possesses ℓ-entropic security (i.e., provides security when the message space has min-entropy $n - \ell$) so long as $\ell(n) \leq k(n) - \omega(\log n)$ and uses keys of length $k(n)$. E^k (and D^k) can be computed in time $O\left(n \log^2 n \log\log n\right)$.

These constructions make use of ϵ-biased sample spaces and universal hash functions, defined below.

2.1 ϵ-biased Sample Spaces

Definition 5. *A set* $S \subseteq \{0,1\}^n$ *is called* ϵ-biased *(or an* ϵ-biased sample space*) if for all nonempty* $\alpha \subset [n] = \{1, \ldots, n\}$, $\left|\mathsf{Exp}_{s \in S}\left[\prod_{a \in \alpha} (-1)^{s_a}\right]\right| \leq \epsilon$.

Small sets with these properties were initially constructed by Naor and Naor [16] and Peralta [18]. We will use a construction, due to Alon, Goldreich, Håstad and Peralta [1], which gives an ϵ-biased sample space in $\{0,1\}^n$ of size about $(\frac{n}{\epsilon})^2$. The sample space is given as the image of a certain function $\sigma_{n,m} : \mathbb{F}_{2^m} \times \mathbb{F}_{2^m} \to \{0,1\}^n$. (Here $\mathbb{F}_{2^n}$ denotes the finite field with 2^n elements.) To define σ, let $\text{bin} : \mathbb{F}_{2^m} \to \{0,1\}^m$ be a bijection satisfying $\text{bin}(0) = 0^m$ and $\text{bin}(x + y) = bin(x) \oplus bin(y)$, where $\alpha \oplus \beta$ denotes the componentwise exclusive or of α and β. Then $\sigma(x, y) = r = (r_0, \ldots, r_{n-1})$, where $r_i = \langle bin(x^i), bin(y)\rangle_2$, the inner product, modulo two, of x^i and y. The size of the sample space is 2^{2m}. Let $S_{n,m} \subset \{0,1\}^n$ be the collection of points so defined. They show that

Theorem 1 ([1]). $S_{n,m} = \textbf{\textit{im}}\, \sigma_{m,n}$ *is* $\frac{n-1}{2^m}$*-biased.*

Observe that when $m = \lceil \log n\epsilon^{-1} \rceil$, $\frac{n-1}{2^m} \leq \epsilon$. As elements of $S_{m,n}$ are constructed during the encryption (and decryption) phase of the 0-entropic encryption system, we analyze the complexity of computing the function above. First, we need to find an irreducible polynomial p of degree m over the finite field $\mathbb{F}_2$. As the degree of the polynomial will correspond to the quantity ϵ, we can be somewhat flexible concerning the degree of the irreducible polynomial and use an explicit family: for each $c \in \mathbb{N}$, the polynomial $p_c(x) = x^{2m} + x^m + 1$, where $m = 3^c$, is irreducible over $\mathbb{F}_2$. (See [14, Exercise 3.96].) Computation of $\sigma = \sigma_{m,n}$ for a pair (x, y) is performed on a component by component basis: given x^i, computation of x^{i+1} requires a single multiplication in $\mathbb{F}_{2^m} \cong \mathbb{F}_2[x]/(p_c)$. Using fast polynomial multiplication, computing this product takes $O(m \log m \log \log m)$ time (see [22], or the discussion in [3, p. 232]). As p_c is sparse (it has only 3 nonzero terms), reducing this result modulo p_c requires $O(m)$ time. Hence computation of $\sigma(x, y)$ requires $O(nm \log m \log \log m))$ time. In order for $S = \mathbf{im}\ \sigma$ to be ϵ-biased, we take m to be the smallest integer of form $2 \cdot 3^c$ larger than $\lceil \log(n/\epsilon) \rceil$; in this case the above running time is $O(n \log(n/\epsilon) \log \log(n/\epsilon) \log \log \log(n/\epsilon))$. To simplify notation, we let $\sigma_{n,\epsilon}$ denote $\sigma_{n,m}$ for this value of m.

2.2 k-wise Independent Permutations

Definition 6. *A family of permutations $\mathcal{P} \subset \{f : X \to X\}$ is a family of* k-wise independent permutations *[24] if for all distinct $s_1, \ldots, s_k \in X$ and all distinct $t_1, \ldots, t_k \in X$, $\mathsf{Pr}_{\phi \in \mathcal{P}} [\forall i, \phi(s_i) = t_i] = \prod_{i=0}^{t-1} \frac{1}{|X|-i}$.*

We will use a family of 3-wise independent permutations, described below. See Rees [20] for a more detailed description.

Let V be a two-dimensional vector space over $\mathbb{F}$, a finite field. For two nonzero vectors $\boldsymbol{v}$ and $\boldsymbol{w}$ in this space, we write $\boldsymbol{v} \sim \boldsymbol{w}$ when $\boldsymbol{v} = c\boldsymbol{w}$ for some $c \in \mathbb{F}$ (so that the two vectors span the same one-dimensional subspace). This is an equivalence relation; we write $[\boldsymbol{v}]$ for the equivalence class containing $\boldsymbol{v}$. Projective 2-space over $\mathbb{F}$ is then $P_2(\mathbb{F}) = \{[\boldsymbol{v}] \mid \boldsymbol{v} \neq \mathbf{0}\}$. We let $\mathrm{GL}_2(\mathbb{F})$ denote the set of non-singular 2×2 matrices over $\mathbb{F}$, and $\mathrm{PGL}_2(\mathbb{F}) = \mathrm{GL}_2(\mathbb{F})/\{cI | c \in \mathbb{F}\}$, where I is the identity matrix. An element ϕ of $\mathrm{PGL}_2(\mathbb{F})$ acts on $P_2(\mathbb{F})$ in a natural (and well-defined) way, mapping $[\boldsymbol{v}]$ to $[\phi(\boldsymbol{v})]$. It is not difficult to show that for any distinct $[\boldsymbol{u_1}], [\boldsymbol{u_2}], [\boldsymbol{u_2}] \in P_2(\mathbb{F})$ and any distinct $[\boldsymbol{v_1}], [\boldsymbol{v_2}], [\boldsymbol{v_2}] \in P_2(\mathbb{F})$, there is in fact a unique $\phi \in \mathrm{PGL}_2(\mathbb{F})$ so that $\phi([\boldsymbol{u_i}]) = [\boldsymbol{v_i}]$ for each i. In particular, $\mathrm{PGL}_2(\mathbb{F})$ is a 3-wise independent family of permutations. As multiplication and inversion in a finite field $\mathbb{F}_p$, for a prime p, may be accomplished in time $O(\log p (\log \log p)^2 \log \log \log p)$ time [23, 21], evaluation of an element $\phi \in \mathrm{PGL}_2(\mathbb{F}_p)$ also has this complexity.

Proposition 1. *$PGL_2(\mathbb{F}_p)$ is a 3-wise independent set of permutations of $P_2(\mathbb{F}_p)$.*

2.3 Fourier Analysis of Boolean Functions

Let $L(\mathbb{Z}_2^n) = \{f : \mathbb{Z}_2^n \to \mathbb{R}\}$ denote the set of real valued functions on $\mathbb{Z}_2^n = \{0,1\}^n$. Though our interest shall be in Boolean functions, it will be temporarily convenient to consider this richer space. $L(\mathbb{Z}_2^n)$ is a vector space over $\mathbb{R}$ of dimension 2^n, and has a natural inner product: for $f, g \in L(\mathbb{Z}_2^n)$, define $\langle f, g \rangle = 2^{-n} \sum_{x \in \{0,1\}^n} f(x)g(x)$. For a subset $\alpha \subset \{1, \ldots, n\}$, define the function $\chi_\alpha : \{0,1\}^n \to \mathbb{R}$ so that $\chi_\alpha(x) = \prod_{a \in \alpha} (-1)^{x_a}$. These functions χ_α are the *characters* of $\mathbb{Z}_2^n = \{0,1\}^n$. Among their many wonderful properties is the fact that *the characters form an orthonormal basis for* $L(\mathbb{Z}_2^n)$. To see this, observe that $\forall \alpha \subset [n]$, $\sum_{x \in \{0,1\}^n} \chi_\alpha(x) = 2^n$ when $\alpha = \emptyset$, and 0 otherwise. Furthermore, for $\alpha, \beta \subset [n]$, $\chi_\alpha(x)\chi_\beta(x) = \chi_{\alpha \oplus \beta}(x)$, where $\alpha \oplus \beta$ denotes the symmetric difference of α and β, so that $\langle \chi_\alpha, \chi_\beta \rangle = 1$ when $\alpha = \beta$, and 0 otherwise. Considering that there are 2^n characters, pairwise orthogonal, they span $L(\mathbb{Z}_2^n)$, as promised. Any function $f : \{0,1\}^n \to \mathbb{R}$ may then be written in terms of this basis: $f = \sum_\alpha \widehat{f}_\alpha \chi_\alpha$, where $\widehat{f}_\alpha = \langle f, \chi_\alpha \rangle$ is the projection of f onto χ_α. These coefficients $\widehat{f}_\alpha$, $\alpha \subset [n]$, are the *Fourier coefficients* of f, and, as we have above observed, uniquely determine the function f.

Given the above, it is easy to establish the *Plancherel* equality:

Proposition 2. *Let $f \in L(\mathbb{Z}_2^n)$. Then $\sum_\alpha \widehat{f}_\alpha^2 = \frac{1}{2^n} \sum_x f(x)^2$.*

As always, $\widehat{f}_\emptyset = \mathsf{Exp}[f]$ and, when the range of f is $\{\pm 1\}$, $\sum_\alpha \widehat{f}_\alpha^2 = \|f\|_2^2 = 1$.

3 Security for Uniformly Distributed Message Spaces

We begin by constructing a simple encryption system offering security in the case when the adversary has no a priori knowledge concerning the message (i.e., the message space is uniform). The next section will develop a more flexible encryption system which can tolerate general message distributions, so long as they have sufficient min-entropy.

Theorem 2. *Let $(\mathcal{E}^u, \mathcal{D}^u)$ be the encryption system given by $D_n^u(m,s) = E_n^u(m,s) = m \oplus \sigma_{n,\epsilon}(s)$ where $|m| = n$ and s is selected randomly in the domain of $\sigma_{n,\epsilon}$ (so $|s| = O(\log n\epsilon^{-1})$). Then for $\epsilon = n^{-\omega(1)}$, this encryption system offers 0-entropic security. E_n^u and D_n^u can be computed in time $O(n \log(n/\epsilon) \log\log(n/\epsilon) \log\log\log(n/\epsilon))$.*

Proof. For $n \geq 1$, let M_n be uniform on $\{0,1\}^n$; when n is understood we drop the subscripts. For simplicity, we treat h as a function with range $\{\pm 1\}$ rather than $\{0,1\}$. We must show that for every $f : \{0,1\}^n \to \{\pm 1\}$, there is a random variable G_f, independent of M, so that for all functions $h : \{0,1\}^n \to \{\pm 1\}$

$$|\mathsf{Pr}[f(M \oplus \sigma(S)) = h(M)] - \mathsf{Pr}[G_f = h(M)]| \leq n^{-\omega(1)} \ .$$

(Here S is uniform on the domain of $\sigma_{n,\epsilon}$.) The random variable G_f is defined in terms of the function f; though G_f is independent of M, G_f will predict $h(M)$ nearly as well as does f. Let M' be uniform on $\{0,1\}^n$ and independent of M (and S). Define $G_f = f(M')$; then $\mathsf{Pr}[G_f = h(M)] = \mathsf{Pr}[f(M') = h(M)]$. We begin with a lemma providing an upper bound on this prediction probability which is independent of f:

Lemma 1. *Let $T(m) = \mathsf{Exp}_S[h(m \oplus \sigma(S))] - \mathsf{Exp}_{M'}[h(M')]$; then*

$$|\mathsf{Pr}\left[f(M') = h(M)\right] - \mathsf{Pr}\left[f(M) = h(M \oplus \sigma(S))\right]| \leq \frac{1}{2} \mathop{\mathsf{Exp}}_{M}\left[|T(M)|\right] \ .$$

Proof. Define $c(m, m')$ so that $c(m, m') = 1$ when $f(m) = h(m')$ and 0 if $f(m) \neq h(m')$. As $h(\cdot)$ takes values in the set $\{\pm 1\}$, we can rewrite $c(m, m') = \frac{1}{2}[1 + f(m)h(m')]$ and

$$\begin{aligned}
&|\mathsf{Pr}\left[f(M') = h(M)\right] - \mathsf{Pr}\left[f(M) = h(M \oplus \sigma(S))\right]| \\
=&\left|\underset{M,M'}{\mathsf{Exp}}\left[c(M, M')\right] - \underset{M,S}{\mathsf{Exp}}\left[c(M, M \oplus \sigma(S))\right]\right| \\
=&\left|\underset{M}{\mathsf{Exp}}\left[\frac{f(M)}{2}\left(\underset{M'}{\mathsf{Exp}}\left[h(M')\right] - \underset{S}{\mathsf{Exp}}\left[h(M \oplus \sigma(S))\right]\right)\right]\right| \\
\leq&\frac{1}{2}\underset{M}{\mathsf{Exp}}\left[\left|\underset{M'}{\mathsf{Exp}}\left[h(M')\right] - \underset{S}{\mathsf{Exp}}\left[h(M \oplus \sigma(S))\right]\right|\right] = \frac{1}{2}\underset{M}{\mathsf{Exp}}\left[|T(M)|\right].
\end{aligned}$$

□

We apply the second moment method to control $\mathsf{Exp}_M[|T(M)|]$. Observe that $\mathsf{Exp}_M[h(M)] = \widehat{h}_\emptyset$, so

$$\begin{aligned}
T(m) =& \underset{S}{\mathsf{Exp}}\left[\sum_{\alpha \neq \emptyset} \widehat{h}_\alpha \chi_\alpha(m \oplus \sigma(S))\right] \\
&= \sum_{\alpha \neq \emptyset} \widehat{h}_\alpha \underset{S}{\mathsf{Exp}}\left[\chi_\alpha(m \oplus \sigma(S))\right] = \sum_{\alpha \neq \emptyset} \widehat{h}_\alpha \chi_\alpha(m) \underset{S}{\mathsf{Exp}}\left[\chi_\alpha(\sigma(S))\right] \quad .
\end{aligned}$$

Then $\mathsf{Exp}_M\left[T(M)\right] = \sum_{\alpha \neq \emptyset} \widehat{h}_\alpha \mathsf{Exp}_S\left[\chi_\alpha(\sigma(S))\right] \mathsf{Exp}_M\left[\chi_\alpha(M)\right] = 0$. Now, the random variables $\widehat{h}_\alpha \chi_\alpha(M \oplus \sigma(S))$ and $\widehat{h}_\beta \chi_\beta(M \oplus \sigma(S))$ are pairwise independent (recall that M is uniform) so that

$$\begin{aligned}
\underset{M}{\mathsf{Var}}\left[T(M)\right] =& \underset{M}{\mathsf{Var}}\left[\sum_{\alpha \neq \emptyset} \widehat{h}_\alpha \chi_\alpha(M) \underset{S}{\mathsf{Exp}}[\chi_\alpha(\sigma(S))]\right] \\
&= \sum_{\alpha \neq \emptyset} \widehat{h}_\alpha^2 \underset{S}{\mathsf{Exp}}[\chi_\alpha(\sigma(S))]^2 \underset{M}{\mathsf{Var}}[\chi_\alpha(M))] \leq \epsilon^2 \sum_{\alpha \neq \emptyset} \widehat{h}_\alpha^2 \leq \epsilon^2
\end{aligned}$$

by the Plancherel equality (see Section 2.3) and the fact that $\mathsf{Var}_M\left[\chi_\alpha(M)\right] = 1$. Now, applying Chebyshev's inequality, we have $\mathsf{Pr}\left[|T(M)| > \lambda\right] < \epsilon^2 \lambda^{-2}$.

Selecting $\lambda = \epsilon^{\frac{2}{3}}$, we have

$$\begin{aligned}
\underset{M}{\mathsf{Exp}}\left[|T(M)|\right] \leq& \mathsf{Pr}\left[|T(M)| > \lambda\right] \cdot \max_m |T(m)| + \mathsf{Pr}\left[|T(M)| \leq \lambda\right] \cdot \lambda \\
\leq& \frac{\epsilon^2}{\lambda^2} \cdot 2 + (1 - \frac{\epsilon^2}{\lambda^2}) \cdot \lambda \leq 3\epsilon^{\frac{2}{3}}.
\end{aligned}$$

Hence $|\mathsf{Pr}[f(M \oplus \sigma(S)) = h(M)] - \mathsf{Pr}[G_f = h(M)]| < \frac{3}{2}\epsilon^{\frac{2}{3}}$. As $\epsilon = n^{-\omega(1)}$, this completes the proof. The bound on $|s|$ and the running time of E^u and D^u follow from Section 2.1. □

4 Security for Entropically Rich Message Spaces

In this section we describe a symmetric encryption system offering ℓ-entropic security; keys of length $\ell + \omega(\log n)$ suffice. In preparation, we will slightly enrich our notion of symmetric encryption system by allowing the encryption function(s) to be stochastic: for each n, E_n may depend on m, the message, s, the secret key, and ϕ, some private random coins of the encryption function. To keep the notation uniform, we let $\varPhi_n$ be the random variable on which E_n may depend. $\varPhi_n$ is independent of M_n and S_n.

For convenience, we will assume that the message space is $\mathbb{Z}_{p+1}$ for a prime p. (So we treat $\mathcal{M} = \{M_p \mid p \text{ prime}\}$, where M_p is a random variable taking values in $\mathbb{Z}_{p+1}$.) To keep our notation uniform, we let $n = \log(p+1)$ and then say that $M_p \in \mathbb{Z}_{p+1}$ has min-entropy $n - \ell$ if $\mathsf{Pr}[M_p = m_0] \leq 2^{-n+\ell}$. Now, we select an artificial bijection $L : \mathbb{Z}_{p+1} \to P_2(\mathbb{F}_p)$, so that

$$L(z) = \left[\begin{pmatrix} 1 \\ z \end{pmatrix}\right], \text{ for } 0 \leq z \leq p-1, \text{ and} \qquad L(p) = \left[\begin{pmatrix} 0 \\ 1 \end{pmatrix}\right] \ .$$

L can be computed in linear time; L^{-1} can be computed by single inversion modulo p. Having fixed this bijection, we will treat the 3-wise independent functions $\mathrm{PGL}_2(\mathbb{F}_p)$, described in Sect. 2.2, as if they act on $\mathbb{Z}_{p+1}$.

Theorem 3. *Let $(\mathcal{E}^k, \mathcal{D}^k)$ be the encryption system with $E_p^k(m, s; \phi) = (\phi(m) + s, \phi)$, where the message $m \in \mathbb{Z}_{p+1}$, s is the secret key, chosen in the set $\{0, 1, \ldots, 2^k - 1\} \subset \mathbb{Z}_{p+1}$, and ϕ is an element of $PGL_2(\mathbb{F}_p)$ selected at random by E_n^k. (Here $+$ is modulo $p+1$.) Decryption is defined analogously. Then if $k = \ell + \omega(\log n)$, this encryption system offers ℓ-entropic security. Furthermore, E_n^k and D_n^k can be computed in time $O(n \log^2 n \log\log n)$.*

Proof. We need to show that for every message space $\mathcal{M}$ of sufficient min-entropy and every $f : \mathbf{im}\ E_p^k \to \{0,1\}$, there is a random variable G_f, independent of M_p, such that for all functions $h : \mathbb{Z}_{p+1} \to \{0,1\}$

$$|\mathsf{Pr}[f(\varPhi(M) + S, \varPhi) = h(M)] - \mathsf{Pr}[G_f = h(M)]| \leq \frac{1}{n^{\omega(1)}} \ .$$

Here S is uniform on $K = \{0, \ldots, 2^k - 1\} \subset \mathbb{Z}_{p+1}$ and $\varPhi$ is uniform on $\mathrm{PGL}_2(\mathbb{F}_p)$; $\varPhi$, S, and M are independent.

We define the random variable G_f, which can predict $h(M)$ nearly as well as can f even without observing $E(M, S; \varPhi)$. As in the proof above, let M', S', and $\varPhi'$ be a random variables with the same distributions as M, S, and $\varPhi$; all independent. Define the random variable $G_f = f(\varPhi'(M') + S', \varPhi')$. Observe that

$$\mathsf{Pr}[G_f = h(M)] = \mathsf{Pr}[f(\varPhi'(M') + S', \varPhi') = h(M)] = \mathsf{Pr}[f(\varPhi(M') + S, \varPhi) = h(M)] \ .$$

We begin by recording an analogue of Lemma 1 for this cryptosystem which allows us to remove the dependence on f.

Lemma 2. *For any* $h : \mathbb{Z}_{p+1} \to \{0,1\}$ *and* $f : \mathbb{Z}_{p+1} \times PGL_2(F_p) \to \{0,1\}$*,*

$$|\mathsf{Pr}\left[f(\Phi(M)+S,\Phi)=h(M')\right] - \mathsf{Pr}\left[f(\Phi(M)+S,\Phi)=h(M)\right]| \le \\ \mathop{\mathsf{Exp}}_{M,\Phi,S}\left[\left|\mathop{\mathsf{Exp}}_{M'}[h(M')] - \mathop{\mathsf{Exp}}_{M',S'}\left[h(M')\,\middle|\,\Phi(M')+S'=\Phi(M)+S\right]\right|\right] \ .$$

Proof. For simplicity, use two new functions $\bar{f}$ and $\bar{h}$ which take values in $\{\pm 1\}$ rather than $\{0,1\}$. Let $\tilde{f}(x) = 2f(x)-1$ and $\tilde{h}(x) = 2h(x)-1$. Then

$$\begin{aligned}
&|\mathsf{Pr}\left[f(\Phi(M)+S,\Phi)=h(M')\right] - \mathsf{Pr}\left[f(\Phi(M)+S,\Phi)=h(M)\right]| \\
=&\frac{1}{2}\left|\mathop{\mathsf{Exp}}_{M,M',\Phi,S}\left[\tilde{f}(\Phi(M)+S,\Phi)\tilde{h}(M')\right] - \mathop{\mathsf{Exp}}_{M,\Phi,S}\left[\tilde{f}(\Phi(M)+S,\Phi)\tilde{h}(M)\right]\right| \\
=&\frac{1}{2}\left|\mathop{\mathsf{Exp}}_{\Phi,M,S}\left[\tilde{f}(\Phi(M)+S,\Phi)\left(\mathop{\mathsf{Exp}}_{M'}[\tilde{h}(M')]-\tilde{h}(M)\right)\right]\right| \\
=&\frac{1}{2}\Big|\mathop{\mathsf{Exp}}_{\Phi}\Big[\sum_{e\in\mathbb{Z}_{p+1}}\mathsf{Pr}[\Phi(M)+S=e]\cdot\tilde{f}(e,\Phi)\cdot \\
&\qquad\qquad \mathop{\mathsf{Exp}}_{M,S}\Big[\big(\mathop{\mathsf{Exp}}_{M'}[\tilde{h}(M')]-\tilde{h}(M)\big) \ \mid\ \Phi(M)+S=e\Big]\Big]\Big| \\
\le&\frac{1}{2}\mathop{\mathsf{Exp}}_{\Phi}\left[\sum_{e}\mathsf{Pr}[\Phi(M)+S=e]\cdot\left|\mathop{\mathsf{Exp}}_{M}[\tilde{h}(M)]-\mathop{\mathsf{Exp}}_{M,S}\left[\tilde{h}(M)\,\middle|\,\Phi(M)+S=e\right]\right|\right]
\end{aligned}$$

Observe now that for any functions $g_1 : \mathbb{Z}_{p+1} \to \mathbb{R}$ and $g_2 : \mathbb{Z}_{p+1} \to \mathbb{Z}_{p+1}$,

$$\sum_{e}\mathop{\mathsf{Pr}}_{x}[g_2(x)=e]\mathop{\mathsf{Exp}}_{x}\left[g_1(g_2(x)) \mid g_2(x)=e\right] = \mathop{\mathsf{Exp}}_{x}[g_1(g_2(x))] \ ;$$

the statement of the lemma follows. □

Now, for an element $s_0 \in K$, let $K_{s_0} = \{s - s_0 \bmod (p+1) \mid s \in K\}$. Then define

$$G^{s_0}_{m,\phi} = \mathop{\mathsf{Exp}}_{M}[h(M)] - \mathop{\mathsf{Exp}}_{M,S'}\left[h(M)\,\middle|\,\phi(M)+(S'-s_0)=\phi(m)\right] \ ,$$

where $S'-s_0$ is uniform on K_{s_0}. From the lemma above, it suffices to show that for every $s_0 \in K$, $\mathsf{Exp}_{M,\Phi}[|G^{s_0}_{M,\Phi}|]$ is small. So fix $s_0 \in K$. We let $p_m = \mathsf{Pr}[M = m]$. For a message $m_0 \in \mathbb{Z}_{p+1}$ and an element $w_0 \in \mathbb{Z}_{p+1}$, let $\mathcal{P}_{w_0\to m_0} = \{\phi \in \mathrm{PGL}_2(\mathbb{F}_p) \mid \phi(m_0) = w_0\}$. We will handle each of these "slices" of $\mathrm{PGL}_2(\mathbb{F}_p)$ separately. For any permutation ϕ in $\mathcal{P}_{w_0\to m_0}$ we can consider the sums

$$A^{\phi}_0 = \sum_{m\in\phi^{-1}(w_0+K_{s_0})} p_m \qquad \text{and} \qquad B^{\phi}_0 = \sum_{m\in\phi^{-1}(w_0+K_{s_0})} p_m\cdot h(m) \ ;$$

then

$$\frac{B^{\phi}_0}{A^{\phi}_0} = \mathop{\mathsf{Exp}}_{M}\left[h(M) \mid \phi(M) \in w_0 + K_{s_0}\right]. \tag{1}$$

(Here $w_0 + K_{s_0}$ denotes the set $\{w_0 + s \bmod (p+1) \mid s \in K_{s_0}\}$.) We would like to see that this quotient is well-behaved (i.e., near the expected value of $h(M)$,) for most ϕ.

Let Φ_0 be a uniform random variable in $\mathcal{P}_{m_0 \to w_0}$ and let X_m be the random variable taking the value p_m if $\Phi_0(m) \in w_0 + K_{s_0}$, and 0 otherwise. Then $\sum_{m \in \mathbb{Z}_{p+1}} X_m = A_0^{\Phi_0}$ and $\sum_{m \in h^{-1}(1)} X_m = B_0^{\Phi_0}$ so that

$$\begin{aligned}\mathsf{Exp}[A_0^{\Phi_0}] &= \mathsf{Exp}\Big[p_{m_0} + \sum_{m \neq m_0} X_m\Big] = p_{m_0} + \sum_{m \neq m_0} (\mathsf{Pr}[\Phi_0(m) \in w_0 + K_{s_0}] \cdot p_m) \\ &= p_{m_0}\left(1 - \frac{2^k - 1}{p+1}\right) + \frac{2^k - 1}{p+1}.\end{aligned}$$

Hence $\frac{2^k-1}{p+1} \leq \mathsf{Exp}[A_0^{\Phi_0}] \leq p_{m_0}\left(1 - \frac{2^k-1}{p+1}\right) + \frac{2^k-1}{p+1} \leq 2^{-n+k} + p_{m_0}$. Let $\overline{h} = \mathsf{Exp}_M[h(M)]$; then, similarly,

$$\overline{h} \cdot \frac{2^k - 1}{p+1} \leq \mathop{\mathsf{Exp}}_{\Phi_0}[B_0^{\Phi_0}] \leq \overline{h} \cdot \frac{2^k - 1}{p+1} + p_{m_0}\left(1 - \frac{2^k - 1}{p+1}\right) \ .$$

Recalling that the distribution of M has min-entropy $n - \ell$,

$$\frac{\mathsf{Exp}_{\Phi_0}[B_0^{\Phi_0}]}{\mathsf{Exp}_{\Phi_0}[A_0^{\Phi_0}]} - \overline{h} \leq \frac{2^\ell}{2^k - 1} \ ,$$

and similarly,

$$\frac{\mathsf{Exp}_{\Phi_0}[B_0^{\Phi_0}]}{\mathsf{Exp}_{\Phi_0}[A_0^{\Phi_0}]} - \overline{h} \geq \frac{\overline{h} \cdot (2^k - 1)}{2^k + (p+1) \cdot p_{m_0} - 1} - \overline{h} \geq -2^{-k+\ell} \ .$$

Hence,

$$\left|\frac{\mathsf{Exp}_{\Phi_0}[B_0^{\Phi_0}]}{\mathsf{Exp}_{\Phi_0}[A_0^{\Phi_0}]} - \overline{h}\right| \leq 2 \cdot 2^{-k+\ell} \ .$$

We wish to insure that $A_0^{\Phi_0}$ and $B_0^{\Phi_0}$ are close to their expected vales. In preparation for applying Chebyshev's inequality, we compute their variances. We have

$$\mathop{\mathsf{Var}}_{\Phi_0}[A_0^{\Phi_0}] = \mathsf{Var}[\sum_{m \in \mathbb{Z}_p} X_m] = \sum_{m \neq m_0} \mathop{\mathsf{Var}}_{\Phi_0}[X_m] + \sum_{m_1 \neq m_2 \neq m_0} \mathop{\mathsf{Cov}}_{\Phi_0}[X_{m_1}, X_{m_2}] \ .$$

Now, $\sum_{m \neq m_0} \mathsf{Var}[X_m] \leq \sum_{m \neq m_0} \mathsf{Exp}[X_m^2] \leq 2^{-2n+k+\ell}$, and these variables are pairwise negatively correlated (so that $\mathsf{Cov}[X_{m_1}, X_{m_2}] < 0$ for $m_1 \neq m_2$, both distinct from m_0). Then, $\mathsf{Var}_{\Phi_0}[A_0^{\Phi_0}] < 2^{-2n+k+\ell}$. Similarly, $\mathsf{Var}_{\Phi_0}[B_0^{\Phi_0}] < \overline{h} \cdot 2^{-2n+k+\ell}$. Observe that 3-wise independence is required here.

By Chebyshev's inequality we have

$$\mathop{\mathsf{Pr}}_{\Phi_0}\left[\left|A_0^{\Phi_0} - \mathsf{Exp}[A_0^{\Phi_0}]\right| \geq \delta_a\right] \leq \frac{\mathsf{Var}[A_0^{\Phi_0}]}{\delta_a^2} < \frac{2^k}{2^{2n-\ell} \cdot \delta_a^2}, \tag{2}$$

and

$$\Pr_{\Phi_0}\left[\left|B_0^{\Phi_0} - \mathsf{Exp}[B_0^{\Phi_0}]\right| \geq \delta_b\right] \leq \frac{\mathsf{Var}[B_0^{\Phi_0}]}{\delta_b^2} < \frac{\overline{h} \cdot 2^k}{2^{2n-\ell} \cdot \delta_b^2} \ . \tag{3}$$

If ϕ is an element of $\mathcal{P}_0$ for which both

$$\left|A_0^{\phi} - \mathop{\mathsf{Exp}}_{\Phi_0}[A_0^{\phi}]\right| \leq \delta_a \qquad \text{and} \qquad \left|B_0^{\phi} - \mathop{\mathsf{Exp}}_{\phi}[B_0^{\Phi_0}]\right| \leq \delta_b \ ,$$

we have in particular, from equation (1),

$$\left|\frac{\mathsf{Exp}_{\Phi_0}[B_0^{\Phi_0}]}{\mathsf{Exp}_{\Phi_0}[A_0^{\Phi_0}]} - \mathop{\mathsf{Exp}}_{M}[h(M) \mid \phi(M) \in w_0 + K_{s_0}]\right| \leq \frac{2(\delta_a + \delta_b)}{\mathsf{Exp}_{\Phi_0}[A_0^{\Phi_0}]} \leq 2(\delta_a+\delta_b)\cdot 2^{n-k} \ ,$$

(assuming that $\delta_a, \delta_b < 1/2$). When $\delta_a = \delta_b = \frac{\epsilon_1}{4\cdot 2^{n-k}}$, this is the statement that

$$\left|\mathop{\mathsf{Exp}}_{M}[h(M) \mid \phi(m) \in w_0 + K_{s_0}] - \mathop{\mathsf{Exp}}_{m}[h(m)]\right| < \epsilon_1 + 2 \cdot 2^{-k+\ell} \ .$$

For this fixed s_0, we say that $(\phi, w) \in \mathrm{PGL}_2(p) \times \mathbb{Z}_{p+1}$ is ϵ-*concealing* if

$$\left|\mathop{\mathsf{Exp}}_{M}[h(M) \mid \phi(M) \in w + K_{s_0}] - \overline{h}\right| < \epsilon \ .$$

From inequalities (2) and (3), for any fixed m_0, w_0 and $\epsilon_1 > 0$ we see that

$$\Pr\left[(\Phi_0, w_0) \text{ is not } (\epsilon_1 + 2 \cdot 2^{-k+\ell})\text{-concealing}\right] \leq \frac{2 \cdot 2^k}{2^{2n-\ell} \cdot \delta_a^2} < \frac{2^5}{\epsilon_1^2 \cdot 2^{k-\ell}} \ .$$

Now, for any fixed m_0 and $\epsilon > 2 \cdot 2^{-k+\ell}$,

$$\begin{aligned}
&\Pr\left[(\Phi, \Phi(m_0)) \text{ is } \epsilon\text{-concealing}\right] \\
&= \sum_w \Pr[\Phi \in \mathcal{P}_{m_0 \to w}] \cdot \Pr[(\Phi, w) \text{ is } \epsilon\text{-concealing} \mid \Phi \in \mathcal{P}_{m_0 \to w}] \\
&\geq 1 - \frac{2^5}{\epsilon^2 \cdot 2^{k-\ell} - 4\epsilon + 4 \cdot 2^{-k+\ell}}.
\end{aligned}$$

For a random pair (m, ϕ), we will (lower) bound the probability that $(\phi, \phi(m))$ is ϵ-concealing pair for s_0, since $\mathsf{Exp}_{M,\Phi}[|G_{M,\Phi}^{s_0}|]$ is no greater than

$$\epsilon \Pr[(\Phi, \Phi(M)) \text{ is } \epsilon\text{-concealing}] + (1 - \Pr[(\Phi, \Phi(M)) \text{ is } \epsilon\text{-concealing}]) \ .$$

For specific m, we define Y_m to be the random variable taking the value 1 if $(\Phi, \Phi(m))$ is ϵ-concealing, and 0 otherwise. Now,

$$\begin{aligned}
&\Pr[(\Phi, \Phi(M)) \text{ is } \epsilon\text{-concealing}] = \mathop{\mathsf{Exp}}_{M,\Phi}[Y_M] = \sum_m p_m \cdot \mathop{\mathsf{Exp}}_{\Phi}[Y_m] \\
&\quad = \sum_m p_m \Pr[(\Phi, \Phi(m)) \text{ is } \epsilon\text{-concealing}] \geq 1 - \frac{2^5}{\epsilon^2 \cdot 2^{k-\ell} - 4\epsilon + 4 \cdot 2^{-k+\ell}},
\end{aligned}$$

so that for all $\epsilon > 2 \cdot 2^{-k+\ell}$, $\mathsf{Exp}_{M,\Phi}\left[\left|G^{s_0}_{M,\Phi}\right|\right] \leq \epsilon(1-\delta)+\delta < \epsilon+\delta$, where $\delta^{-1} = 2^{-5}\cdot(\epsilon^2\cdot 2^{k-\ell} - 4\epsilon + 4\cdot 2^{-k+\ell})$. Select $\epsilon = 4\cdot 2^{\frac{-k+\ell}{3}}$. As $k = \ell + \omega(\log n)$, we can be guaranteed that $\epsilon = n^{-\omega(1)}$ and so for all s, $\mathsf{Exp}_{M,\Phi}\left[\left|G^{s}_{M,\Phi}\right|\right] = n^{-\omega(1)}$. Hence, $\mathsf{Exp}_{M,\Phi,S}\left[\left|G^{S}_{m,\phi}\right|\right] = n^{-\omega(1)}$, which, considering the above lemma, completes the proof. Note that if, in general, the keys have length $\ell + s$, the advantage is bounded by $2^{-\Theta(s)}$. □

5 Applications to Asymmetric Encryption

In many practical situations requiring public-key cryptography, encryption and decryption (with, e.g., RSA) are so expensive that they are used only to exchange a session key, which is then fed into some cheaper symmetric system. A similar approach is possible with the systems above.

As mentioned in the introduction, a natural analogue of the notion of perfect secrecy for resource-bounded adversaries is the notion of *semantic security.* This is the guarantee that no bounded adversary can predict any piece of partial information about the message with non-negligible advantage. There are a range of complexity-theoretic assumptions which can give rise to such systems. In general, stronger assumptions allow implementations with improved efficiency. Firstly, constructions are possible under "generic" assumptions, e.g., existence of a one-way trapdoor permutation [7]. Under the stronger assumption that factoring is hard, a system of Blum and Goldwasser [6] based on the Rabin functions ($x \mapsto x^2 \bmod pq$) encrypts (in a semantically secure fashion) an n-bit message in time $O(nk \operatorname{poly}(\log k))$. Here k is the security parameter of the system. (It is interesting to note that under assumptions of a presumably stronger flavor, Cramer and Shoup [8] show that a constant number of exponentiations over a group suffice to encrypt a group element, in such a way that the resulting system is secure against even (adaptive) chosen ciphertext attack. In particular, hardness of the Diffie-Hellman decision problem is sufficient.)

As the encryption schemes of Theorem 3 are quite efficient, it is interesting to consider the (public-key) system obtained by applying an extant public key system to securely transmit the (short) shared key of E^b. Specifically, if E_{pub} is the encryption algorithm for a public key system offering semantic security, one can study the behavior of the scheme which encrypts a message m as $(\phi(m) + z, \phi, E_{\text{pub}}(z; P, R))$; here P denotes the public key, R the random string required for E, and z the and ϕ the variables of Theorem 3. If, for example, the public key system is taken to be that of Blum and Goldwasser mentioned above, the hybrid system has running time $O([\ell + w(k)\log k]k \operatorname{poly}(\log k) + n \operatorname{poly}(\log n))$, where w is any function which tends to infinity. (Here k is the security parameter of the system.) *Note, however, that the resulting security guarantee will be weaker than that of semantic security: it requires message spaces of min-entropy $n - \ell$.*

The proof of security for this system follows the proof of Theorem 3 except that one needs to initially argue that availability of the semantically secure encryption of the secret key does not interfere with the security guarantee. This fact relies on a variant of an "elision" lemma originally proved in [11], for which we give a new, streamlined proof.

For definitions of public-key cryptosystem, semantic security, and indistinguishability of encryptions we refer the reader to [11].

We shift notation in this section to agree with [12]: when x is a variable and S a random variable, $x \leftarrow S$ denotes the assignment of x according to S. If S is simply a set, we abuse the notation by allowing S to represent the random variable uniform on S. In the sequel, we will use the term "algorithm" to refer to a probabilistic polynomial time Turing machine. Furthermore, a "message generator" is an algorithm which, given 1^k, produces a output in the set $\{0,1\}^n$ (determined by the random coins of M), where n is polynomially bounded in k. Whenever a probability is expressed it is understood that the random coins of any algorithm appearing inside the brackets are to be included in the probability space. When the underlying probability space of a variable x is clear from context, we may simply write $\mathsf{Pr}_x[P(x)]$, or elide x altogether. A public-key encryption scheme is described by a triple (G, E, D): here G is a key generator algorithm which, given 1^k, generates a pair (P, S); $E(m, P, R)$ is the encryption algorithm, operating on a message m, the public key P, and a random string R; and $D(c, S)$ is the decryption algorithm, operating on a ciphertext c and the private key S.

The following lemma, which generalizes the original elision lemma of [11], is due to [10]. We give a streamlined proof which improves upon previous proofs in the sense that it *requires no sampling* on the part of the constructed algorithm (F, in the proof below). It gives an error bound which depends only on a natural 2-norm of the message distribution. Roughly, the lemma asserts that a cryptosystem offering indistinguishability of encryptions possesses the property that any efficient computation performed with observation of $E(m)$, an encryption, (and, perhaps, some related information) may as well have been performed without it. In the system mentioned above (coupling E^b with E_{pub}), this allows us to disregard the fact that the bounded adversary has witnessed a semantically secure encryption of the shared secret key s; the result is a security guarantee of the form appearing in Theorem 3 but for efficient adversaries.

Lemma 3. *Let (G, E, D) denote an encryption system possessing indistinguishability of encryptions. Then for every message space M and algorithm A, there is an algorithm B so that for all polynomials Q_1, all efficiently computable $f : \{0,1\}^* \to \{0,1\}^*$, and every polynomial Q_2, $\exists k_0, \forall k > k_0$ and $\forall h : \{0,1\}^* \to \{0,1\}^*$,*

$$\mathsf{Pr}[A(1^k, P, f(s,m), E(m; P, R)) = h(s,m)] \leq \mathsf{Pr}[B(1^k, f(s,m)) = h(s,m)] + \frac{1}{Q_2(k)}.$$

The first probability is taken over $m \leftarrow M(1^k)$, $(P, S) \leftarrow G(1^k)$, $s \leftarrow \{0,1\}^{Q_1(k)}$, and R. The second probability is taken over $m \leftarrow M(1^k)$ and $s \leftarrow \{0,1\}^{P(k)}$.

Proof. The algorithm B uses A as a black box: given 1^k and $f(s,m)$, B proceeds as follows: generate $m' \leftarrow M(1^k)$, $(P,S) \leftarrow G(1^k)$, and a random string R of appropriate length; return $v = A(1^k, P, f(s,m), E(m';P,R))$.

Observe that $\mathsf{Pr}[B(1^k, f(s,m)) = h(s,m)]$ is exactly the probability $\mathsf{Pr}[A(1^k, P, f(s,m), E(m';P,R)) = h(s,m)]$. In this case, the lemma is a consequence of the following claim:

Claim. For every message space M, efficient algorithm A, every polynomial Q_1, efficiently computable $f : \{0,1\}^* \to \{0,1\}^*$, and every polynomial Q_2, $\exists k_0, \forall k > k_0$ and $\forall h : \{0,1\}^* \to \{0,1\}^*$,

$$\mathsf{Pr}[A(1^k, P, f(s,m), E(m;P,R)) = h(s,m)] \leq \\ \mathsf{Pr}[A(1^k, P, f(s,m), E(m';P,R)) = h(s,m)] + \frac{1}{Q_2(k)},$$

where each probability is taken over $m \leftarrow M(1^k)$, $m' \leftarrow M(1^k)$, $(P,S) \leftarrow G(1^k)$, $s \leftarrow \{0,1\}^{Q_1(k)}$, and R.

Proof (of Claim). Suppose not. Then there is a polynomial Q_2, a message space M, and an algorithm A, a polynomial Q_1 and a function f so that $\forall k_0, \exists k > k_0$,

$$\mathop{\mathsf{Pr}}_{s,m,R,P}\left[A(1^k, P, f(s,m), E(m;P,R)) = h(s,m)\right] > \\ \mathop{\mathsf{Pr}}_{s,m,m',R,P}\left[A(1^k, P, f(s,m), E(m';P,R)) = h(s,m)\right] + \epsilon$$

where $\epsilon = \epsilon(k) = \frac{1}{Q_2(k)}$. For a pair of messages m, m', define $P_{m,m'}$ to be the probability $\mathsf{Pr}_{s,R,P}[A(1^k, P, f(s,m), E(m';P,R)) = h(s,m)]$ and $P_{m,*} = \mathsf{Exp}_{m'}[P_{m,m'}]$. Observe, then, that

$$\mathop{\mathsf{Pr}}_{s,m,R,P}\left[A(1^k, P, f(s,m), E(m;P,R)) = h(s,m)\right] = \mathop{\mathsf{Exp}}_{m}[P_{m,m}], \text{ and} \\ \mathop{\mathsf{Pr}}_{s,m,m',R,P}\left[A(1^k, P, f(s,m), E(m';P,R)) = h(s,m)\right] = \mathop{\mathsf{Exp}}_{m}[P_{m,*}].$$

In particular, $\mathsf{Exp}_m[P_{m,m}] - \mathsf{Exp}_m[P_{m,*}] > \epsilon$.

Now, we build an algorithm F which, given random m_0 and m_1, can distinguish an encryption of m_0 from one of m_1. The algorithm F proceeds as follows: given m_0, m_1 and $\alpha = E(m_i; P, R)$, *(i.)* j is chosen uniformly in $\{0,1\}$, s is chosen uniformly in $\{0,1\}^{P(n)}$, and R is chosen uniformly among strings of appropriate length, *(ii.)* $E(m_j; P, R)$ and $f(s, m_0)$ are computed, *(iii.)* $A(1^k, f(s,m_0), E(m_j;P,R))$ is simulated, resulting in the value v_j, and $A(1^k, f(s,m_0), \alpha)$ is simulated, resulting in the value v, and, finally, *(iv.)* if $v = v_j$, the j is output; otherwise $1-j$ is output.

Let $I_n = \{A(1^k, P, f(s,m'), E(m;P,R)) \mid m, m' \in \{0,1\}^n, s \in \{0,1\}^{Q_1(n)}, R\}$ be values that algorithm A can take, when restricted to those inputs possible when $|m| = n$. Then, for $v \in I_n$, let

$$D^s_{m',m}(v) = \mathop{\mathsf{Pr}}_{R,P}[A(1^k, P, f(s,m'), E(m;P,R)) = v] \ ,$$

so that $P_{m',m} = \mathsf{Exp}_s[D^s_{m',m}(h(s,m'))]$. Now, for a particular pair m_0, m_1, the probability $\mathsf{Pr}[F(m_0,m_1,\alpha) = i]$ is

$$\sum_{i'=0}^{1}\sum_{j'=0}^{1} \mathsf{Pr}[i = i' \wedge j = j'] \cdot \mathsf{Pr}[F(m_0,m_1,\alpha) = i' \mid i = i', j = j']$$

$$= \underset{s}{\mathsf{Exp}}\left[\frac{1}{4}\sum_v D^s_{m_0,m_0}(v)^2 + 2(1 - \sum_v D^s_{m_0,m_0}(v) \cdot D^s_{m_0,m_1}(v)) + \sum_v D^s_{m_0,m_1}(v)^2\right]$$

$$\overset{*}{\geq} \frac{1}{2} + \frac{1}{4}(\underset{s}{\mathsf{Exp}}\left[D^s_{m_0,m_0}(h(s,m_0)) - D^s_{m_0,m_1}(h(s,m_0))\right])^2.$$

which is $\frac{1}{2} + \frac{1}{4}(P_{m_0,m_0} - P_{m_0,m_1})^2$. Here inequality $\overset{*}{\geq}$ follows because $\mathsf{Exp}[X]^2$ never exceeds $\mathsf{Exp}[X^2]$ for any random variable. Then

$$\underset{m_0,m_1}{\mathsf{Pr}}[F(m_0,m_1,\alpha) = i] \geq \underset{m_0,m_1}{\mathsf{Exp}}\left[\frac{1}{2} + \frac{1}{4} \cdot (P_{m_0,m_0} - P_{m_0,m_1})^2\right]$$

$$\geq \frac{1}{2} + \frac{1}{4} \cdot (\underset{m_0,m_1}{\mathsf{Exp}}[P_{m_0,m_0} - P_{m_0,m_1}])^2 = \frac{1}{2} + \frac{1}{4} \cdot (\underset{m_0}{\mathsf{Exp}}[P_{m_0,m_0} - \underset{m_1}{\mathsf{Exp}}[P_{m_0,m_1}]])^2$$

$$== \frac{1}{2} + \frac{1}{4} \cdot (\underset{m}{\mathsf{Exp}}[P_{m,m}] - \underset{m}{\mathsf{Exp}}[P_{m,*}])^2 \geq \frac{1}{2} + \frac{\epsilon^2}{4}.$$

Hence (E, D) does not offer indistinguishability of encryptions. □

As mentioned above, the Lemma follows immediately from the Claim. □

References

[1] Noga Alon, Oded Goldreich, Johan Håstad, and René Peralta. Simple constructions of almost k-wise independent random variables. In *31st Annual Symposium on Foundations of Computer Science*, volume II, pages 544–553, St. Louis, Missouri, 22–24 October 1990. IEEE.

[2] Yonatan Aumann and Michael O. Rabin. Information theoretically secure communication in the limited storage space model. In Michael Wiener, editor, *Advances in Cryptology—CRYPTO '99*, volume 1666 of *Lecture Notes in Computer Science*, pages 65–79. Springer-Verlag, 1999.

[3] Eric Bach and Jeffrey Shallit. *Algorithmic number theory. Vol. 1.* MIT Press, Cambridge, MA, 1996. Efficient algorithms.

[4] M. Bellare, A. Desai, A. Pointcheval, and P. Rogaway. Relations among notions of public-key cryptosystems. In Krawczyk [13], page 540.

[5] Mihir Bellare and Phillip Rogaway. Optimal asymmetric encryption. In Alfredo De Santis, editor, *Advances in Cryptology—EUROCRYPT 94*, volume 950 of *Lecture Notes in Computer Science*, pages 92–111. Springer-Verlag, 1995, 9–12 May 1994.

[6] Manuel Blum and Shafi Goldwasser. An *efficient* probabilistic public-key encryption scheme which hides all partial information. In G. R. Blakley and David Chaum, editors, *Advances in Cryptology: Proceedings of CRYPTO 84*, volume 196 of *Lecture Notes in Computer Science*, pages 289–299. Springer-Verlag, 1985, 19–22 August 1984.

[7] Manuel Blum and Silvio Micali. How to generate cryptographically strong sequences of pseudo-random bits. *SIAM Journal on Computing*, 13(4):850–864, November 1984.

[8] Ronald Cramer and Victor Shoup. A practical public key cryptosystem provably secure against adaptive chosen ciphertext attack. In Krawczyk [13], pages 13–25.

[9] Danny Dolev, Cynthia Dwork, and Moni Naor. Non-malleable cryptography (extended abstract). In *Proceedings of the Twenty Third Annual ACM Symposium on Theory of Computing*, pages 542–552, New Orleans, Louisiana, 6–8 May 1991.

[10] Oded Goldreich. A uniform-complexity treatment of encryption and zero-knowledge. *Journal of Cryptology*, 6(1):21–53, 1993.

[11] Shafi Goldwasser and Silvio Micali. Probabilistic encryption. *Journal of Computer and System Sciences*, 28(2):270–299, April 1984.

[12] Shafi Goldwasser, Silvio Micali, and Ronald L. Rivest. A digital signature scheme secure against adaptive chosen-message attacks. *SIAM Journal on Computing*, 17(2):281–308, April 1988.

[13] Hugo Krawczyk, editor. *Advances in Cryptology—CRYPTO '98*, volume 1462 of *Lecture Notes in Computer Science*. Springer-Verlag, 23–27 August 1998.

[14] Rudolf Lidl and Harald Niederreiter. *Finite Fields*, volume 20 of *Encyclopedia of Mathematics and its Applications*. Addison-Wesley Publishing Company, Reading, Massachusetts, 1983.

[15] Ueli M. Maurer. Conditionally-perfect secrecy and a provably-secure randomized cipher. *Journal of Cryptology*, 5(1):53–66, 1992.

[16] Joseph Naor and Moni Naor. Small-bias probability spaces: Efficient constructions and applications. *SIAM Journal on Computing*, 22(4):838–856, August 1993.

[17] Moni Naor and Moti Yung. Public-key cryptosystems provably secure against chosen ciphertext attacks. In *Proceedings of the Twenty Second Annual ACM Symposium on Theory of Computing*, pages 427–437, Baltimore, Maryland, 14–16 May 1990.

[18] Rene Peralta. On the distribution of quadratic residues and nonresidues modulo a prime number. *Mathematics of Computation*, 58(197):433–440, 1992.

[19] Charles Rackoff and Daniel R. Simon. Non-interactive zero-knowledge proof of knowledge and chosen ciphertext attack. In J. Feigenbaum, editor, *Advances in Cryptology—CRYPTO '91*, volume 576 of *Lecture Notes in Computer Science*, pages 433–444. Springer-Verlag, 1992, 11–15 August 1991.

[20] E. G. Rees. *Notes on Geometry*. Springer-Verlag, 1983.

[21] A. Schönhage. Schnelle berechnung von kettenbruchentwicklungen. *Acta Informatica*, 1:139–144, 1971.

[22] A. Schönhage. Schnelle Multiplikation von Polynomen über Körpern der Charakteristik 2. *Acta Informat.*, 7(4):395–398, 1976/77.

[23] A. Schönhage and V. Strassen. Schnelle multiplikation groβer zahlen. *Computing*, 7:281–292, 1971.

[24] Mark N. Wegman and J. Lawrence Carter. New classes and applications of hash functions. In *20th Annual Symposium on Foundations of Computer Science*, pages 175–182, San Juan, Puerto Rico, 29–31 October 1979. IEEE.

[25] Dominic Welsh. *Codes and cryptography*. The Clarendon Press Oxford University Press, New York, 1988.