Small Pseudo-Random Families of Matrices: Derandomizing Approximate Quantum Encryption

Andris Ambainis[1] and Adam Smith[2] ⋆

[1] Institute for Advanced Study, Princeton, NJ, USA. ambainis@ias.edu.
[2] MIT Computer Science and AI Lab, Cambridge, MA, USA. asmith@csail.mit.edu.

Abstract. A *quantum encryption scheme* (also called *private quantum channel*, or *state randomization protocol*) is a one-time pad for quantum messages. If two parties share a classical random string, one of them can transmit a quantum state to the other so that an eavesdropper gets little or no information about the state being transmitted. *Perfect* encryption schemes leak no information at all about the message. *Approximate* encryption schemes leak a non-zero (though small) amount of information but require a shorter shared random key. Approximate schemes with short keys have been shown to have a number of applications in quantum cryptography and information theory [8].

This paper provides the first deterministic, polynomial-time constructions of quantum approximate encryption schemes with short keys. Previous constructions [8] are probabilistic—that is, they show that if the operators used for encryption are chosen at random, then with high probability the resulting protocol will be a secure encryption scheme. Moreover, the resulting protocol descriptions are exponentially long. Our protocols use keys of the same length as the probabilistic constructions; to encrypt n qubits approximately, one needs $n + o(n)$ bits of shared key [8], whereas $2n$ bits of key are necessary for perfect encryption [3].

An additional contribution of this paper is a connection between classical combinatorial derandomization and constructions of pseudo-random matrix families in a continuous space.

1 Introduction

A *quantum encryption scheme* (or *private quantum channel*, or *state randomization protocol*) allows Alice, holding a *classical* key[3], to scramble a quantum state and send it to Bob (via a quantum channel) so that (1) Bob, given the key, can recover Alice's state exactly and (2) an adversary Eve who intercepts the ciphertext learns nothing about the message, as long as she doesn't know the key. We do not assume any shared quantum states between Alice and Bob, nor any back channels from Bob to Alice.[4]

⋆ A.A. supported by NSF grant DMS-0111298. A.S. supported by a Microsoft Ph.D. Fellowship.

[3] Classical keys are inherently easier to store, distribute and manipulate, since they can be copied. More subtly, encryption with a shared quantum key is in many ways a dual problem to encryption with a classical key; see [8, 5] for more discussion.

[4] A back channel from Bob to Alice would allow using quantum key distribution to generate a long secret key. However, such interaction is often impossible, e.g. if Alice wants to encrypt stored data for her own later use.

There are two variants of this definition. An encryption scheme is called *perfect* if Eve learns zero information from the ciphertext, and *approximate* if Eve can learn some non-zero amount of information. A perfect encryption ensures that the distributions (density matrices) of ciphertexts corresponding to different messages are exactly identical, while an approximate scheme only requires that they be very close; we give formal definitions further below. In the classical case, both perfect and approximate encryption require keys of roughly the same length—n bits of key for n bits of message. In the quantum case, the situation is different.

For perfect encryption, Ambainis et al. [3] showed that $2n$ bits of key are necessary and sufficient to encrypt n qubits. The construction consists of applying two classical one-time pads—one in the "standard" basis $\{|0\rangle, |1\rangle\}$ and another in the "diagonal" basis $\{\frac{1}{\sqrt{2}}(|0\rangle + |1\rangle), \frac{1}{\sqrt{2}}(|0\rangle - |1\rangle)\}$.

Approximate encryption was studied by Hayden, Leung, Shor and Winters [8]. They introduced an additional, useful relaxation: they showed that if the plaintext is not entangled with Eve's system to begin with, then one can get *approximate* quantum encryption using only $n + o(n)$ bits of key—roughly half as many as are necessary for perfect encryption.[5] The assumption that Eve's system is unentangled with the message is necessary for this result; otherwise roughly $2n$ bits are needed, even for approximate encryption. The assumption holds in the quantum counterpart of the one-time pad situation (one party prepares a quantum message and sends it to the second party, using the encryption scheme) as long as the message is not part of a larger cryptographic protocol.

Hayden et al. [8] showed that a *random* set of $2^{n+o(n)}$ unitary matrices leads to a good encryption scheme with high probability (to encrypt, Alice uses the key to choose one of the matrices from the set and applies the corresponding operator to her input). However, verifying that a particular set of matrices yields a good encryption scheme is not efficient; even writing down the list of matrices is prohibitive, since there are exponentially many of them.

This paper presents the first polynomial time constructions of approximate quantum encryption schemes (to relish the oxymoron: derandomized randomization protocols). The constructions run in time $O(n^2)$ when the message ρ consists of n qubits. That is, given the key and the input message, Alice can produce the output using $O(n^2)$ steps on a quantum computer. The key length we achieve is slightly better than that of the probabilistic construction of [8]. Our results apply to the trace norm on matrices; exact results are stated further below.

The main tools in our construction are small-bias sets [10] of strings in $\{0,1\}^{2n}$. Such sets have proved useful in derandomizing algorithms [10], constructing short PCPs [6] and the encryption of high-entropy messages [12]. Thus, one of the contributions of this paper is a connection between classical combinatorial derandomization and constructions of pseudo-random matrix families in a continuous space. Specifically, we connecti Fourier analysis over $\mathbb{C}^{\mathbb{Z}_2^{2n}}$ to Fourier analysis over the matrices $\mathbb{C}^{2^n \times 2^n}$. This parallels to some extent the connection between quantum error-correcting codes over n qubits and classical codes over $GF(4)^n$.

[5] The result of [8] highlights an error in the proof of a lower bound on key length of authentication schemes in [4]. The results of that paper remain essentially correct, but the definition of authentication requires some strengthening, and the proof of the lower bound is more involved.

Definitions We assume that the reader is familiar with the basic notation of quantum computing (see [11] for an introduction). Syntactically, an approximate quantum encryption scheme is a set of 2^k invertible operators $\left\{E_\kappa | \kappa \in \{0,1\}^k\right\}$. The E_κ's may be unitary, but need not be: it is sufficient that one be able to recover the input ρ from the output $E_\kappa(\rho)$, which may live in a larger-dimensional space than ρ. Each E_κ takes n qubits as input and produces $n' \geq n$ qubits of output. If $n' = n$ then each operator E_κ corresponds to a unitary matrix U_κ, that is $E_\kappa(\rho) = U_\kappa \rho U_\kappa^\dagger$.

For an input density matrix[6] ρ, the density matrix of the ciphertext from the adversary's point of view is:

$$\mathcal{E}(\rho) = \mathbb{E}_\kappa\left[E_\kappa(\rho)\right] = \frac{1}{2^k} \sum_{\kappa \in \{0,1\}^k} E_\kappa(\rho)$$

When the scheme is length-preserving, this yields $\mathcal{E}(\rho) = \frac{1}{2^k} \sum_\kappa U_\kappa \rho U_\kappa^\dagger$.

Definition 1. *The set of operators $\{E_\kappa\}$ is an approximate quantum encryption scheme with leakage ϵ (also called "ϵ-randomizing scheme") for n qubits if*

$$\textit{for all density matrices } \rho \textit{ on } n \textit{ qubits:} \qquad D(\mathcal{E}(\rho), \tfrac{1}{2^{n'}}\mathbb{I}) = \left\| \mathcal{E}(\rho) - \tfrac{1}{2^{n'}}\mathbb{I} \right\|_{tr} \leq \epsilon. \tag{1}$$

Here $\mathbb{I}$ refers to the identity matrix in dimension $2^{n'}$, and $D(\cdot,\cdot)$ refers to the trace distance between density matrices. The trace norm of a matrix σ is the trace of the absolute value of σ (equivalently, the sum of the absolute values of the eigenvalues). The *trace distance* between two matrices ρ, σ is the trace norm of their difference:

$$D(\rho, \sigma) \stackrel{\triangle}{=} \|\rho - \sigma\|_{tr} = \mathrm{Tr}(|\rho - \sigma|)$$

This distance plays the same role for quantum states that statistical difference plays for probability distributions: the maximum probability of distinguishing between two quantum states ρ, σ via a single measurement is $\frac{1}{2} + \frac{1}{4}D(\rho, \sigma)$. One can also measure leakage with respect to other norms; see below.

Remark 1. This definition of quantum encryption implicitly assumes that the message state ρ is not entangled with the adversary's system. Without that assumption the definition above is not sufficient, and it is *not* possible to get secure quantum encryption using $n(1 + o(1))$ bits of key (roughly $2n$ bits are provably necessary[7]). Thus, this sort of construction is not universally applicable, and must be used with care.

Previous Work Ambainis et al. [3] considered perfect encryption; this corresponds to the case where $\epsilon = 0$. The choice of matrix norm is irrelevant there, since $\mathcal{E}(\rho) = \frac{1}{2^{n'}}\mathbb{I}$. As mentioned above, they showed that $2n$ bits of key are necessary and sufficient. The

[6] Recall that for a pure state $|\phi\rangle$, the density matrix ρ is $|\phi\rangle\langle\phi|$.

[7] This folklore result appears more or less explicitly in both [4, 8]. Similar arguments show that n bits of key are necessary to encrypt n classical bits, even with access to quantum computers (but not interaction).

construction uses the key to choose one of 2^{2n} Pauli operators (defined below) and applies that to the input state.

Hayden et al. [8] showed that a set of $O(n2^n/\epsilon^2)$ unitary operators suffices. They showed this both for the trace norm, and for the "operator norm," discussed below. For the trace norm, they also showed that a random set of Pauli matrices of the same size would suffice. This means that for encrypting n qubits, they gave a non-polynomial-time, randomized scheme requiring $n + \log n + 2\log(1/\epsilon) + O(1)$ bits of key.

Our Results We present three explicit, polynomial time constructions of approximate state randomization protocols for the trace norm. All are based on existing constructions of δ-biased sets [10, 2, 1], or on families of sets with small average bias. The three constructions are explained and proven secure in Sections 3.1, 3.2 and 3.3, respectively.

The first construction is length-preserving, and requires $n + 2\log n + 2\log(1/\epsilon) + O(1)$ bits of key, roughly matching the performance of the non-explicit construction. The second construction is length-increasing: it encodes n qubits into n qubits and $2n$ classical bits but uses a shorter key: only $n + 2\log(1/\epsilon)$ bits of key are required. Both of these constructions are quite simple, and are proven secure using the same Fourier-analytic technique.

The final construction has a more sophisticated proof, but allows for a length-preserving scheme with slightly better dependence on the number of qubits:

$$n + \min\{2\log n + 2\log(1/\epsilon), \log n + 3\log(1/\epsilon)\} + O(1)$$

bits of key. The right-hand term provides a better bound when $\epsilon > \frac{1}{n}$.

Randomization Schemes for Other Norms? Definition 1 measures leakage with respect to the trace norm on density matrices, $\|\cdot\|_{tr}$. This is good enough for encryption since the trace norm captures distinguishability of states. However, Hayden et al. [8] also considered randomization schemes which give guarantees with respect to a different norm, the operator norm.

A guarantee on the operator norm implies a guarantee for the trace norm, but schemes with the operator norm guarantee also have a host of less cryptographic applications, for example: constructing efficient quantum data hiding schemes in the LOCC (local operation and classical communication) model; exhibiting "locked" classical correlations in quantum states [8]; relaxed authentication of quantum states using few bits of key [9]; and transmitting quantum states over a classical channel using $n + o(n)$ bits of communication, rather than the usual $2n$ bits required for quantum teleportation [5].

More formally, for a $d \times d$ Hermitian matrix A with eigenvalues $\{\lambda_1, \lambda_2, ..., \lambda_d\}$, the *operator norm* (or ∞-norm) is the largest eigenvalue, $\|A\|_\infty = \max|\lambda_i|$, the Frobenius norm is the Euclidean length of the vector of eigenvalues, $\|A\|_2 = (\sum_i \lambda_i^2)^{1/2}$, and the trace norm is the sum of the absolute values of the eignvalues, $\|A\|_{tr} = \sum_i |\lambda_i|$. It is easy to see the chain of inequalities:

$$\|A\|_{tr} \le \sqrt{d}\,\|A\|_2 \le d\,\|A\|_\infty .$$

We can then state the condition for a map $\mathcal{E}$ to be ϵ-randomizing map for n qubits in three forms of increasing strength. For all input states ρ on n qubits:

$$\left\|\mathcal{E}(\rho) - \tfrac{1}{2^n}\mathbb{I}\right\|_{tr} \leq \epsilon; \qquad \left\|\mathcal{E}(\rho) - \tfrac{1}{2^n}\mathbb{I}\right\|_{2} \leq \epsilon/\sqrt{2^n}; \qquad \left\|\mathcal{E}(\rho) - \tfrac{1}{2^n}\mathbb{I}\right\|_{\infty} \leq \epsilon/2^n.$$

Our constructions satisfy the definition with respect to the Frobenius norm, but they are not known to satisfy the stronger operator-norm definition. This suggests two interesting questions. First, is it possible to prove that the other applications of state randomization schemes require only a guarantee on the Frobenius norm? Second, is it possible to design explicit (i.e. polynomial-time, deterministic) randomization schemes that give good guarantees with respect to the operator norm?

The remainder of this paper describes our constructions and their proofs of security.

2 Preliminaries

Small-Bias Spaces The bias of a random variable A in $\{0,1\}^n$ with respect to a string $\alpha \in \{0,1\}^n$ is the distance from uniform of the bit $\alpha \odot A$, where $\odot$ refers to the standard dot product on $\mathbb{Z}_2^n$:

$$\hat{A}(\alpha) = \mathbb{E}_A\left[(-1)^{\alpha \odot A}\right] = 2\Pr[\alpha \odot A = 0] - 1.$$

The function $\hat{A}$ is the Fourier transform of the probability mass function of the distribution, taken over the group $\mathbb{Z}_2^n$.

The bias of a set $S \in \{0,1\}^n$ with respect to α is simply the bias of the uniform distribution over that set. A set S is called δ-biased if the absolute value of its bias is at most δ for all $\alpha \neq 0^n$.

Small-bias sets of size polynomial in n and $1/\delta$ were first constructed by Naor and Naor [10]. Alon, Bruck et al. (ABNNR, [1]) gave explicit (i.e. deterministic, polynomial-time) constructions of δ-biased sets in $\{0,1\}^n$ with size $O(n/\delta^3)$. Constructions with size $O(n^2/\delta^2)$ were provided by Alon, Goldreich, et al. (AGHP, [2]). The AGHP construction is better when $\delta = o(1/n)$. In both cases, the i^{th} string in a set can be constructed in roughly n^2 time (regardless of δ).

One can sample a random point from a δ-biased space over $\{0,1\}^n$ using either $\log n + 3\log(1/\delta) + O(1)$ bits of randomness (using ABNNR) or using $2\log n + 2\log(1/\delta)$ bits (using AGHP).

Small-bias Set Families One can generalize small bias to *families* of sets (or random variables) by requiring that on average, the bias of a random set from the family with respect to every α is low [7]. Specifically, the expectation of the *squared* bias must be at most δ^2. Many results on δ-biased sets also hold for δ-biased families, which are easier to construct.

Definition 2. *A* family *of random variables (or sets)* $\{A_i\}_{i \in I}$ *is δ-biased if*

$$\mathbb{E}_{i \leftarrow I}\left[\hat{A}_i(\alpha)^2\right] \leq \delta^2 \textit{ for all } \alpha \neq 0^n.$$

Note that this is *not* equivalent, in general, to requiring that the expected bias be less than δ. There are two important special cases:

1. If S is a δ-biased set, then $\{S\}$ is a δ-biased set family with a single member;
2. A family of linear spaces $\{C_i\}_{i\in I}$ is δ-*biased* if no particular word is contained in the dual $C_i^\perp$ of a random space C_i from the family with high probability. Specifically:
$$\hat{C}_i(\alpha) = \begin{cases} 0 \text{ if } \alpha \notin C_i^\perp \\ 1 \text{ if } \alpha \in C_i^\perp \end{cases}$$
Hence a family of codes is δ-biased if and only if $\Pr_{i\leftarrow I}[\alpha \in C_i^\perp] \le \delta^2$, for every $\alpha \neq 0^n$. Note that to meet the definition, for linear codes the expected bias must be at most δ^2, while for a single set the bias need only be δ.

One can get a good δ-biased family simply by taking $\{C_i\}$ to be the set of all linear spaces of dimension k. The probability that any fixed non-zero vector α lies in the dual of a random space is exactly $\delta^2 = \frac{2^{n-k}-1}{2^n-1}$, which is at most 2^{-k}.

One can save some randomness in the choice of the space using a standard pairwise independence construction. View $\{0,1\}^n$ as $GF(2^n)$, and let $K \subseteq GF(2^n)$ be an additive subgroup of size 2^k. For every non-zero string a, let the space C_a be given by all multiples $a\kappa$, where $\kappa \in K$, and the product is taken in $GF(2^n)$. The family $\{C_a \mid a \in GF(2^n), a \neq 0\}$ has the same bias as the set of all linear spaces ($\delta < 2^{-k/2}$). To see this, let $\{\kappa_1, ..., \kappa_k\}$ be a basis of K (over $GF(2)$). A string α is in $C_a^\perp$ if and only if $\alpha \odot (a\kappa_1) = \cdots = \alpha \odot (a\kappa_k) = 0$. This is a system of k linearly independent constraints on a, and so it is satisfied with probability $\delta^2 = 2^{-k}$ when $a \leftarrow GF(2^n)$, and even lower probability when we restrict a to be non-zero. Choosing a set C_a from the family requires n bits of randomness.

Entropy of Quantum States As with classical distributions, there are several ways to measure the entropy of a quantum density matrix. We'll use the analogue of collision entropy (a.k.a. Renyi entropy).

For a classical random variable A on $\{0,1\}^n$, the collision probability of two independent samples of X is $p_c = \sum_a \Pr[A = a]^2$. The Renyi entropy of A is $-\log p_c$.

For a quantum density matrix ρ, the analogous quantity is $-\log \mathrm{Tr}(\rho^2)$. If the eigenvalues of ρ are $\{p_x\}$, then the eigenvalues of ρ^2 are $\{p_x^2\}$, and so $\mathrm{Tr}(\rho^2)$ is exactly the collision probability of the distribution obtained by measuring ρ in a basis of eigenvectors. $\sqrt{\mathrm{Tr}(\rho^2)}$ is called the Frobenius norm of ρ.

If ρ is the completely mixed state in d dimensions, $\rho = \frac{1}{d}\mathbb{I}$, then $\mathrm{Tr}(\rho^2)$ is $1/d$. The following fact states that any other density matrix for which this quantity is small must be very close to $\mathbb{I}$. The fact follows by applying the (standard) inequality $\mathrm{Tr}(|\Delta|)^2 \le d\mathrm{Tr}(\Delta^2)$ to the Hermitian matrix $\Delta = \rho - \mathbb{I}/d$.

Fact 1. *If ρ is d-dimensional quantum state and Tr$(\rho^2) \le \frac{1}{d}(1+\epsilon^2)$, then $D(\rho, \frac{1}{d}\mathbb{I}) \le \epsilon$.*

Pauli matrices The 2×2 Pauli matrices are generated by the matrices:

$$X = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \qquad Z = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}$$

The Pauli matrices are the four matrices $\{\mathbb{I}, X, Z, XZ\}$. These form a basis for the space of all 2×2 complex matrices. Since $XZ = -ZX$, and $Z^2 = X^2 = 1$, the set generated by X and Z is given by the Pauli matrices and their opposites: $\{\pm\mathbb{I}, \pm X, \pm Z, \pm XZ\}$.

If u and v are n-bit strings, we denote the corresponding tensor product of Pauli matrices by $X^u Z^v$. That is, if we write $u = (u1, ..., u_n)$ and $v = (v_1, ..., v_n)$, then

$$X^u Z^v = X^{u_1} Z^{v_1} \otimes \cdots \otimes X^{u_n} Z^{v_n}.$$

(The strings x and z indicate in which positions of the tensor product X and Z appear, respectively.) The set $\{X_u Z_v \mid u, v \in \{0,1\}^n\}$ forms a basis for the $2^n \times 2^n$ complex matrices. The main facts we will need are given below:

1. Products of Pauli matrices obey the group structure of $\{0,1\}^{2n}$ up to a minus sign. That is, $(X^u Z^v)(X^a Z^b) = (-1)^{a \odot v} X^{u \oplus a} Z^{v \oplus b}$.
2. Any pair of Pauli matrices either commutes or anti-commutes. Specifically, $(X^u Z^v)(X^a Z^b) = (-1)^{u \odot b + v \odot a}(X^a Z^b)(X^u Z^v)$.
3. The trace of $X^u Z^v$ is 0 if $(u, v) \neq 0^{2n}$ (and otherwise it is $\text{Tr}(\mathbb{I}) = 2^n$).
4. $(X^u Z^v)^\dagger = Z^v X^u = (-1)^{u \odot v} X^u Z^v$

Pauli matrices and Fourier Analysis The Pauli matrices form a basis for the set of all $2^n \times 2^n$ matrices. Given a density matrix ρ, we can write $\rho = \sum_{u,v \in \{0,1\}^n} \alpha_{u,v} X^u Z^v$. This basis is orthonormal with respect to the inner product given by $\frac{1}{2^n}\text{Tr}(A^\dagger B)$, where A, B are square matrices. That is, $\frac{1}{2^n}\text{Tr}((X^u Z^v)^\dagger X^a Z^b) = \delta_{a,u}\delta_{b,v}$.

Thus, the usual arithmetic of orthogonal bases (and Fourier analysis) applies. One can immediately deduce certain properties of the coefficients $\alpha_{u,v}$ in the decomposition of a matrix ρ. First, we have the formula $\alpha_{u,v} = \frac{1}{2^n}\text{Tr}(Z^v X^u \rho)$. Second, the squared norm of ρ is given by the squared norm of the coefficients, that is $\frac{1}{2^n}\text{Tr}(\rho^\dagger \rho) = \sum_{u,v} |\alpha_{u,v}|^2$. Since ρ is a density matrix, it is Hermitian ($\rho^\dagger = \rho$). One can use this fact, and the formula for the coefficients $\alpha_{u,v}$, to get a compact formula for the Renyi entropy (or Frobenius norm) in terms of the decomposition in the Pauli basis:

$$\text{Tr}(\rho^2) = \frac{1}{2^n} \sum_{u,v} |\text{Tr}(X^u Z^v \rho)|^2.$$

3 State Randomization and Approximate Encryption

3.1 Encrypting with a Small-Bias Space

The ideal quantum one-time pad applies a random Pauli matrix to the input [3]. Consider instead a scheme which first chooses a $2n$-bit string from some set with small bias δ (we will set δ later to be $\epsilon 2^{-n/2}$). If the set of strings is B we have:

$$\mathcal{E}(\rho_0) = \frac{1}{|B|} \sum_{(a,b) \in B} X^a Z^b \rho_0 Z^b X^a = \mathbb{E}_{a,b}\left[X^a Z^b \rho_0 Z^b X^a\right]$$

That is, we choose the key from the set B, which consists of $2n$-bit strings. To encrypt, we view a $2n$-bit string as the concatenation (a, b) of two strings of n bits, and apply the corresponding Pauli matrix.

(The intuition comes from the proof that Cayley graphs based on ϵ-biased spaces are good expanders: applying a Pauli operator chosen from a δ-biased family of strings to ρ_0 will cause all the Fourier coefficients of ρ_0 to be reduced by a factor of δ, which implies that the "collision probability" (Frobenius norm) of ρ_0 also gets multiplied by δ. We expand on this intuition below.)

As a first step, we can try to see if a measurement given by a Pauli matrix $X^u Z^v$ can distinguish the resulting ciphertext from a totally mixed state. More explicitly, we perform a measurement which projects the ciphertext onto one of the two eigenspaces of the matrix $X^u Z^v$. We output the corresponding eigenvalue. (All Pauli matrices have two eigenvalues with eigenspaces of equal dimension. The eigenvalues are always either -1 and 1 or $-i$ and i.)

To see how well a particular Pauli matrix $X^u Z^v$ will do at distinguishing, it is sufficient to compute

$$|\mathrm{Tr}(X^u Z^v \mathcal{E}(\rho_0))|.$$

This is exactly the statistical difference between the Pauli measurement's outcome and a uniform random choice from the two eigenvalues. We can compute it explicitly:

$$\begin{aligned}
\mathrm{Tr}(X^u Z^v \mathcal{E}(\rho_0)) &= \mathrm{Tr}\left(X^u Z^v \mathbb{E}_{(a,b)\in B}\left[X^a Z^b \rho_0 Z^b X^a\right]\right) \\
&= \mathbb{E}_{a,b}\left[\mathrm{Tr}(X^u Z^v X^a Z^b \rho_0 Z^b X^a)\right] \\
&= \mathbb{E}_{a,b}\left[\mathrm{Tr}(Z^b X^a X^u Z^v X^a Z^b \rho_0)\right] \\
&= \mathbb{E}_{a,b}\left[(-1)^{a\odot v + b\odot u}\right] \mathrm{Tr}(X^u Z^v \rho_0)
\end{aligned}$$

Since $a \odot v + b \odot u$ is linear in the concatenated $2n$-bit vector (a, b), we can take advantage of the small bias of the set B to get a bound:

$$|\mathrm{Tr}(X^u Z^v \mathcal{E}(\rho_0))| \leq \delta |\mathrm{Tr}(X^u Z^v \rho_0)| \qquad \text{when } (u, v) \neq 0^{2n}$$

Equivalently: if we express ρ_0 in the basis of matrices $X^u Z^v$, then each coefficient shrinks by a factor of at least δ after encryption. We can now bound the distance from the identity by computing $\mathrm{Tr}(\mathcal{E}(\rho_0)^2)$:

$$\begin{aligned}
\mathrm{Tr}(\mathcal{E}(\rho_0)^2) &= \frac{1}{2^n} \sum_{u,v} |\mathrm{Tr}(X^u Z^v \mathcal{E}(\rho_0))|^2 \\
&\leq \frac{1}{2^n} + \frac{\delta^2}{2^n} \sum_{(u,v)\neq 0^{2n}} |\mathrm{Tr}(X^u Z^v \rho_0)|^2 \leq \frac{1}{2^n}(1 + \delta^2 2^n \mathrm{Tr}(\rho_0^2))
\end{aligned}$$

Setting $\delta = \epsilon 2^{-n/2}$, we get approximate encryption for all states (since $\mathrm{Tr}(\rho_0^2) \leq 1$). Using the constructions of AGHP [2] for small-bias spaces, we get a polynomial-time scheme that uses $n + 2\log n + 2\log(1/\epsilon)$ bits of key.

3.2 A Scheme with Shorter Key Length

We can improve the key length of the previous scheme using δ-biased *families* of sets. The tradeoff is that the resulting states are longer: the ciphertext consists of n qubits and $2n$ classical bits. In classical terms, the encryption algorithm uses additional randomness which is not part of the shared key; in the quantum computing model, however, that randomness is "free" if one is allowed to discard ancilla qubits.

Lemma 1. *If $\{A_i\}_{i\in\mathcal{I}}$ is a family of subsets of $\{0,1\}^{2n}$ with average square bias δ^2, then the operator*

$$\mathcal{E}(\rho_0) = \mathbb{E}_{i\in\mathcal{I}}\left[|i\rangle\langle i| \otimes \mathbb{E}_{ab\in A_i}\left[X^a Z^b \rho_0 Z^b X^a\right]\right]$$

is an approximate encryption scheme for n qubits with leakage ϵ whenever $\delta \leq \epsilon 2^{-n/2}$.

Before proving the lemma, we give an example using the small-bias set family from the preliminaries. View the key set $\{0,1\}^k$ as an additive subgroup K of the field $\mathbb{F} = GF(2^{2n})$. For every element $a \in \mathbb{F}$, define the set $C_a = \{a\kappa | \kappa \in K\}$. The family $\{C_a\}$ has bias $\delta < 2^{-k/2}$ (Section 2). The corresponding encryption scheme takes a key $\kappa \in \{0,1\}^k \subseteq GF(2^{2n})$:

$$\mathcal{E}(\rho_0;\kappa) = \left[\begin{array}{l}\text{Choose } \alpha \leftarrow_R GF(2^{2n}) \setminus \{0\} \\ \text{Compute the product } \alpha\kappa \in GF(2^{2n}) \\ \text{Write } \alpha\kappa \text{ as a concatenation } (a,b), \text{ where } a,b \in \{0,1\}^n \\ \text{Output the classical string } \alpha \text{ and the quantum state } X^a Z^b \rho_0 Z^b X^a\end{array}\right.$$

With a quantum computer, random bits are not really necessary for choosing α; it is sufficient to prepare $2n$ EPR pairs and discard one qubit from each pair. For the scheme to be secure, the bias δ should be less than $\sqrt{\epsilon/2^n}$, and so the key only needs to be $n + 2\log(1/\epsilon)$ bits long. The main disadvantage is that the length of the ciphertext has increased by $2n$ classical bits.

Proof. As before, the proof will use elementary Fourier analysis over the hypercube $\mathbb{Z}_2^{2n}$, and intuition comes from the proof that Cayley graphs based on ϵ-biased set families are also expanders.

Think of the output of the encryption scheme as a single quantum state consisting of two systems: the first system is a classical string describing which member of the δ-biased family will be used. The second system is the encrypted quantum state. To complete the proof, it is enough to bound the collision entropy of the entire system by $\frac{1}{2^n|\mathcal{I}|}(1+\epsilon^2)$.

For each $i \in \mathcal{I}$ (that is, for each member of the set family), let ρ_i denote the encryption of ρ_0 with a random operator from the set A_i. The first step of the proof is to show that the collision entropy of the entire system is equal to the average collision entropy of the states ρ_i.

Claim. $\mathrm{Tr}(\mathcal{E}(\rho_0)^2) = \dfrac{1}{|\mathcal{I}|}\mathbb{E}_{i\leftarrow I}\left[\mathrm{Tr}(\rho_i^2)\right]$

Proof. We can write $\mathcal{E}(\rho_0) = \frac{1}{|\mathcal{I}|}\sum_i |i\rangle\langle i| \otimes \rho_i$. Then we have

$$\mathrm{Tr}(\mathcal{E}(\rho_0)^2) = \frac{1}{|\mathcal{I}|^2}\sum_{i,j} \mathrm{Tr}\big((|i\rangle\langle i||j\rangle\langle j|) \otimes \rho_i\rho_j\big)$$

Since $\langle i||j\rangle = \delta_{i,j}$, we get $\mathrm{Tr}(\mathcal{E}(\rho_0)^2) = \frac{1}{|\mathcal{I}|^2}\sum_i \mathrm{Tr}(\rho_i^2)$, as desired. □

Take any string $w = (u,v) \in \{0,1\}^{2n}$, where $u, v \in \{0,1\}^n$. Recall that $\hat{A}_i(u,v)$ is the ordinary Fourier coefficient (over $\mathbb{Z}_2^{2n}$) of the uniform distribution on A_i, that is $\hat{A}_i(u,v) = \mathbb{E}_{a \leftarrow A_i}\left[(-1)^{a \odot w}\right]$. From the previous proof, we know that

$$\mathrm{Tr}(X^u Z^v \rho_i) = \hat{A}_i(v,u) \cdot \mathrm{Tr}(X^u Z^v \rho_0).$$

We can now compute the average collision entropy of the states ρ_i. Using linearity of expectations:

$$\begin{aligned}
\mathbb{E}_i\left[\mathrm{Tr}(\rho_i^2)\right] &= \mathbb{E}_i\left[\frac{1}{2^n} + \frac{1}{2^n}\sum_{(u,v)\neq 0} |\mathrm{Tr}(X^u Z^v \rho_i)|^2\right] \\
&= \frac{1}{2^n} + \frac{1}{2^n}\sum_{(u,v)\neq 0} \mathbb{E}_i\left[|\mathrm{Tr}(X^u Z^v \rho_i)|^2\right] \\
&= \frac{1}{2^n} + \frac{1}{2^n}\sum_{(u,v)\neq 0} \mathbb{E}_i\left[\hat{A}_i(v,u)^2\right] |\mathrm{Tr}(X^u Z^v \rho_0)|^2
\end{aligned}$$

The expression $\mathbb{E}_i\left[\hat{A}_i(v,u)^2\right]$ is exactly the quantity bounded by the (squared) bias δ^2. As in the previous proof, the entropy $\mathrm{Tr}(\mathcal{E}(\rho_0)^2)$ is bounded by $\frac{1}{2^n|\mathcal{I}|}(1 + \delta^2 2^n \mathrm{Tr}(\rho_0^2))$. By our choice of δ, the entropy is at most $\frac{1}{2^n|\mathcal{I}|}(1+\epsilon^2)$, and so $\mathcal{E}(\rho_0^2)$ is within trace distance ϵ of the completely mixed state. □

3.3 Hybrid Construction

Let d be a prime between 2^n and 2^{n+1}. Then, it suffices to show how to randomize a state in a d-dimensional space $\mathcal{H}_d$ spanned by $|i\rangle$, $i \in \{0, 1, \ldots, d-1\}$, since a state on n qubits can be embedded into $\mathcal{H}_d$. We define X and Z on this space by $X|j\rangle = |(j+1) \bmod d\rangle$ and $Z|j\rangle = e^{2\pi i j/d}|j\rangle$. Notice that $X^j Z^k = e^{2\pi i(jk)/d} Z^k X^j$ and $(X^j Z^k)^\dagger = Z^{-k}X^{-j}$. (The definitions of X and Z are different than in the previous sections, since we are operating on a space of prime dimension).

We start with a construction that uses $n+1$ bits of randomness and achieves approximate encryption for $\epsilon = 1$. (Notice that this is a non-trivial security guarantee. The trace distance between perfectly distinguishable states is 2. Distance 1 means that the state cannot be distinguished from $\frac{I}{d}$ with success probability more than 3/4.) We will then extend it to any $\epsilon > 0$, using more randomness.

Let

$$\mathcal{E}(\rho) = \frac{1}{d}\sum_{a=1}^{d-1} X^a Z^{a^2} \rho Z^{-a^2} X^{-a}.$$

Claim.
$$Tr(\mathcal{E}(\rho)^2) \leq \frac{1}{d}(1 + Tr(\rho^2)).$$

Proof. Let $\rho' = \mathcal{E}(\rho)$.
$$Tr(\rho')^2 = \sum_{ij} \rho'_{ij}(\rho'_{ij})^* = \sum_i \rho'_{ii}(\rho'_{ii})^* + \sum_{i,j:i\neq j} \rho'_{ij}(\rho'_{ij})^*.$$

The first sum is equal to $d\frac{1}{d^2} = \frac{1}{d}$ because $\rho'_{ii} = \frac{1}{d}\sum_{k=1}^{d} \rho_{kk} = \frac{1}{d}$. To calculate the second sum, we split it into sums $S_t = \sum_i \rho'_{i,i+t}(\rho'_{i,i+t})^*$ for $t = 1, 2, \ldots, d-1$. (In the indices for ρ_{ij} and ρ'_{ij}, we use $i + t$ as a shortcut for $(i + t) \bmod d$.) We have
$$\rho'_{i,i+t} = \frac{1}{d}\sum_{a=0}^{d-1} w^{a^2 t}\rho_{i-a,i-a+t},$$

where w is the d^{th} root of unity.
$$\rho'_{i,i+t}(\rho'_{i,i+t})^* = \frac{1}{d^2}\left(\sum_{a=0}^{d-1} |\rho_{i+a,i+t+a}|^2 + \sum_{a,b,a\neq b} w^{(b^2-a^2)t}\rho_{i-a,i+t-a}(\rho_{i-b,i+t-b})^*\right)$$

Therefore,
$$S_t = \frac{1}{d}\sum_{i=1}^{d} |\rho_{i,i+t}|^2 + \frac{1}{d^2}\sum_{i\neq j} c_{i,j}\rho_{i,i+t}(\rho_{j,j+t})^*$$

where
$$c_{i,j} = \sum_a w^{((i+a)^2-(j+a)^2)t} = \sum_a w^{(i^2-j^2+2a(i-j))t} = w^{(i^2-j^2)t}\sum_a w^{a*2(i-j)t}.$$

Since d is a prime, $2(i-j)t$ is not divisible by d. Therefore, $\sum_a w^{a*2(i-j)t} = 0$, $c_{ij} = 0$, $S_t = \frac{1}{d}\sum_{i=1}^{d} |\rho_{i,i+t}|^2$ and
$$Tr((\rho')^2) = \frac{1}{d} + \frac{1}{d}\sum_{i\neq j} |\rho_{ij}|^2.$$

□

By Fact 1, $D(\mathcal{E}(\rho), \frac{I}{d}) \leq 1$.

We now improve this construction to any ϵ. Let B be an ϵ-biased set on $m = \lceil \log d \rceil$ bits. For $b \in \{0,1\}^m$, define a unitary transformation U_b as follows. Identify numbers $0, 1, \ldots, d-1$ with strings $x \in \{0,1\}^m$. Define $U_b|x\rangle = (-1)^{b\odot x}|x\rangle$, with $b \odot x$ being the usual (bitwise) inner product of b and x. (Note that U_b is just to the Z operator over a different group. It is the same Z operator used in the previous sections). Let
$$\mathcal{E}'(\rho) = \sum_{b\in B} U_b\rho U_b^\dagger \text{ and } \mathcal{E}''(\rho) = \mathcal{E}(\mathcal{E}'(\rho)).$$

We claim that $\mathcal{E}''$ is ϵ-approximate encryption scheme. W.l.o.g., assume that ρ is a pure state $|\psi\rangle = \sum_i c_i|i\rangle$. Then $\rho_{ij} = c_i c_j^*$. Let $\rho' = \frac{1}{|B|}\sum_{b\in B} U_b \rho U_b^\dagger$ be the result of encrypting ρ by $\mathcal{E}'$. Then,

$$\rho'_{xy} = \frac{1}{|B|}\sum_{b\in B}(-1)^{b\odot x + b\odot y}\rho_{xy} = \frac{1}{|B|}\sum_{b\in B}(-1)^{b\odot(x+y)}\rho_{xy}.$$

Since B is ϵ-biased, $|\rho'_{xy}| \le \epsilon|\rho_{xy}|$ for any x, y, $x \neq y$. Therefore, $\sum_{x\neq y}|\rho'_{xy}| \le \epsilon\sum_{x\neq y}|\rho_{xy}|$. Together with the Claim above and Fact 1, this implies that $\mathcal{E}''$ is ϵ-randomizing. The number of key bits used by $\mathcal{E}''$ is $n + \log|B| + O(1)$ which is $n + 2\log n + 2\log\frac{1}{\epsilon} + O(1)$ if the AGHP scheme is used and $n + \log n + 3\log\frac{1}{\epsilon} + O(1)$ if ABNNR is used. The first bound is the same as the one achieved by using small-bias spaces directly (Section 3.1). The second bound gives a better result when $\epsilon > \frac{1}{n}$.

Acknowledgements

We are grateful for helpful discussions with and comments from Claude Crépeau, Daniel Gottesman, Patrick Hayden, Debbie Leung, Sofya Raskhodnikova, Alex Samorodnitsky, and an anonymous referee.

References

1. Noga Alon, Jehoshua Bruck, Joseph Naor, Moni Naor, and Ronny Roth. Construction of asymptotically good low-rate error-correcting codes through pseudo-random graphs. IEEE Transactions on Information Theory, 38:509-516, 1992.
2. Noga Alon, Oded Goldreich, Johan Håstad, René Peralta. Simple Construction of Almost k-wise Independent Random Variables. Random Structures and Algorithms 3(3): 289-304.
3. Andris Ambainis, Michele Mosca, Alain Tapp, Ronald de Wolf. Private Quantum Channels. FOCS 2000: 547-553.
4. Howard Barnum, Claude Crépeau, Daniel Gottesman, Adam Smith, Alain Tapp. Authentication of Quantum Messages. FOCS 2002: 449-458.
5. Charles Bennett, Patrick Hayden, Debbie Leung, Peter Shor and Andreas Winter. Remote preparation of quantum states. ArXiv e-Print quant-ph/0307100.
6. Eli Ben-Sasson, Madhu Sudan, Salil P. Vadhan, Avi Wigderson. Randomness-efficient low degree tests and short PCPs via epsilon-biased sets. STOC 2003: 612-621.
7. Yevgeniy Dodis and Adam Smith. Encryption of High-Entropy Sources. Manuscript, 2003.
8. Patrick Hayden, Debbie Leung, Peter Shor and Andreas Winter. Randomizing quantum states: Constructions and applications. Comm. Math. Phys., to appear. Also ArXiv e-print quant-ph/0307104.
9. Debbie Leung, personal communication, 2004.
10. Joseph Naor, Moni Naor. Small-Bias Probability Spaces: Efficient Constructions and Applications. SIAM J. Comput. 22(4): 838-856 (1993).
11. Michael Nielsen, Isaac Chuang. *Quantum Computation and Quantum Information*. Cambridge University Press, 2000.
12. Alexander Russell, Hong Wang. How to Fool an Unbounded Adversary with a Short Key. EUROCRYPT 2002: 133-148.