Conditional Expectation Bounds with Applications in Cryptography

Kevin J. Compton∗

Computer Science and Engineering Division
University of Michigan - Ann Arbor
Ann Arbor, MI 48109-2212, USA

July 9, 2018

Abstract

We present two conditional expectation bounds. In the first bound, \(Z\) is a random variable with \(0 \leq Z \leq 1\), \(U_i (i < t)\) are i.i.d. random objects with each \(U_i \sim U\), and \(W_i = \mathbb{E}[Z|U_i]\) are conditional expectations whose average is \(W = (W_0 + \cdots + W_{t-1})/t\). We show for \(0 < \varepsilon \leq 1\) that \(\mathbb{E}[Z] \leq \frac{\mathbb{P}_U\{W > \varepsilon\}^t}{t} + \frac{t\varepsilon}{t}\). In the second bound we replace the i.i.d. property with a weaker property, the so-called \(\beta\)-i.i.d. property, where \(0 < \beta < 1\). The conclusion then is that \(\mathbb{E}[Z] \leq \alpha + \frac{\beta\mathbb{P}\{W > \varepsilon\}^t}{t} + \frac{t\varepsilon}{t}\), where \(\alpha = 1 - \beta\). We show how to produce \(\beta\)-i.i.d. random objects from random walks on hybrid expander-permutation directed graphs where the transition matrix of the expander graph has spectral gap \(\beta\). These results underlie many security proofs in cryptography, for example, the classical Yao-Goldreich result that strongly one-way functions exist if weakly one-way functions exist, and the result of Goldreich et al. showing security preserving reductions from weakly to strongly one-way functions.

1 Introduction

Consider a random variable \(Z\) on a probability space \((\Omega, \mathcal{F}, \mathbb{P})\) with \(0 \leq Z \leq 1\) and i.i.d. random objects \(U_i, 0 \leq i < t\), where \(U_i \sim U\) for some fixed \(U\). Take conditional expectations \(W_i = \mathbb{E}[Z|U_i]\) and let their average be \(W = (W_0 + \cdots + W_{t-1})/t\). From the law of iterated expectations and linearity of expectation, we know that \(\mathbb{E}[W] = \mathbb{E}[Z]\), but we would expect \(W\) to be more concentrated around this expectation since it is an average. However, it is not an average of independent random variables, so this concentration is difficult to quantify.

As a concrete example, let \(\Omega = [0, 1]^t\), the unit \(t\)-cube, and \(\mathbb{P}\) be the usual product measure on the \(\sigma\)-algebra \(\mathcal{F}\) of Borel sets in \(\Omega\). Fix \(p\) strictly between 0 and 1, and let \(Z(y_0, y_1, \ldots, y_{t-1})\) be the indicator function of the event \([0, p^{1/t}]^t\). Clearly, \(\mathbb{E}[Z] = p\). The \(t\) projection functions \(U_i(y_0, y_1, \ldots, y_{t-1}) = y_i\) are i.i.d. random objects. A simple computation shows that \(W_i = \mathbb{E}[Z|U_i]\) is given by

\[
W_i(x) = \begin{cases} 
p^{1-1/t}, & \text{if } x \leq p^{1/t}, \\
0, & \text{otherwise.}
\end{cases}
\]

∗kjc@umich.edu
The random variables $W_i$ are identical, so $W = W_i$.

Now compare: $Z$ takes the value 1 with probability $p$; $W$ takes the value $p^{1-1/t}$ with probability $p^{1/t}$. For large $t$, $W$ is more concentrated, taking a value closer to the expectation $p$ on an event of probability $\mathbb{E}[Z]^{1/t}$. In fact, for any $\varepsilon$ such that $0 < \varepsilon < p^{1-1/t}$, $\mathbb{P}\{W > \varepsilon\} = \mathbb{E}[Z]^{1/t}$.

Our first main result, Theorem 3.1(i), says that something close to this holds in general: when $0 \leq Z \leq 1$ and $0 < \varepsilon < 1$,

$$\mathbb{E}[Z] \leq \mathbb{P}_U\{W > \varepsilon\}^t + t\varepsilon.$$  

(1)

Another result, Theorem 3.2(i), dispenses with the hypothesis that the random objects $U_i$ are identically distributed and the bound is given in terms of product of tail probabilities $\mathbb{P}_{U_i}\{W_i > \varepsilon_i\}$ plus a correction term $t\varepsilon$, where $\varepsilon$ is the average of the values $\varepsilon_i$.

Our second main result, Theorem 3.1(ii), weakens the hypothesis by replacing independence of the random objects $U_i$ with a property we call $\beta$-independence (where $0 < \beta \leq 1$). By this we mean that for all events $T_i$, if we set $\mathbb{P}\{U_i \in T_i\} = 1 - \nu_i$, then

$$\mathbb{P}\{\bigwedge_{0 \leq i < t} U_i \in T_i\} \leq \prod_{0 \leq i < t} (1 - \beta \nu_i).$$

(2)

With this hypothesis (but still requiring identical distribution of the random objects $U_i$) we conclude that

$$\mathbb{E}[Z] \leq (\alpha + \beta \mathbb{P}_U\{W > \varepsilon\})^t + t\varepsilon,$$

(3)

where $\alpha = 1 - \beta$. It is not difficult to see that 1-independence is equivalent to independence, so that part (i) of Theorem 3.1 is a special case of part (ii). We should note that Theorem 3.1(ii) does not require full $\beta$-independence. It suffices that (2) should hold when the events $T_i$ are identical:

$$\mathbb{P}\{\bigwedge_{0 \leq i < t} U_i \in T\} \leq (1 - \beta \nu)^t,$$

(4)

where $\mathbb{P}\{U_i \in T\} = 1 - \nu$. However, we show a related result, Theorem 3.2, which dispenses with the condition that the random objects $U_i$ be identically distributed, but requires full $\beta$-independence given by (2).

We regard these results as useful probabilistic tools similar to Chernoff bounds. But whereas Chernoff bounds are upper bounds on tail distributions for averages of independent random variables, Theorem 3.1(i) (formula (1) above) is a lower bound on tail distributions for averages of conditional expectations with respect to independent random objects. It is not surprising, then, that conditional expectation bounds turn up in security proofs for cryptographic constructions. A Chernoff bound will show that an efficient probabilistic algorithm has a high probability of returning the correct result. A conditional expectations bound will show that all efficient probabilistic algorithms have a low probability of breaking a cryptographic construction.

The earliest example of such a security proof was for Yao’s construction [38] of a strongly one-way function as a direct power of a weakly one-way function. (See Section 5 for precise definitions.) The idea is simple. From $F$, a weakly one-way function, define

$$F'(x_0 x_1 \cdots x_{t-1}) = (F(x_0), F(x_1), \ldots, F(x_{t-1})).$$
where $x_0, x_1, \ldots, x_{t-1} \in \{0, 1\}^n$ and $t = t(n)$ is suitably chosen polynomial. The proof that $F'$ is strongly one-way is not so simple.

The proof appeared first in an online draft of a text by Goldreich [13]. Let us formulate it in terms of a reduction between cryptographic primitives (as found in [27, 9, 25]). For one-way functions, a reduction is a pair $(R, R^*)$ where $R$ is an efficient transformation taking $F$ to $F'$ and $R^*$ is an efficient transformation taking each randomized function $G'$ attempting to invert $F'$ to a randomized function $G$ attempting to invert $F$. There is also a condition, detailed below, relating the probability that $G$ inverts $F$ to the probability that $G'$ inverts $F'$. In this situation we say $(R, R^*)$ reduces $F$ to $F'$. (It would be more accurate to say this is a reduction of the invertibility problem for $F$ to the invertibility problem for $F'$.)

Goldreich’s proof implicitly gives such a reduction. $R$ is the the direct power construction. The crux of the proof is to specify $R^*$ so that if $G'$ is probabilistic polynomial time (p.p.t.) computable then so is $G$, and whenever $G$ inverts $F$ with probability significantly less than 1, $G'$ inverts $F'$ with negligible probability. Reformulated, this is an instance of [1]. One then observes that to any p.p.t. function $G'$ attempting to invert $F'$, we may apply $R^*$ to obtain $G$. If $F$ is weakly one-way, $G$ inverts $F$ with probability significantly less than 1 and, thus, $G'$ inverts $F'$ with negligible probability; i.e., $F'$ is strongly one-way.

Goldreich et al. [15] later pointed out a drawback to the direct power reduction: it is not security preserving (see section 7 for the precise definition of security preserving). The reason is that the forward transformation $R$ takes $F$ with input size $n$ to $F'$ with input size $nt(n)$; $R^*$, then, takes $G'$ with input size $nt(n)$ to $G$ with input size $n$. It follows that if $S(n)$ is the security of $F$ against $G$ and $S'(n)$ is the security of $F'$ against $G'$, then $S(n)$ is of the same order as $S'(nt(n))$, which is much larger than $S'(n)$ when $S'$ has superpolynomial growth.

Goldreich et al. [15] (cf. also [13]) remedied this deficiency by replacing the direct power with a more elaborate expander graph construction which controls input size blowup. It gives a security preserving reduction in the restricted case where $F$ is a weakly one-way permutation and $F'$ is a strongly one-way permutation, but now rather than $S'(nt(n))$, the security is now $S'(n + \omega(\log n))$, which is of the same order as $S'(n)$ (in a sense made precise in section 7). In this case, Theorem 3.1(ii) does not suffice. If we frame the argument of [15] in terms of conditional expectation averages, we see that we no longer have independent random objects, but we do have random objects exhibiting a weaker property we call $\beta$-independence. This is the motivation for Theorem 3.1(ii).

The $\beta$-independence property is closely related to the hitting property of expander graphs. Goldreich et al. [15] employ directed graphs $G'$ which combine expander graphs with a weakly one-way permutation. They show that when $U_i$ is the $i$-th point on a random directed $t$-walk in $G'$, then a hitting property somewhat along the lines of [1] holds, where $T$ is a set of vertices in $G'$ and we need to compute the probability that it is hit by a random $t$-path.

Later work [9, 18, 5] provided security preserving reductions, again in restricted cases, using hash functions rather than expander graphs to control input size blowup.

Subsequent investigations showed hardness amplification using similar methods applied to various cryptographic primitives such as collision-resistant hash functions [7], encryption schemes [11], weakly verifiable puzzles [6, 22, 24], signature schemes and message authentication codes [10], commitment schemes [19, 8], pseudorandom functions and pseudorandom generators [10, 31], block ciphers [28, 32, 30, 37], and interactive protocols [3, 33, 20, 17].
The outline of the paper is as follows. In Section 2 we review terminology and results from probability theory, particularly those concerning conditional expectations. In Section 4 we prove the main results. In Section 5 we address the question of whether the \( \beta \)-independence property holds in hybrid expander-permutation directed graphs and show that the answer is yes. Although this result is not needed for the applications that follow, it is of independent interest and generalizes earlier work on hitting properties of expander graphs. In Section 6 we define and review the properties of weakly and strongly one-way functions. In Section 7 we use the conditional expectation inequality to show Goldreich's result that existence of weakly one-way functions implies the existence of strongly one-way functions. In Section 8 Goldreich et al. make a stronger assumption that there is a polynomial-time computable edge-coloring. We will also show in in Section 7 that the edge-coloring assumption is not necessary.

2 Probability Background

We recall some basic terminology and notation from probability theory. For more details see Ash and Loève.

Throughout \((\Omega, \mathcal{F}, \mathbb{P})\) is a probability space where \(\Omega\) the sample space, \(\mathcal{F}\) is a \(\sigma\)-algebra of events on \(\Omega\), and \(\mathbb{P}: \mathcal{F} \rightarrow [0, 1]\) is a \(\sigma\)-additive probability measure on \(\mathcal{F}\). A pair \((\Omega, \mathcal{F})\) where \(\mathcal{F}\) is a \(\sigma\)-algebra of events on \(\Omega\) is a measurable algebra.

Event \(S\) holds almost everywhere (a.e.) if \(\mathbb{P}(S) = 1\). The indicator function for an event \(S\) is

\[
1_S(x) = \begin{cases} 
1, & \text{if } x \in S, \\
0, & \text{otherwise,}
\end{cases}
\]

\(\mathbb{P}\) is uniformly distributed (on a finite space) if \(\Omega\) is finite, \(\mathcal{F} = 2^\Omega\) and \(\mathbb{P}(S) = |S|/|\Omega|\) for every \(S \subseteq \Omega\).

A random object on \((\Omega, \mathcal{F}, \mathbb{P})\) is a measurable function \(X\) from measurable algebra \((\Omega, \mathcal{F})\) to measurable algebra \((\Psi, \mathcal{F}')\); elements of \(\Psi\) are objects. When \(\Psi = \mathbb{R}\) and \(\mathcal{F}'\) is the \(\sigma\)-algebra of Borel sets \(\mathbb{R}\) \(X\) is a random variable. When \(\Psi = \mathbb{R} \cup \{-\infty, \infty\}\) and \(\mathcal{F}'\) is \(\sigma\)-algebra of extended Borel sets, \(X\) is an extended random variable. (In general, we may substitute other terms for object. A random \(n\)-bit string, for instance, is a random object \(X\) where \(\Psi = \{0, 1\}^n\) and \(\mathcal{F}'\) is the power set of \(\Psi\).)

For \(S \subseteq \Psi\) denote the inverse image \(X^{-1}[S] = \{a \in \Omega : X(a) \in S\}\) by \(\{X \in S\}\). Similarly abbreviate other events determined by \(X\). For example, \(\{X = b\}\) is the event \(\{a \in \Omega : X(a) = b\}\). Also, delete superfluous parentheses whenever possible, e.g., by writing \(\mathbb{P}\{X \in S\}\) rather than \(\mathbb{P}\{\{X \in S\}\}\). For clarity, we also write

\[
\bigcap_{i<t} \{X_i \in S_i\}
\]

rather than

\[
\bigcap_{i<t} \{X_i \in S_i\}.
\]

\[1\] The set of Lebesgue measurable sets would serve as well.
A simple consequence of our definition is that
\[
\{Y \in \{X \in S\}\} = \{X \circ Y \in S\},
\]
where \(\circ\) is functional composition.

For a random object \(X: (\Omega, \mathcal{F}) \to (\Psi, \mathcal{F}')\), the probability measure \(\mathbb{P}_X\) induced by \(X\) (or the distribution of \(X\)) is
\[
\mathbb{P}_X(S) = \mathbb{P}\{X \in S\}.
\]
The \(\sigma\)-algebra induced by \(X\) (a subalgebra of \(\mathcal{F}\)) is
\[
\sigma(X) = \{\{X \in S\} : S \in \mathcal{F}'\}.
\]

Random objects \(U_i: \Omega_i \to \Psi, i \in I\), are identically distributed if they induce the same probability measure, i.e., if for all \(i \in I\), \(\mathbb{P}_{U_i} = \mathbb{P}_U\) for some fixed \(U\); we write \(U_i \sim U\) to indicate and \(U_i\) and \(U\) are identically distributed. Random objects \(U_i: (\Omega, \mathcal{F}) \to (\Psi_i, \mathcal{F}'_i), i \in I\), are independent if for all events \(S_i \in \mathcal{F}'_i, i \in I\), events \(\{U_i \in S_i\}\) are mutually independent. Random objects that are independent and identically distributed are \(\text{i.i.d.}\).

The following definition extending the notion of independent random objects is not found in standard probability texts.

**Definition.** Fix \(\beta, 0 \leq \beta \leq 1\) and let \(\alpha + \beta = 1\). Suppose \(S_i, i \in I\), are events with
\[
\mathbb{P}(S_i) = \mu_i = 1 - \nu_i.
\]
Events \(S_i\) are mutually \(\beta\)-independent if for all \(J \subseteq I\)
\[
\mathbb{P}(\bigcap_{i \in J} S_i) \leq \prod_{i \in J} (1 - \beta \nu_i) = \prod_{i \in J} (\alpha + \beta \mu_i).
\]
Random objects \(U_i: (\Omega, \mathcal{F}) \to (\Psi_i, \mathcal{F}'_i), i \in I\), are \(\beta\)-independent if for all events \(S_i \in \mathcal{F}'_i, i \in I\), events \(\{U_i \in S_i\}\) are mutually \(\beta\)-independent. If random objects \(U_i: (\Omega, \mathcal{F}) \to (\Psi, \mathcal{F}')\) are \(\beta\)-independent and identically distributed, they are \(\beta\)-\(\text{i.i.d.}\).

It is not difficult to show that 1-independence of random objects is equivalent to independence. Also, 0-independence holds in all cases.

\(\mathbb{E}[Z]\) is the expectation of a random variable \(Z\) and \(\mathbb{E}_S[Z]\) denotes \(\mathbb{E}[1_S \cdot Z]\). The change of variables formula for expectation says that when \(X: (\Omega, \mathcal{F}) \to (\Psi, \mathcal{F}')\) is a random object and \(Z\) is an extended random variable on \((\Psi, \mathcal{F}', \mathbb{P}_X)\), then
\[
\mathbb{E}[Z] = \mathbb{E}[Z \circ X].
\]
Substituting \(1_S \cdot Z\) for \(Z\) gives
\[
\mathbb{E}_S[Z] = \mathbb{E}_{\{X \in S\}}[Z \circ X]. \tag{5}
\]
The notation for expectation obscures a fundamental distinction in these two equations (and also several equations below). The expectation on the left side of each equation is with respect to \(\mathbb{P}_X\), while the one on the right side is with respect to \(\mathbb{P}\).

As usual, \(\mathbb{P}(T|S)\) is the conditional probability of an event \(T\) given event \(S\) and \(\mathbb{E}[Y|S] = \mathbb{E}_S[Y]/\mathbb{P}(S)\) is the conditional expectation of an extended random variable \(Y\) given event
When the expectations are defined. (When \( \mathbb{P}(S_i) = 0 \) we take \( \mathbb{E}[Y|S_i] \) to be 0.)

The conditional expectation \( \mathbb{E}[Y|X] \), where \( Y \) is an extended random variable on \((\Omega, \mathcal{F}, \mathbb{P})\) and \( X: (\Omega, \mathcal{F}) \to (\Psi, \mathcal{F}')\) is a random object, is an extended random variable \( Z \) on \((\Psi, \mathcal{F}', \mathbb{P}_X)\). The diagram in Figure 1 summarizes the relationship between \( Y \) and \( \mathbb{E}[Y|X] \). This diagram does not commute. Rather, \( Z \) is the random variable on \((\Psi, \mathcal{F}', \mathbb{P}_X)\) that minimizes the mean square difference between \( Y \) and \( Z \circ X \). It can be shown that such a \( Z \) exists and is obtained by first defining a measure \( \mathbb{M}(S) = \mathbb{E}_{\{X \in S\}}[Y] \) on \((\Psi, \mathcal{F}', \mathbb{P}_X)\). \( \mathbb{M} \) is absolutely continuous with respect to \( \mathbb{P}_X \) so by the Radon-Nikodym Theorem (see Ash [2]) there is a random variable \( Z \) such that \( \mathbb{M}(S) = \mathbb{E}_S[Z] \). \( Z \) is unique a.e. with respect to \( \mathbb{P}_X \); i.e., if \( \mathbb{M}(S) = \mathbb{E}_S[Z'] \) for all \( S \in \mathcal{F}' \), then \( Z = Z' \) a.e. Let \( Z = \mathbb{E}[Y|X] \). (Some texts define \( \mathbb{E}[Y|X] \) to be \( Z \circ X \) rather than \( Z \), but the more usual notation for \( Z \circ X \) is \( \mathbb{E}[Y|S] \), where \( S = \sigma(X) \), since \( \sigma(X) \) determines \( Z \circ X \) a.e.) (When \((\Omega, \mathcal{F}, \mathbb{P})\) is a discrete space (i.e., \( \Omega \) is countable and \( \mathcal{F} = 2^\Omega \)), the definition of conditional expectation simplifies: \( Z = \mathbb{E}[Y|X] \) is given by \( Z(a) = \mathbb{E}[Y|X = a] \) whenever \( \mathbb{P}\{X = a\} > 0 \) and is assigned an arbitrary value, such as 0, otherwise.)

By construction we have
\[
\mathbb{E}_{\{X \in S\}}[Y] = \mathbb{E}_S[\mathbb{E}[Y|X]]
\]
for all \( S \in \mathcal{F}' \). Taking \( S = \Omega \) gives the law of iterated expectation
\[
\mathbb{E}[Y] = \mathbb{E}[\mathbb{E}[Y|X]].
\]

Conditional expectation, like expectation, is a linear operator so for random object \( X \), random variables \( Y, Z \), and real numbers \( a, b \).

\[
\mathbb{E}[aY + bZ|X] = a \mathbb{E}[Y|X] + b \mathbb{E}[Z|X].
\]

It is also monotonic: \( Y \leq Z \) implies \( \mathbb{E}[Y|X] \leq \mathbb{E}[Z|X] \). A simple consequence of linearity and monotonicity is that when \( S = \bigcup_{i \in J} S_i \) and \( Z \geq 0 \), then
\[
\mathbb{E}_S[Z|X] \leq \sum_{i \in J} \mathbb{E}_{S_i}[Z|X]. \quad (6)
\]
3 Conditional Expectation Bounds

We prove the main results of the paper.

**Theorem 3.1.** Let $Z$ be a random variable such that $0 \leq Z \leq 1$, and

$$U_i: (\Omega, \mathcal{F}) \to (\Psi, \mathcal{F}'), i < t,$$

be identically distributed with $U_i \sim U$ for some fixed $U$. Set $W_i = \mathbb{E}[Z|U_i]$, $W = (W_0 + W_1 + \cdots + W_{t-1})$ and take any $0 < \varepsilon < 1$.

(i). If the random objects $U_i$ are independent,

$$\mathbb{E}[Z] \leq \mathbb{P}_U\{W > \varepsilon\}^t + t\varepsilon.$$

(ii). If the random objects $U_i$ are $\beta$-independent,

$$\mathbb{E}[Z] \leq (\alpha + \beta \mathbb{P}_U\{W > \varepsilon\})^t + t\varepsilon,$$

where $\alpha + \beta = 1$.

**Proof.** Let $T = \{W > \varepsilon\}$, $S_i = \{U_i \in T\}$, and $S = \bigcap_{i < t} S_i$. By the law of total expectation

$$\mathbb{E}[Z] = \mathbb{E}[Z|S] \mathbb{P}(S) + \mathbb{E}[Z|\bar{S}] \mathbb{P}(\bar{S}),$$

bound the left side of this equation as follows.

$$\mathbb{E}[Z|S] \leq 1$$

since $Z \leq 1$. Independence of the random objects $U_i$ in part (i) of the theorem implies

$$\mathbb{P}(S) = \mathbb{P}(\bigcap_{i < t} S_i) = \prod_{i < t} \mathbb{P}_{U_i}(T) = \mathbb{P}_U\{W > \varepsilon\}^t,$$

and $\beta$-independence in part (ii) implies

$$\mathbb{P}(S) \leq \prod_{i < t} (\alpha + \beta \mathbb{P}_{U_i}(T)) = (\alpha + \beta \mathbb{P}_U\{W > \varepsilon\})^t.$$

From $\bar{S} = \bigcup_{i < t} \bar{S_i}$ and (6) we have

$$\mathbb{E}[Z|\bar{S}] \mathbb{P}(\bar{S}) = \mathbb{E}_{\bar{S}}[Z] \leq \sum_{i < t} \mathbb{E}_{\bar{S_i}}[Z],$$

which by (5) is equal to

$$\sum_{i < t} \mathbb{E}_{\bar{T}}[\mathbb{E}[Z|U_i]] = \sum_{i < t} \mathbb{E}_{\{W \leq \varepsilon\}}[W_i]$$

$$= \mathbb{E}_{\{W \leq \varepsilon\}}[\sum_{i < t} W_i]$$

$$= \mathbb{E}_{\{W \leq \varepsilon\}}[tW]$$

$$\leq t\varepsilon,$$

since $\bar{S_i} = \{U_i \in \bar{T}\}$ and $\bar{T} = \{W \leq \varepsilon\}$.

Substitution into (7) completes the proof. \(\square\)
Remark. We may relax the hypotheses of Theorem 3.1 so that rather than independence in part (i), we need only require that for every event $T$,

$$\mathbb{P}\{\bigwedge_{i<t} U_i \in T\} = \prod_{i<t} \mathbb{P}\{U_i \in T\},$$

and similarly for $\beta$-independence in part (ii).

We do not use the following theorem in this paper, but it may prove useful in other contexts where the random variables $U_i$ are not identically distributed.

**Theorem 3.2.** Let $Z$ be a random variable such that $0 \leq Z \leq 1$,

$$U_i: (\Omega, \mathcal{F}) \to (\Psi, \mathcal{F}', i < t),$$

be random objects with $W_i = \mathbb{E}[Z|U_i]$, $0 < \varepsilon_i < 1$ for $i < t$, and $\varepsilon = t^{-1} \sum_{i<t} \varepsilon_i$.

(i). If the random objects $U_i$ are independent,

$$\mathbb{E}[Z] \leq \prod_{i<t} \mathbb{P}_{U_i}\{W_i > \varepsilon_i\} + t\varepsilon.$$

(ii). If the random objects $U_i$ are $\beta$-independent,

$$\mathbb{E}[Z] \leq \prod_{i<t} (\alpha + \beta \mathbb{P}_{U_i}\{W_i > \varepsilon_i\}) + t\varepsilon,$$

where $\alpha + \beta = 1$.

**Proof.** Let $T_i = \{W_i > \varepsilon_i\}$, $S_i = \{U_i \in T_i\}$, and $S = \bigcap_{i<t} S_i$. As in the previous theorem, we bound the right side of

$$\mathbb{E}[Z] = \mathbb{E}[Z|S] \mathbb{P}(S) + \mathbb{E}[Z|\overline{S}] \mathbb{P}(\overline{S}).$$

As before, $\mathbb{E}[Z|S] \leq 1$. Also, for part (i) of the theorem,

$$\mathbb{P}(S) = \mathbb{P}\left(\bigcap_{i<t} S_i\right) = \prod_{i<t} \mathbb{P}_{U_i}(T_i),$$

and for part (ii),

$$\mathbb{P}(S) \leq \prod_{i<t} (\alpha + \beta \mathbb{P}_{U_i}(T_i)).$$

We have

$$\mathbb{E}[Z|\overline{S}] \mathbb{P}(\overline{S}) \leq \sum_{i<t} \mathbb{E}_{\overline{T_i}}[Z],$$

This last summation is equal to

$$\sum_{i<t} \mathbb{E}_{\overline{T_i}}[\mathbb{E}[Z|U_i]] = \sum_{i<t} \mathbb{E}_{\{W_i \leq \varepsilon_i\}}[W_i] \leq \sum_{i<t} \varepsilon_i,$$

since $\overline{T_i} = \{U_i \in T_i\}$ and $\overline{T_i} = \{W_i \leq \varepsilon_i\}$. The theorem follows by substitution. \qed
4 Expander Graphs and $\beta$-Independence.

Theorem 4.5, the main result of this section, provides a natural construction of $\beta$-independent random objects based on hybrid expander-permutation directed graphs.

Let $A$ be a Hermitian matrix of dimension $N$ (cf. Horn and Johnson [21] for basic results concerning Hermitian matrices). $A$ has $N$ real eigenvalues. List them (with repetitions according to multiplicities) in nonincreasing order:

$$\lambda_0(A) \geq \lambda_1(A) \cdots \geq \lambda_{N-1}(A).$$

We write $\lambda_i$ rather than $\lambda_i(A)$ when matrix $A$ is clear from context. Because $A$ is Hermitian, there is an orthonormal basis $u_0, u_1, \ldots, u_{N-1} \in \mathbb{R}^n$, where $u_i$ is an eigenvector associated with eigenvalue $\lambda_i(A)$.

We consider only real-valued matrices; under this restriction, the Hermitian matrices are precisely the symmetric matrices. We also consider only real-valued vectors. We write vectors in lower case boldface and denote the transpose of vector $v$ by $v^T$, so the inner product of column vectors $v$ and $w$ is $\langle v, w \rangle = v^T \cdot w$ and $|v|^2 = \langle v, v \rangle$. The Cauchy-Schwarz inequality states that $|\langle v, w \rangle| \leq |v| |w|$.

Let $G = (V, E)$ be a $d$-regular undirected graph and $A$ be its transition matrix; i.e., $A = (a_{i,j})$, where

$$a_{i,j} = \begin{cases} 1/d, & \text{if there is an edge from vertex } i \text{ to vertex } j, \\ 0, & \text{otherwise.} \end{cases}$$

$A$ is symmetric, nonnegative, and doubly stochastic. If $G$ is connected, then $A$ is irreducible (cf. Seneta [36] for basic results concerning nonnegative matrices). By the Perron-Frobenius Theorem, the largest eigenvalue of $A$ is $\lambda_0 = 1$, the common row sum. Also, this is a simple eigenvalue so $\lambda_0 > \lambda_1$ and no other eigenvalue is larger in magnitude than $\lambda_0$; it follows that $\lambda_{N-1} \geq -1$. If $G$ is not bipartite, then $\lambda_{N-1} > -1$. Under these conditions, define $\alpha = \alpha(A)$ to be the second largest eigenvalue magnitude of $A$, and $\beta$ to be the spectral gap of $A$, i.e., the difference between the largest and second largest eigenvalue magnitudes. That is,

$$\alpha = \max(|\lambda_1(A)|, |\lambda_{N-1}(A)|)$$

$$\beta = 1 - \alpha$$

**Definition.** A connected, non-bipartite graph $G = (V, E)$ is an $(N, d, \alpha)$-expander graph if $|V| = N$, $G$ is $d$-regular, and the second largest eigenvalue magnitude of its transition matrix is at most $\alpha$.

For the remainder of the section, $A$ is the transition matrix for an $(N, d, \alpha)$-expander graph $G$, $\alpha + \beta = 1$, and vectors $u_0, u_1, \ldots, u_{N-1} \in \mathbb{R}^n$ form an orthonormal basis, where $u_i$ is an eigenvector associated with eigenvalue $\lambda_i(A)$. We take $u_0$ to be $N^{-1/2}(1, 1, \ldots, 1)^T$.

For reference, we collect a few simple facts.

**Proposition 4.1.** Let $V_0$ be the subspace of $\mathbb{R}^n$ spanned by $u_0$ and $V_1$ its orthogonal space, the subspace spanned by $u_1, \ldots, u_{N-1}$.

(i). $V_0$ and $V_1$ are invariant under the action of $A$. 
(ii). Every vector \( \mathbf{w} \in \mathbb{R}^n \) can be decomposed as a sum of two vectors \( \mathbf{w} = \mathbf{x} + \mathbf{y} \) where \( \mathbf{x} \in V_0 \) and \( \mathbf{y} \in V_1 \).

(iii). If \( \mathbf{x} \in V_0 \), then \( A\mathbf{x} = \mathbf{x} \). If \( \mathbf{y} \in V_1 \), then \( |A\mathbf{y}| \leq \alpha|\mathbf{y}| \).

**Definition.** The projection matrix for set \( S \subseteq \{0,1,\ldots,N-1\} \) is the matrix \( \mathbf{P} = (p_{i,j}) \) where \( p_{i,i} = 1 \) if \( i \in S \) and all other entries \( p_{i,j} \) are 0.

Applying \( \mathbf{P} \) to a vector zeros out any coordinate in a position \( i \notin S \) so, in particular, \( |\mathbf{P}\mathbf{u}_0|^2 = |S|/N \).

**Lemma 4.2.** Let \( \mathbf{A}, \quad N, \quad d, \quad \alpha, \) and \( \beta \) be as above. Take projection matrices \( \mathbf{P} \) and \( \mathbf{P}' \) for \( S \) and \( S' \subseteq \{0,1,\ldots,N-1\} \) and set \( \mu = |S|/N, \mu' = |S'|/N \). Then for all vectors \( \mathbf{v} \)

\[
|\mathbf{P}\mathbf{P}'\mathbf{v}| \leq (\alpha + \beta \mu)^{1/2} (\alpha + \beta \mu')^{1/2} |\mathbf{v}|
\]

**Proof.** Observe that

\[
|\mathbf{P}\mathbf{P}'\mathbf{v}| = \langle \mathbf{u}, \mathbf{P}\mathbf{P}'\mathbf{v} \rangle.
\]

where \( \mathbf{u} \) is a unit vector parallel to \( \mathbf{P}\mathbf{P}'\mathbf{v} \). Put \( \mathbf{w} = \mathbf{P}\mathbf{u}, \mathbf{w}' = \mathbf{P}'\mathbf{v} \), and note that

\[
\langle \mathbf{u}, \mathbf{P}\mathbf{P}'\mathbf{v} \rangle = \langle \mathbf{P}\mathbf{u}, \mathbf{A}\mathbf{P}'\mathbf{v} \rangle = \langle \mathbf{w}, \mathbf{A}\mathbf{w}' \rangle.
\]

Decompose \( \mathbf{w} = \mathbf{x} + \mathbf{y} \) and \( \mathbf{w}' = \mathbf{x}' + \mathbf{y}' \) as in Proposition 4.1 ii). From bilinearity of the inner product,

\[
\langle \mathbf{w}, \mathbf{A}\mathbf{w}' \rangle = \langle \mathbf{x}, \mathbf{A}\mathbf{x}' \rangle + \langle \mathbf{x}, \mathbf{A}\mathbf{y}' \rangle + \langle \mathbf{y}, \mathbf{A}\mathbf{x}' \rangle + \langle \mathbf{y}, \mathbf{A}\mathbf{y}' \rangle.
\]

Proposition 4.1 implies \( \langle \mathbf{x}, \mathbf{A}\mathbf{y}' \rangle = \langle \mathbf{y}, \mathbf{A}\mathbf{x}' \rangle = 0 \), \( \langle \mathbf{x}, \mathbf{A}\mathbf{x}' \rangle = \langle \mathbf{x}, \mathbf{x}' \rangle = |\mathbf{x}| |\mathbf{x}'| \), and \( \langle \mathbf{y}, \mathbf{A}\mathbf{y}' \rangle \leq |\mathbf{y}| |\mathbf{y}'| \leq \alpha |\mathbf{y}| |\mathbf{y}'| \), so

\[
\langle \mathbf{w}, \mathbf{A}\mathbf{w}' \rangle \leq |\mathbf{x}| |\mathbf{x}'| + \alpha |\mathbf{y}| |\mathbf{y}'|.
\]

By the Cauchy-Schwarz inequality

\[
|\mathbf{x}| |\mathbf{x}'| + \alpha |\mathbf{y}| |\mathbf{y}'| \leq (|\mathbf{x}|^2 + |\mathbf{y}|^2)^{1/2} (|\mathbf{x}'|^2 + |\mathbf{y}'|^2)^{1/2} = |\mathbf{w}| |\mathbf{w}'|
\]

so \( |\mathbf{y}| |\mathbf{y}'| \leq |\mathbf{w}| |\mathbf{w}'| - |\mathbf{x}| |\mathbf{x}'| \). Thus,

\[
|\mathbf{x}| |\mathbf{x}'| + \alpha |\mathbf{y}| |\mathbf{y}'| \leq |\mathbf{x}| |\mathbf{x}'| + \alpha (|\mathbf{w}| |\mathbf{w}'| - |\mathbf{x}| |\mathbf{x}'|)
= \alpha |\mathbf{w}| |\mathbf{w}'| + \beta |\mathbf{x}| |\mathbf{x}'|.
\]

Clearly, \( |\mathbf{w}| = |\mathbf{P}\mathbf{u}_0| \leq |\mathbf{u}_0| = 1 \) and \( |\mathbf{w}'| = |\mathbf{P}'\mathbf{v}| \leq |\mathbf{v}| \).

Now \( \langle \mathbf{u}_0, \mathbf{w} \rangle = \langle \mathbf{u}_0, \mathbf{x} \rangle + \langle \mathbf{u}_0, \mathbf{y} \rangle = |\mathbf{u}_0| |\mathbf{x}| = |\mathbf{x}| \). But \( \langle \mathbf{u}_0, \mathbf{w} \rangle = \langle \mathbf{u}_0, \mathbf{P}\mathbf{u}_0 \rangle = \langle \mathbf{P}\mathbf{u}_0, \mathbf{u} \rangle \leq |\mathbf{P}\mathbf{u}_0| |\mathbf{u}| = \mu^{1/2} |\mathbf{u}| \), so \( |\mathbf{x}| \leq \mu^{-1/2} |\mathbf{u}| \).

Also, \( \langle \mathbf{u}_0, \mathbf{w}' \rangle = \langle \mathbf{u}_0, \mathbf{x}' \rangle + \langle \mathbf{u}_0, \mathbf{y}' \rangle = |\mathbf{u}_0| |\mathbf{x}'| = |\mathbf{x}'| \). But \( \langle \mathbf{u}_0, \mathbf{w}' \rangle = \langle \mathbf{u}_0, \mathbf{P}'\mathbf{v} \rangle = \langle \mathbf{P}'\mathbf{u}_0, \mathbf{v} \rangle \leq |\mathbf{P}'\mathbf{u}_0| |\mathbf{v}| = \mu^{1/2} |\mathbf{v}| \), so \( |\mathbf{x}'| \leq \mu^{-1/2} |\mathbf{v}| \). This shows

\[
\alpha |\mathbf{w}| |\mathbf{w}'| + \beta |\mathbf{x}| |\mathbf{x}'| \leq \alpha |\mathbf{v}| + \beta (\mu \mu')^{1/2} |\mathbf{v}|.
\]

Hence, we can conclude

\[
|\mathbf{P}\mathbf{P}'\mathbf{v}| = \langle \mathbf{u}, \mathbf{P}\mathbf{P}'\mathbf{v} \rangle \leq (\alpha + \beta (\mu \mu')^{1/2}) |\mathbf{v}|.
\]
By the Cauchy-Schwarz inequality,
\[ \alpha + \beta (\mu \mu')^{1/2} = \alpha^{1/2} \alpha^{1/2} + (\beta \mu)^{1/2}(\beta \mu')^{1/2} \leq (\alpha + \beta \mu)^{1/2}(\alpha + \beta \mu')^{1/2}, \]
which completes the proof.

Random walks on expander graph construction give rise to \( \beta \)-i.i.d. random objects.

In the simplest case we take an \((N, d, \alpha)\)-expander graph \( G = (V, E) \) with \( \alpha + \beta = 1 \) and an integer \( t \geq 0 \). A \( t \)-walk on \( G \) is a \((t + 1)\)-tuple \( y' = (y_0, y_1, \ldots, y_t) \) of (not necessarily distinct) vertices such that \( \{y_i, y_{i+1}\} \in E \) for \( i < t \). Let \( \Omega \) be the set of all \( t \)-walks on \( G \) and \( \mathbb{P} \) be the uniform probability measure on \( \Omega \). We shall see that the projection functions \( U_i \), defined by \( U_i(y') = y_i \) are \( \beta \)-independent. We could, of course, just take the projection functions defined in the same way on the Cartesian product \( V^{t+1} \), but this would entail a considerable increase in the number of bits needed to represent a sample point. We know that \( |\Omega| = N d^t \) because we generate each \( t \)-walk in a unique way by picking an initial vertex \( y_0 \) from \( V \) and then choosing each successive vertex \( y_{i+1} \) by traversing one of the \( d \) edges incident with \( y_i \). On the other hand, \( |V^{t+1}| = N^{t+1} \). Hence, representing a point in \( \Omega \) requires \([\lg N] + t \lg d \) bits while representing a point in \( V^{t+1} \) require \((t + 1)[\lg N] \) bits.

Now we generalize this scheme. Let \( F \) be a permutation on \( V \). Define \( E' \) to be the composition of \( F \) and the edge relation \( E \); i.e.,
\[ E' = F \circ E = \{(u, F(v)) \mid \{u, v\} \in E\} \]
giving a directed graph \( G' = (V, E') \) (possibly with loops). (Goldreich et al. use \( E' = E \circ F \).) \( G' \) is a \( d \)-regular directed graph in the sense that every vertex has both indegree and outdegree \( d \). Let \( \Omega \) be the set of directed \( t \)-walks \( y' = (y_0, y_1, \ldots, y_t) \), where \( \{y_i, y_{i+1}\} \in E' \) for \( i < t \), \( \mathbb{P} \) be the uniform probability measure on \( \Omega \), and \( U_i(y') = y_i \). Theorem \ref{thm:beta-independence} below shows that the projection functions \( U_i \) are \( \beta \)-independent. Before proving this, we need some preliminary results.

If \( G \) has transition matrix \( A \), then \( G' \) has transition matrix \( A' = AB \), where \( B \) is the permutation matrix for \( F \) (i.e., the \((i, j)\) entry of \( B \) is 1 if \( F(i) = j \) and 0 otherwise).

**Corollary 4.3.** Let \( A \) be the transition matrix for an \((N, d, \alpha)\)-expander graph with \( \alpha + \beta = 1 \) and \( A' = AB \), where \( B \) is a permutation matrix. Also let \( P \) and \( P' \) be projection matrices for \( S \) and \( S' \subseteq \{0, 1, \ldots, N - 1\} \) with \( \mu = |S|/N \) and \( \mu' = |S'|/N \). Then
\[ |v^T P A' P'| \leq |v^T| (\alpha + \beta \mu)^{1/2}(\alpha + \beta \mu')^{1/2}. \]

**Proof.** We have
\[
\begin{align*}
PA'P' &= PABP' \\
&= PA(BP'B^{-1})B \\
&= PAB,
\end{align*}
\]

\footnote{Some sources assert that since \( F \) is a permutation, \( G \) and \( G' \) have the same mixing properties. This is true, but since the adjacency matrix of \( G \) is not Hermitian, this assertion does not follow directly.}
where \( \mathbf{P''} = \mathbf{BP'B}^{-1} \). It is easy to see that \( \mathbf{P''} \) is the projection matrix for \( F^{-1}[S'] \). Now \( |F^{-1}[S']|/N = |S'|/N = \mu' \), so by Lemma 4.2
\[
|\mathbf{PAP''v}| \leq (\alpha + \beta \mu)^{1/2}(\alpha + \beta \mu')^{1/2}|v|.
\]
Thus, since \( \mathbf{B} \) is a permutation matrix,
\[
|\mathbf{vPA'P'}| = |\mathbf{vPAP''B}|
\leq |\mathbf{v}|(\alpha + \beta \mu)^{1/2}(\alpha + \beta \mu')^{1/2}.
\]
\[ \square \]

\( A' \) represents a step of a random walk on \( S' \). When it acts on a row vector representing a probability distribution on \( V \), the result is the succeeding probability distribution along the random walk. \( A' \) may also act on an improper probability distribution (i.e., one whose coordinates are nonnegative and sum to at most 1). Fix \( t \) and let \( S \) be a set of \( t \)-walks. The terminal probability vector \( \mathbf{v} \) of \( S \) is a row vector \( (p_0, p_1, \ldots, p_{N-1}) \) formed by partitioning \( S \) into events \( S_j = S \cap \{U_t = j\} \), for each vertex \( j \in V \) and setting \( p_j = \mathbb{P}(S_j) \). This is an improper probability distribution whose coordinates sum to \( \mathbb{P}(S) \). The action of \( A' \) on \( \mathbf{v} \) results in another improper probability distribution characterized in the following lemma. Here the truncation function \( Y \) is given by
\[
Y(y_0, y_1, \ldots, y_{t-1}) = (y_0, y_1, \ldots, y_{t-1}).
\]

**Lemma 4.4.** Let \( A' \) be a transition matrix for some \( d \)-regular directed graph. Let \( \Omega \) and \( \Psi \) be the the sets of directed \( t \)-walks and \((t-1)\)-walks, and \( \mathbb{P} \) and \( \mathbb{P}' \) be the uniform probability measures on \( \Omega \) and \( \Psi \), respectively. Note that \( Y : \Omega \to \Psi \) is a random object on \((\Omega, \mathbb{P})\). Let \( S' \subseteq \Psi \) and \( S = \{Y \in S'\} \). Then \( \mathbb{P}(S) = \mathbb{P}'(S') \) and \( \mathbf{v} = \mathbf{v}'A' \), where \( \mathbf{v} \) and \( \mathbf{v}' \) are the terminal probability vectors for \( S \) and \( S' \), respectively.

The proof is straightforward.

We come now to the main result of this section giving the construction of \( \beta \)-independent random objects.

**Theorem 4.5.** Let \( S = (V, E), (N, d, \alpha) \) and \( \beta \) be as above. Given \( F \), a permutation on \( V \), form directed graph \( S' = (V, E') \) with \( E' = F \circ E \). Let \( \Omega \) be the set of all directed \( t \)-walks in \( S' \) and \( \mathbb{P} \) be the uniform probability measure on \( \Omega \). Then the projection functions \( U_i \) are \( \beta \)-independent random objects on \((\Omega, \mathbb{P})\).

**Proof.** We need to show that for all \( S_i \subseteq V, i \leq t \),
\[
\mathbb{P}\{\bigwedge_{i \leq t} U_i \in S_i\} \leq \prod_{i \leq t} (\alpha + \beta \mu_i), \tag{9}
\]
where \( \mu_i = \mathbb{P}\{U_i \in S_i\} \).

We claim that it is enough to show this in the special case where \( S_0 = S_t = V \). In this case \( \{U_0 \in S_0\} = \{U_t \in S_t\} = \Omega, \mu_0 = \mu_t = 1, \) and \( \alpha + \beta \mu_0 = \alpha + \beta \mu_t = 1 \). In effect, this eliminates constraints on the initial and terminal vertices of walks. Then (assuming \( t > 2 \)) we may delete the initial and terminal vertices to obtain \( \mathbb{P}' \) for \((t-2)\)-walks rather than \( t \)-walks. That is, letting \( \mathbb{P}' \) be the uniform measure on the set of \((t-2)\)-walks,
\[
\mathbb{P}'\{\bigwedge_{1 \leq i \leq t-1} U_i \in S_i\} = \mathbb{P}\{\bigwedge_{0 \leq i \leq t} U_i \in S_i\} = \prod_{1 \leq i \leq t-1} (\alpha + \beta \mu_i).
\]
Put \( u = n^{-1}(1, 1, \ldots, 1) \), let \( A' \) be the transition matrix for \( G' \), and \( P_i \) be the projection matrices for \( S_i \) for \( i \leq t \).

By induction on \( t \), the terminal probability vector for
\[
\{ \bigwedge_{i \leq t} U_{i} \in S_{i} \}
\]  
(10)
in the space of \( t \)-walks on \( G' \) is
\[
u P_0 A' P_1 \cdots P_{t-1} A' P_t.\]  
(11)

For the base step, \( u \) is the uniform probability distribution and \( u P_0 \) is the terminal probability vector for \( \{ U_0 \in S_0 \} \) in the space of 0-walks since multiplication by the projection matrix \( P_0 \) selects precisely the coordinates at positions \( i \in S_0 \).

Now suppose the terminal probability vector for \( S' = \{ \bigwedge_{i \leq t-1} U_{i} \in S_{i} \} \) (in the space of \( (t-1) \)-walks) is given by
\[
u P_0 A' P_1 \cdots P_{t-2} A' P_{t-1}.\]

Following Lemma 4.4, compute the terminal probability vector of \( \{ Y \in S' \} \), the set of length 1 extensions of walks in \( S' \), by multiplying by \( A' \). Then select the \( t \)-walks with terminal vertices in \( S_t \) by multiplying by \( P_t \). This gives (11) as the terminal probability vector of (10) and completes the induction proof.

We may rewrite (11) as
\[
u (P_0 A' P_1)(P_1 A' P_2) \cdots (P_{t-1} A' P_t)
\]
since \( P_i P_i = P_i \) for \( 1 \leq i \leq t-1 \). By Corollary 4.3, multiplication by \( P_i A' P_{i+1} \) changes the magnitude of a vector by at most a factor of \((\alpha + \beta \mu_i)^{1/2}(\alpha + \beta \mu_{i+1})^{1/2}\), so (11) is bounded in magnitude by
\[
|u| \prod_{i \leq t-1} (\alpha + \beta \mu_i)^{1/2}(\alpha + \beta \mu_{i+1})^{1/2}
\]
which is equal to
\[
|u| (\alpha + \beta \mu_0)^{1/2}(\alpha + \beta \mu_t)^{1/2} \prod_{1 \leq i \leq t-1} (\alpha + \beta \mu_i).
\]

By assumption, \( \alpha + \beta \mu_0 = \alpha + \beta \mu_t = 1 \), so we can further simplify this to
\[
|u| \prod_{i \leq t} (\alpha + \beta \mu_i).
\]

Finally, compute the probability of (10) by summing the coordinates of its terminal probability vector (11); we do this by taking the inner product of (10) with the vector \( v = (1, 1, \ldots, 1)^T \). The result is bounded in magnitude by
\[
|u| |v| \prod_{i \leq t} (\alpha + \beta \mu_i).
\]

But \( |u| = n^{-1/2} \) and \( |v| = n^{1/2} \), so we have established (9).
5 Invertibility and One-Way Functions

In this section we present definitions and notation concerning one-way functions.

For a set $S$, let $I_S$ be the identity function on $S$. Consider functions $F: S \to T$ and $G: T \to S$. We say $G$ is a right inverse of $F$ if $F \circ G = I_T$ and is a left inverse of $F$ if $G \circ F = I_S$.  

It is useful to extend these notions to partial functions. A partial function $F$ from $S$ to $T$ (written $F: S \rightharpoonup T$) is a function that maps a subset of $S$ (the domain of $F$, denoted $\text{dom}(F)$) onto a subset of of $T$ (the range of $F$, denoted $\text{ran}(F)$). If $F: S \rightharpoonup T$ and $G: T \rightharpoonup U$, the composition of $F$ and $G$, written $G \circ F$, is a partial function from $S$ to $U$ that maps $s$ to $u$ if there is a $t$ such that $F(s) = t$ and $G(t) = u$. For partial functions $F: S \rightharpoonup T$ and $G: T \rightharpoonup S$, $G$ is a partial right inverse of $F$ if $F \circ G = I_{\text{ran}(F)}$ and is a partial left inverse of $F$ if $G \circ F = I_{\text{dom}(F)}$.

Like functions, partial functions are injective if and only if they have a (partial) left inverse. Unlike functions, which are surjective if and only if they have a left inverse, partial functions always have a partial right inverse. In cryptography, invertibility almost always refers to existence of an efficiently computable partial right inverse of some kind.

To make precise the notion of an efficiently computable partial right inverse we need probabilistic computation and, in particular, the concept of negligibility. Let $\varphi(n)$ be a proposition concerning the natural numbers $n \in \mathbb{N} = \{0, 1, 2, \ldots\}$. We say that $\varphi(n)$ holds infinitely often, and write $\varphi(n)$ i.o. or $(\exists^\infty n) \varphi(n)$, if

$$\forall m \ (\exists n \geq m) \varphi(n).$$

We say that $\varphi(n)$ holds almost always, and write $\varphi(n)$ a.a. or $(\forall^\infty n) \varphi(n)$, if

$$\exists m \ (\forall n \geq m) \varphi(n).$$

$\exists^\infty$ and $\forall^\infty$ are dual quantifiers: $\lnot (\exists^\infty n) \varphi(n)$ is equivalent to $(\forall^\infty n) \lnot \varphi(n)$ and $\lnot (\forall^\infty n) \varphi(n)$ is equivalent to $(\exists^\infty n) \lnot \varphi(n)$. It is helpful to keep this in mind when negating statements.

A function $p: \mathbb{N} \to \mathbb{R}^+$ is negligible if $p(n) = n^{-\omega(1)}$, i.e.,

$$(\forall c > 0) \ (\forall^\infty n) \ (p(n) \leq n^{-c}).$$

(Some sources use the term superpolynomially small rather than negligible.)

In contrast, $p$ is significant if $p(n) = n^{-O(1)}$, i.e.,

$$(\exists c > 0) \ (\forall^\infty n) \ (p(n) \geq n^{-c}).$$

($\mathbb{R}^+$ is the set of nonnegative reals.) The term negligible is standard in cryptography; the term significant is not.

Write $p(n) \approx q(n)$ if $|p(n) - q(n)|$ is negligible. Write $q(n) \gg p(n)$, or $p(n) \ll q(n)$, if $q(n) - p(n)$ is significant.

---

3 Other common terms for left inverse are retract and retraction. Other common terms for right inverse are coretraction and section.

4 The existence of partial right inverses for partial functions is equivalent to the Axiom of Choice. This is almost the same result as Axiom of Choice equivalent AC 5 in Rubin and Rubin [35], which says that the existence of right inverses for onto functions is equivalent to the Axiom of Choice.
Clearly, \( \approx \) is an equivalence relation on the set of functions from \( \mathbb{N} \) to \( \mathbb{R}^+ \). Also, \( \gg \) is irreflexive and transitive. It is easy to see that \( \approx \) is a congruence with respect to \( \gg \); i.e., if \( p(n) \approx p'(n) \) and \( q(n) \approx q'(n) \), then \( p(n) \gg q(n) \) if and only if \( p'(n) \gg q'(n) \).

Let \( F \) be a function from \( \{0, 1\}^\ast \) to \( \{0, 1\}^\ast \). Define the auxiliary function of \( F \) to be \( F(x) = (1^{|x|}, F(x)) \), where \( 1^{|x|} \) is the unary representation of the input length. Auxiliary functions are convenient when defining weakly and strongly one-way functions because they make public to an adversary attempting to invert \( F \) information about the length of a preimage. It is reasonable to suppose that an adversary would have this information.

We will take a probabilistic approach where the arguments of \( F \) are uniformly distributed random bit strings \( X_n \in \{0, 1\}^n \) and \( Y_n = F(X_n) \). \( Y_n \) may not be uniformly distributed; indeed, it need not have a fixed length for a given \( n \).

We require another modification: a probabilistic adversary. Thus, the adversary attempting to find \( X_n \) such that \( F(X_n) = (1^n, Y_n) \) is a p.p.t. partial function \( G(1^n, Y_n, R_n) \) computable in time polynomial in \( |(1^n, Y_n)| \), where \( R_n \) is a random bit string independent of \( X_n \). We may assume that \( R_n \) is uniformly distributed on \( \{0, 1\}^q(n) \) for some polynomial \( q(n) \).

For each \( n > 0 \) and polynomial \( q(n) \), \( (1^n, Y_n, R_n) \) is a random vector and \( P_{(1^n, Y_n, R_n)} \), denoted for the sake of simplicity as \( P^n \), is an induced probability measure on \( \Psi = \{1\}^\ast \times \{0, 1\}^n \). (Strictly speaking, this notation should specify the particular polynomial \( q(n) \) used.) In the definitions below the function \( I(1^n, y, r) = (1^n, y) \) acts as an identity function when the adversary is the p.p.t. function \( G \).

**Definition.** A polynomial time computable function \( F : \{0, 1\}^\ast \rightarrow \{0, 1\}^\ast \) is a **weakly one-way function** if

\[
(\exists \delta \gg 0) (\forall \text{p.p.t. } G) \left( P^n \{ F \circ G = I \} \leq 1 - \delta \right),
\]

or equivalently, there is a \( c > 0 \) such that for all p.p.t. \( G \), \( P^n \{ F \circ G = I \} \leq 1 - n^{-c} \) a.a.

A polynomial time computable function \( F \) is a **strongly one-way function** if

\[
(\forall \text{p.p.t. } G) \left( P^n \{ F \circ G = I \} \approx 0 \right),
\]

or equivalently, for all \( c > 0 \) and all p.p.t. \( G \), \( P^n \{ F \circ G = I \} \leq n^{-c} \) a.a. If, in addition to either of the conditions above, \( F \) is a length-preserving permutation (i.e., it is a bijection on \( \{0, 1\}^n \) when restricted to strings of length \( n \)), we say it is a **weakly or strongly one-way permutation**.

**Remark.** The notation used in this paper differs from the notation used in other sources. A typical definition (similar to the one found in [13]) says that \( F \) is weakly one-way if there is a \( c > 0 \) such that for every p.p.t. \( G \) and all large \( n \),

\[
P\{G(1^n, F(x)) \notin F^{-1}F(x)\} > n^{-c}
\]

where the probability is taken uniformly over \( x \in \{0, 1\}^n \) and the random bits used by \( G \). This is equivalent to the definition above.

Given \( F \) and \( G \) as above, let

\[
W = P^n \{ F \circ G = I \}.
\]
$W(1^n, y)$ is the probability that $\overline{G}$ successfully finds a length $n$ inverse of $y$. Taking the conditional probability with respect to $I$ averages over the random bit strings used by $\overline{G}$. Using this notation, we may prove a standard amplification result which yields a useful tail bound for $W$ when $F$ is a weakly one-way function \cite{15, 13}.

**Proposition 5.1.** Let $F$ be an auxiliary function and $\overline{G}$ be a randomized partial function with $W = \mathbb{P}^n \{ F \circ \overline{G} = I \}$. 

(i). Let $\overline{G}'$ be the randomized partial function computed by independently computing $\overline{G}$ $k$ times on a given input $(1^n, y)$ (with fresh random bits each time) and returning the first value $x$ such that $F(x) = (1^n, y)$ (if there is such a value). Let $W' = \mathbb{P}^n \{ F \circ \overline{G}' = I \}$. Then for any $0 < \varepsilon < 1$,

$$\mathbb{P}^n_I \{ W > \varepsilon \} = \mathbb{P}^n_I \{ W' > 1 - (1 - \varepsilon)^k \}.$$

(ii). If $F$ is a weakly one-way function where for every p.p.t. $\overline{G}$, $\mathbb{P}^n \{ F \circ \overline{G} = I \} \leq 1 - \delta$, and $\varepsilon = \varepsilon(n)$ and $\delta = \delta(n)$ are significant, then

$$\mathbb{P}^n_I \{ W > \varepsilon \} \ll 1 - \delta/2.$$

*Proof.* Fix a value $(1^n, y)$ in the range of $F$. $W(1^n, y) > \varepsilon$ asserts that the probability that $F \circ \overline{G}(1^n, y) = (1^n, y)$ is greater than $\varepsilon$ or, equivalently, the probability that $F \circ \overline{G}(1^n, y) \neq (1^n, y)$ is less than $1 - \varepsilon$. This happens if and only if the probability that $k$ independent computations of $F \circ \overline{G}(1^n, y)$ fail to yield $(1^n, y)$ is less than $(1 - \varepsilon)^k$; and this happens if and only if $W'(1^n, y) > 1 - (1 - \varepsilon)^k$. This proves part (i).

Now suppose $\varepsilon$ is significant. There is an integer $d$ such that $\varepsilon > n^{-d}$ a.a. In (i) take $k = n^{d+1}$ so almost always

$$\mathbb{P}^n_I \{ W > \varepsilon \} = \mathbb{P}^n_I \{ W' > 1 - (1 - \varepsilon)^k \} \leq \mathbb{P}^n_I \{ W' > 1 - (1 - n^{-d})^{n^{d+1}} \} \leq \mathbb{P}^n_I \{ W' > 1 - e^{-n} \}.$$

By Markov’s inequality and the weakly one-way assumption, we have almost always

$$\mathbb{P}^n_I \{ W' > 1 - e^{-n} \} \leq \mathbb{E}[W']/(1 - e^{-n}) \leq (1 - \delta)/(1 - e^{-n}) = 1 - \delta + e^{-n}(1 - \delta)/(1 - e^{-n}) \leq 1 - \delta + e^{-n} \ll 1 - \delta/2,$$

from which (ii) follows. $\square$

We often deal with partial functions defined only on arguments of certain prescribed lengths. The following technical result shows that under certain circumstances we can obtain weakly and strongly one-way functions (and permutations) from hard-to-invert partial functions by filling in undefined values.
Proposition 5.2. Let $F: \{0,1\}^* \rightarrow \{0,1\}^*$ be polynomial time computable with domain $\bigcup_{m \geq 0} \{0,1\}^{\tau(m)}$, where $\tau(m)$ is strictly increasing, computable in time polynomial in $m$, and for some $k > 0$, $\tau(m+1) \leq \tau(m)^k$. Define $F'(x)$ on strings $x$ of length $n$ as follows. Let $m$ be the largest integer such that $\tau(m) \leq n$, put $x = x'z$ where $|x'| = \tau(m)$, and set $F'(x) = F(x')z$ (the concatenation of $F(x')$ and $z$).

(i). If there is a $c > 0$ such that for all p.p.t. $\overline{G}$, $\mathbb{P}^{\tau(m)}\{\overline{F} \circ \overline{G} = 1\} \leq 1 - \tau(m)^{-c}$ a.a., then $F'$ is a weakly one-way function.

(ii). If for all $c > 0$ and all p.p.t. $\overline{G}$, $\mathbb{P}^{\tau(m)}\{\overline{F} \circ \overline{G} = 1\} \leq \tau(m)^{-c}$ a.a., then $F'$ is strongly one-way.

(iii). If for every $m \geq 0$, $F$ is a permutation (i.e., its restriction to each domain $\{0,1\}^{\tau(m)}$ is a bijection), then $F'$ is also a permutation.

Proof. Part (iii) is obvious.

For parts (i) and (ii) we derive an upper bound for $\mathbb{P}^n\{\overline{F}' \circ \overline{G}' = 1\}$, where $\overline{G}'$ is an arbitrary p.p.t. function.

From $\overline{G}$ construct a p.p.t. function $\overline{G}$ that attempts to invert $\overline{F}$ as follows on input $(1^{\tau(m)}, y)$. It takes successive values of $n$ in the interval $\tau(m) \leq n < \tau(m + 1)$, each time choosing a random bit string $z$ of length $n - \tau(m)$ and applying $\overline{G}$ to $(1^n, yz)$; if the result is of the form $xz$ and $F(x) = y$, it returns the value $x$ and terminates the computation.

In (i) of the proposition, there is a $c > 0$ such that

$$\mathbb{P}^{\tau(m)}\{\overline{F} \circ \overline{G} = 1\} \leq 1 - \tau(m)^{-c} \text{ a.a.}$$

$\overline{G}$ finds an inverse image $x$ (with respect to $\overline{F}$) of $(1^{\tau(m)}, y)$ only if there is an $n$ in the range $\tau(m) \leq n < \tau(m + 1)$ such that $\overline{G}'$ finds an inverse image $xz$ of $(1^n, yz)$. But

$$1 - \tau(m)^{-c} \leq 1 - n^{-c},$$

so

$$\mathbb{P}^n\{\overline{F}' \circ \overline{G}' = 1\} \leq 1 - n^{-c} \text{ a.a.}$$

Therefore, $F'$ is a weakly one-way function.

In part (ii) of the proposition, for all $c > 0$ and all p.p.t. $\overline{G}$,

$$\mathbb{P}^{\tau(m)}\{\overline{F} \circ \overline{G} = 1\} \leq \tau(m)^{-c} \text{ a.a.}$$

Again, $\overline{G}$ finds an inverse $x$ of $(1^{\tau(m)}, y)$ only if there is an $n$ in the range $\tau(m) \leq n < \tau(m + 1)$ such that $\overline{G}'$ finds an inverse $xz$ of $(1^n, yz)$. But $n < \tau(m + 1) \leq \tau(m)^k$ a.a., so

$$\mathbb{P}^n\{\overline{F}' \circ \overline{G}' = 1\} \leq n^{-c/k} \text{ a.a.}$$

and, therefore, $F'$ is a strongly one-way function. \qed

6 From Weakly to Strongly One-Way.

We now show that the existence of a weakly one-way function implies the existence of a strongly one-way function. As noted in the introduction, the first published proof \cite{13} implicitly defines a reduction $(\mathcal{R}, \mathcal{R}^*)$ between function inversion problems. We are in a position to give a more accurate account of $\mathcal{R}$ and $\mathcal{R}^*$. 

17
Recall that from $F$ we define $F'$ by

$$F'(x_0x_1 \cdots x_{t-1}) = (F(x_0), F(x_1), \ldots, F(x_{t-1}))$$

where $t = t(n)$ is a polynomially bounded function and $|x_i| = n$ for all $i$. This defines $F'$ only for inputs $x' = x_0x_1 \cdots x_{t-1}$ of length $nt(n)$, but the Proposition 5.2 allows us to extend $F'$ to a total function. It is convenient to view $R$ as a polynomial time oracle Turing machine that computes the auxiliary function $F'$ by making queries to evaluate $F$ at $x_0, x_1, \ldots, x_{t-1}$. Similarly, $R^*$ is a probabilistic polynomial time oracle Turing machine that computes $G$ by making queries to evaluate the function $F$ and the randomized function $G'$ at various values.

We now show how the probability inequality in Theorem 3.1(i) figures in Goldreich’s proof [13].

**Theorem 6.1.** Suppose $F$ is a weakly one-way function, so that for some integer $c > 0$,

$$(\forall \text{ p.p.t. } G) \left( \Pr^n \{ F \circ G = I \} \leq 1 - n^{-c} \text{ a.a.} \right).$$

Then $F'$, defined as above with $t = n^{c+1}$, is a strongly one-way function.

**Proof.** Take an arbitrary p.p.t. $G'$. By Proposition 5.2(ii) it is enough to show that for all $d > 0$,

$$\Pr^{nt} \{ F' \circ G' = I \} \leq (nt)^{-d} \text{ a.a.}$$

Let $Z = \Pr^{nt} \{ F' \circ G' = I \}$ so $\mathbb{E}[Z] = \Pr^{nt} \{ F' \circ G' = I \}$ by the law of iterated expectation.

$R^*$ computes a randomized function $G$ which attempts to find the inverse of $F$ at $(1^n, y)$ as follows.

1. $R^*$ forms $y = (y_0, y_1, \ldots, y_{t-1})$ by choosing a random $i < t$ and setting $y_i = y$; then generating random $x'_j \in \{0, 1\}^n$ for each $j \neq i$ and putting $y_j = F(x'_j)$. (Here it queries the $F$ oracle).

2. It queries the $G'$ oracle on $(1^{nt}, y)$ and receives an answer $x_0x_1 \cdots x_{t-1}$, where $|x_j| = n$.

3. It checks that $F(x_j) = y_j$ for all $j < t$ (again, by querying the $F$ oracle) and, if so, returns $x_i$; otherwise, the function is undefined.

When this procedure returns a value $x_i$, $(1^n, y) = F(x_i)$.

Define $U_i(1^{nt}, y) = (1^n, y_i)$. The random objects $U_i$ are i.i.d. The probability that $G'$ successfully finds an inverse for $(1^{nt}, y)$ given that $y = y_i$ is $W_i(y)$, where $W_i = \mathbb{E}[Z | U_i]$. Therefore, $\Pr^n \{ F \circ G = I \}$ is $W = t^{-1} \sum W_i$. By Theorem 3.1(i)

$$\mathbb{E}[Z] \leq \Pr^t \{ W > \epsilon \} + t\epsilon.$$

Let $d$ be an arbitrary positive integer and put $\epsilon = n^{-d}t^{-d-1}/2$. By Proposition 5.1 $\Pr^t \{ W > \epsilon \} \leq 1 - n^{-c}/2$ a.a. Thus,

$$\Pr^t \{ W > \epsilon \} \ll (1 - n^{-c}/2)^{n^{c+1}} \leq (e^{-n^{c}/2})^{n^{c+1}} = e^{-n/2} \leq (nt)^{-d}/2.$$  

Also, $\epsilon t = (nt)^{-d}/2$. Consequently, $\Pr^{nt} \{ F' \circ G' = I \} = \mathbb{E}[Z] \leq (nt)^{-d}$ a.a.  

\[ \square \]
Using Proposition 5.2, we have the following corollary.

**Corollary 6.2.**

(i). If weakly one-way functions exist, then strongly one-way functions exist.

(ii). If weakly one-way permutations exist, then strongly one-way permutations exist.

### 7 A Security Preserving Reduction for One-Way Permutations.

In this section we restrict our attention to one-way permutations. Since $|F(x)| = |x|$, it is not necessary to to use an auxiliary function $F(x) = (1^{|x|}, F(x))$ – an adversary can infer this information from $|F(x)|$.

Definitions of *security preserving reduction* differ on details [15, 9, 25, 16, 14], but follow the same general pattern. $R$, with oracle access to cryptographic primitive $F$, computes $F'$. $R^*$, with oracle access to $G'$, computes $G$. The probability of $G$ breaking $F$ depends on the probability of $G'$ breaking $F'$. $(R, R^*)$ is a security preserving reduction if the security of $F'$ against $G'$ is of the same order or dominates the security of $F$ against $G$ (supposing the security function satisfies some reasonable growth condition, for example, being exponentially bounded).

When the cryptographic primitive is a one-way permutation, *breaking* means inverting. In this case let $G$ be a randomized function computed in time $T(n)$ (not necessarily a polynomial) and $\varepsilon(n)$ be the probability that $G$ inverts $F$. We define the *security* of $F$ against $G$ as $S(n) = T(n)/\varepsilon(n)$. This is essentially the expected time to invert $F$ by applying $G$ independently to random elements in the range of $F$. In this context we will say that $S'(n)$ dominates or is of the same order as $S(n)$ (written $S'(n) \succeq S(n)$) if there are positive constants $k$ and $c$ such that $S'(cn) \geq S(n)/n^k$ a.a. $S'(n)$ and $S(n)$ are of the same order (written $S'(n) \asymp S(n)$) if $S'(n) \asymp S(n)$ and $S(n) \asymp S'(n)$.

The reduction used in Theorem 6.1 is not security preserving because $R^*$ makes one query, of length $n^{c+2}$, to the $G'$ oracle. If the security of $F'$ against $G'$ is $S'(n)$, then the security of $F$ against $G$ is of the same order as $S'(n^{c+2})$, which is not of the same order as $S'(n)$ when $S'(n)$ is reasonably fast-growing.

We apply Theorem 5.1(ii) to show that the reduction of Goldreich et al. [15] is security preserving. In particular, $S(n)$ is essentially $S'(n + \omega(\log n))$, which is of the same order as $S'(n)$. This reduction applies just to one-way permutations rather than arbitrary one-way functions. It uses the set of $t$-walks in a hybrid expander-permutation directed graph in place of a direct power and $\beta$-independence in place of independence.

The expander graphs used for this reduction must be from a fully explicit family, defined as follows.

**Definition.** Let $G_m, m \geq 0,$ be a family of $d$-regular graphs where $G_m$ has vertex set $V_m = \{0, 1, \ldots, N_m - 1\}$ with $N_0 < N_1 < N_2 < \cdots$. A *rotation function* $R(N, u, j)$ for this family satisfies the following conditions.

1. It is defined if and only if for some $m$, $N = N_m$, $0 \leq u < N_m$, and $0 \leq j < d$.

2. For each edge $\{u, v\}$ in $G_m$, there is a unique pair $j, k$ such that $R(N_m, u, j) = (v, k)$ and $R(N_m, v, k) = (u, j)$.
A rotation function gives an implicit linear order (not necessarily the lexicographic order) on the \(d\) edges incident with vertex \(u\) in \(\mathbb{G}_m\). Intuitively, \(R(N_m, u, j) = (v, k)\) asserts that the \(j\)-th edge incident with \(u\) is the same as the \(k\)-th edge incident with \(v\).

A family of \(d\)-regular graphs is **fully explicit** if it has a polynomial time computable rotation function.

If \(\mathbb{G}_m, m \geq 0\), has a rotation function \(R(N, u, j)\) such that for every edge \(\{u, v\}\) in \(\mathbb{G}_m\) there is a \(j\) such that \(R(N_m, u, j) = (v, j)\), we let \(\kappa(u, v) = j\). Thus, \(\kappa\) defines an edge coloring. That is, the incident edges at each vertex have distinct colors. When this occurs for a polynomial time computable \(R(N, u, j)\) we have a fully explicit edge coloring.

There is an extensive literature on the construction of fully explicit families of \((N, d, \alpha)\)-expander graphs [29, 12, 23, 34, 1, 4].

**Remark.** Most applications involving fully explicit expander graph families require that the gap between \(N_m\) and \(N_{m+1}\) not grow too quickly. We require more, viz., that \(N_0, N_1, N_2, \ldots\) be a smoothly growing sequence of powers of two with \(N_m = 2^{cn}\) for some constant \(c\), and that \(d\) be a fixed power of two, say \(2^c\). For the remainder of this section we will assume that \(\mathbb{G}_m, m \geq 0\), is a fully explicit \((N, d, \alpha)\)-expander graph family satisfying the above conditions, with \(d\) fixed and \(\alpha < 1\). Hence, \(\mathbb{G}_m = (\{0, 1\}^n, E_m)\) with \(n = cm\). As a notational convenience, we will take \(E = \bigcup_{m \geq 0} E_m\) and write \(\mathbb{G}_m = (\{0, 1\}^n, E)\) rather than \(\mathbb{G}_m = (\{0, 1\}^n, E_m)\).

One example of an explicit family of expander graphs satisfying these conditions is the affine torus expander graph family of Margulis [29]. Gabber and Galil [12] established an upper bound for the second largest eigenvalue magnitude of graphs in this family, later improved by Jimbo and Maruoka [23]. Using these results, we may take \(n = 2m\) (so \(N_m = 2^{2m}\)), \(d = 8 = 2^3\), and \(\alpha = 5\sqrt{2}/8 \approx 0.88388\). Other constructions may give a better bound for \(\alpha\). The argument in [15] requires that \(\alpha \leq 1/2\), but the approach here based on \(\beta\)-independence works for any fixed bound less than 1.

Goldreich et al. [15] require an expander graph family which has a fully explicit edge coloring (but use different terminology). Many explicit expander graph constructions do, in fact, have a fully explicit edge coloring, but we will extend the proof in [15] so that we may dispense with this assumption.

Let \(t = t(n)\) be a polynomially bounded, strictly increasing function. We first describe the transformation \(\mathcal{R}\) taking \(F\), a weakly one-way permutation, to \(F'\), a slightly harder to invert permutation.

Take \(E' = F \circ E\). This gives a family of directed graphs \(\mathbb{G}'_m = (\{0, 1\}^n, E')\). Note that for each directed edge \((u, v)\) in \(\mathbb{G}'_m\), there is a unique \(v\) such that \(\{u, v\} \in E\) and \(F(v) = w\). This suggests two ways to color the directed edges of \(\mathbb{G}'_m\). If \(R(N_m, u, j) = (v, k)\), we have the coloring \(\kappa(u, w) = j\) and the coloring \(\kappa'(u, w) = k\). Thus, \(\kappa\) is an explicit out-edge coloring in the sense that at every vertex \(u\) of \(\mathbb{G}'_m\), the \(d\) out-edges are differently colored; and \(\kappa'\) is an explicit in-edge coloring in the sense that at every vertex \(w\), the \(d\) in-edges are differently colored.

Let \(x = (x_0, x_1, \ldots, x_t)\) be a directed \(t\)-walk in \(\mathbb{G}'_m\). The **forward representation** of \(x\) is

\[
\varphi(x) = (x_0, \kappa(x_0, x_1), \kappa(x_1, x_2), \ldots, \kappa(x_{t-1}, x_t)).
\]

In effect, we regard \(\mathbb{G}'_m\) and its out-edge-coloring as a finite automaton with alphabet \(\{0, 1, \ldots, d-1\}\), and \(\kappa(x_0, x_1)\kappa(x_1, x_2)\cdots\kappa(x_{t-1}, x_t)\) as the unique string causing this finite state machine to reach state \(x_t\).
automaton to transition through the states \( x_0, x_1, \ldots, x_t \). This walk representation uses fewer bits than just listing vertices. Clearly, \( \varphi \) is a bijection from the set of directed \( t \)-walks in \( G \) to \( V \times \{0, 1, \ldots, d-1\}^t \). We will identify \( V \times \{0, 1, \ldots, d-1\}^t \) with \( \{0, 1\}^{n+te} \). Since \( r \) and \( F \) are polynomial time computable, so are \( \varphi \) and \( \varphi^{-1} \).

The reverse representation of \( x \) is

\[
\rho(x) = (x_t, \kappa'(x_{t-1}, x_t), \kappa'(x_{t-2}, x_{t-1}), \ldots, \kappa'(x_0, x_1)).
\]

View this as taking \( G'_m \) together with its in-edge-coloring, reversing the edge directions to form another finite automaton, and specifying a succinct walk representation as before. As with \( \varphi \), \( \rho \) is a bijection from the set of directed \( t \)-walks in \( G \) to \( \{0, 1\}^{n+te} \). It is easy to see that \( \rho \) is polynomial time computable, but since \( F \) is weakly one-way, it does not follow that \( \rho^{-1} \) is polynomial time computable.

We now describe how \( R \) computes \( F' \), a permutation on \( \{0, 1\}^{n+te} \), from \( F \), a permutation on \( \{0, 1\}^n \), where \( n = cm \). For each \( t \)-walk \( x \) in \( G'_m \), \( F' \) maps \( \varphi(x) \) to \( \rho(x) \). In other words, \( F' = \rho \circ \varphi^{-1} \). To compute \( F'(x_0, k_1, \ldots, k_t) \), \( R \) begins at vertex \( x_0 \) in \( G'_m \), \( R \) repeatedly follows the \( k_i \)-th edge from the current vertex, then applies \( F \) to jump to a new vertex. Clearly, \( F' \) is a permutation. By this definition, \( F' \) is defined only on strings of length \( cm + te \) for \( m \geq 0 \), but we may use Proposition 5.2 to extend \( F' \) so that it is defined on strings of any length.

**Lemma 7.1.** From a weakly one-way permutation \( F \), construct \( F' \) as above with polynomially bounded \( t = t(n) \). Suppose \( \delta = \delta(n) \) is significant and that for all p.p.t. \( G \),

\[
P^n\{F \circ G = I\} \leq 1 - \delta \text{ a.a.}
\]

Then the following hold.

(i). For every p.p.t. \( G' \),

\[
P^{n+te}\{F' \circ G' = I\} \leq (1 - \beta \delta(n)/2)^t \text{ a.a.}
\]

(ii). If \( t \geq 7/\beta \), then for every p.p.t. \( G' \),

\[
P^{n+te}\{F' \circ G' = I\} \leq \max(1 - 2\delta(n), 1/2) \text{ a.a.}
\]

(iii). If \( \delta(n) \geq 1/2 \) a.a. and \( t = \omega(\log n) \), then for every p.p.t. \( G' \),

\[
P^n\{F' \circ G' = I\} \approx 0.
\]

**Proof.** (i) Let \( Z = P^{n+te}\{F' \circ G' = I\} \) so \( E[Z] = P^{n+te}\{F' \circ G' = I\} \). As in the proof of Theorem 6.1 we have \( R \) taking \( F \) to \( F' \) and must specify \( R^* \) taking each p.p.t. function \( G' \), which attempts to invert \( F' \), to another p.p.t. function \( G \), which attempts to invert \( F \). On a given input \( y \), \( R^* \), querying oracles for \( G' \) and \( F \), computes \( G \) as follows.

1. \( R^* \) chooses a random \( i \) in the interval \( 1 \leq i < t \), then generates a random sequence of integers \( k_1, k_2, \ldots, k_{t-i} \), where each \( k_j \) is in the range \( 0 \leq k_j < d \); \( (y_1, k_1, k_2, \ldots, k_{t-i}) \) is the forward walk representation of a random \( (t-i) \)-walk in \( G'_m \) with initial vertex \( y \). \( R^* \) applies \( F \) to obtain the reverse walk representation \( (y_t, k'_1, k'_2, \ldots, k'_{t-i}) \) then generates random integers \( k'_{t-i+1}, k'_{t-i+2}, \ldots, k'_{t} \) in the range \( 0 \leq k'_j < d \) to obtain \( y = (y_t, k'_1, k'_2, \ldots, k'_t) \), the reverse walk representation of a \( t \)-walk \( (y_0, y_1, \ldots, y_t) \) chosen randomly from \( t \)-walks such that \( y_i = y \).
2. \( R^* \) queries the \( G' \) oracle on \( y \) and receives an answer
\[
x = (x_0, k'_1, k'_2, \ldots, k'_n).
\]
3. \( R^* \) applies \( \varphi^{-1} \) to \( x \) to obtain a purported walk \((x_0, x_1, \ldots, x_t)\), and then applies \( \rho \). If the result matches \( y \), \( R^* \) has verified that \((x_0, x_1, \ldots, x_t)\) is indeed a \( t \)-walk and that \( x_i = y \) and \((x_{i-1}, x_i)\) is an edge in \( \mathcal{G}_m' \). In this case, there is a \( v \) such that \( \{x_{i-1}, v\} \) is an edge in \( \mathcal{G}_m \) and \( F(v) = x_i \), so \( R^* \) returns \( v = R(2^n, x_{i-1}, k'_n) \); otherwise, the function is undefined.

When this procedure returns a value \( v \), \( F(v) = y \).

Define \( U_i(y) = y_i \), where \( \rho^{-1}(y) = (y_0, y_1, \ldots, y_t) \). By Theorem 4.3 and the bijectivity of \( \rho \), the random objects \( U_i \) are \( \beta \)-i.i.d. The probability that \( G' \) successfully finds an inverse of \( y \), given that \( y = y_i \) is \( W_i(y) \), where \( W_i = \mathbb{E}[Z|U_i] \). Therefore, \( \mathbb{P}^n \{F \circ G = I|I\} \) is \( W = t^{-1} \sum W_i \). By Theorem 5.1(ii),
\[
\mathbb{E}[Z] \leq (\alpha + \beta \mathbb{P}^n \{W > \varepsilon\})^t + t \varepsilon.
\]

By Proposition 5.1 \( \mathbb{P}^n \{W > \varepsilon\} \ll 1 - \delta/2 \) Thus, almost always
\[
(\alpha + \beta \mathbb{P}^n \{W > \varepsilon\})^t \leq (\alpha + \beta(1 - \delta/2))^t = (1 - \beta\delta/2)^t.
\]

Hence, for every significant \( \varepsilon \), \( \mathbb{E}[Z] \ll (1 - \beta\delta/2)^t + t \varepsilon \), which proves (i).

(ii) The function \( f(x) = (1 - \beta x)^t \) is convex for \( x \geq 0 \) and
\[
f((\beta t)^{-1}) = (1 - 1/t)^t < e^{-1}.
\]

so it lies below the continuous piecewise linear function
\[
g(x) = \begin{cases} 
1 - (1 - e^{-1})\beta tx, & \text{if } 0 \leq x \leq (\beta t)^{-1}, \\
1 - e^{-1}, & \text{if } x > (\beta t)^{-1}.
\end{cases}
\]

In other words, \( f(x) \leq g(x) = \max \left(1 - (1 - e^{-1})\beta tx, e^{-1}\right) \) (see Figure 2). Thus, setting \( x = \delta/2 \) gives
\[
(1 - \beta\delta/2)^t \leq \max \left(1 - (1 - e^{-1})\beta t\delta/2, e^{-1}\right).
\]

But \( \beta t \geq 7 \) and \( (1 - e^{-1})/2 = 0.31606 \ldots \), so from part (i) of the theorem
\[
\mathbb{P}^{n+te} \{F' \circ G' = I\} \ll \max(1 - 2\delta, 1/2)
\]

(iii) This also follows from part (i). Since \( \delta \geq 1/2 \) a.a., \( 1 - \beta\delta/2 \leq 1 - \beta/4 \) a.a. We know \( t = \omega(\log n) \) a.a., so \( (1 - \beta\delta/2)^t \) is negligible.

To conclude, we show that the reduction of Goldreich et al. \[15\] from a weakly one-way permutation to a strongly one-way permutation is security preserving.

First, consider the reduction \((R_0, R_0')\) described in the proof of Lemma 7.1(ii) where \( t \geq 1/\beta \) is an even integer. (Recall that \( n = 2m \) and \( e = 3 \), so an even \( t \) ensures that \( n + te \)
is even.) If $F$ is a weakly one-way permutation, where there is a significant $\delta$ such that for all p.p.t. $G$, $\mathbb{P}^n\{ F \circ G = I \} \leq 1 - \delta$ a.a., then applying $R_0$ $s$ times to $F$ results in a permutation $F'$ such that for any p.p.t. $G'$ attempting to invert $F'$

$$\mathbb{P}^{n+\text{rte}}\{ F' \circ G' = I \} \leq \max(1 - 2^s\delta, 1/2) \text{ a.a.}$$

We know that for some $c > 0$, $\delta(n) < n^{-c}$ a.a., so, setting $s = \lceil c \log n \rceil$, we have that for every p.p.t. $G'$, $\mathbb{P}^n\{ F' \circ G' = I \} \leq 1/2$ a.a.

Next, consider the reduction $(R_1, R^*)$ described in the proof of Lemma 7.1(iii) where $t = \omega(n)$. Applying $R_1$ to $F'$ results in a permutation $F''$ such that for all p.p.t. $G''$ attempting to invert $F''$,

$$\mathbb{P}^n\{ F'' \circ G'' = I \} \approx 0.$$ 

Thus, a transformation $R$ consisting of $s$ applications of $R_0$ followed by an application of $R_1$ takes $F$ to $F''$, thereby increasing input length from $n$ to $n + \omega(\log n)$. Transformation $R^*$ consisting of an application of $R_1^*$ followed by $s$ applications of $R_0^*$ takes p.p.t. function $G''$ to p.p.t. function $G$, thereby decreasing input length from $n + \omega(\log n)$ to $n$. $R^*$, computing $G$, queries the oracle for $G''$ just once and the probability that $G$ inverts $F$ is precisely the probability that $G''$ inverts $F''$. $R^*$ runs in polynomial time (assuming constant time to answer a query). Thus, we have the following result.

**Theorem 7.2.** For reduction $(R, R^*)$ described above, if the security of $F''$ against $G''$ is $S''(n)$, then the security of $F$ against $G$ is $S(n) \asymp S''(n + \omega(\log n))$. Thus, $S(n) \asymp S''(n)$. That is, $(R, R^*)$ is a security preserving reduction taking weakly to strongly one-way functions.

**Remark.** The reduction of Theorem 7.2 takes every weakly one-way permutation to a strongly one-way permutation. In contrast, the proof of Theorem 6.1 shows that for every weakly one-way function there is a reduction to a strongly one-way function. It is not apparent from this proof that there is just one reduction that takes every weakly one-way function to a strongly one-way function. Thus, the reduction of Theorem 7.2 is security preserving but is also uniform in this sense.
References

[1] Noga Alon, Oded Schwartz, and Asaf Shapira. An elementary construction of constant-degree expanders. *Combinatorics, Probability and Computing*, 17:319–327, 2008.

[2] Robert B. Ash. *Real Analysis and Probability*. Probability and Mathematical Statistics; a Series of Monographs and Textbooks. Academic Press, New York, 1972.

[3] Mihir Bellare, Russell Impagliazzo, and Moni Naor. Does parallel repetition lower the error in computationally sound protocols? In *Proc. 38th IEEE Symp. on Foundations of Computer Science, FOCS 1997*, pages 374–383. IEEE Computer Society Press, 1997.

[4] Avraham Ben-Aroya and Amnon Ta-Shma. A combinatorial construction of almost-Ramanujan graphs using the zig-zag product. In *Proc. 40th Annual ACM Symp. on Theory of Computing*, pages 325–334. ACM, 2008.

[5] Andre Bogdanov and Alon Rosen. Input locality and hardness amplification. *J. Cryptology*, 26:144–171, 2013.

[6] Ran Canetti, Shai Halevi, and Michael Steiner. Hardness amplification of weakly verifiable puzzles. In Joe Kilian, editor, *2nd Theory of Cryptography Conference - TCC 2005*, volume 3378 of *Lecture Notes in Computer Science*, pages 17–33. Springer, 2005.

[7] Ran Canetti, Madhu Sudan, Luca Trevisan, Salil Vadhan, and Hoeteck Wee. Amplifying collision resistance: A complexity-theoretic treatment. In Alfred Menezes, editor, *Advances in Cryptology - Crypto 2007*, volume 4622 of *Lecture Notes in Computer Science*, pages 264–283. Springer, 2007.

[8] Kai-Min Chung, Feng-Hao Liu, Chi-Jen Lu, and Bo-Yin Yang. Efficient string-commitment from weak bit-commitment. In Masayuki Abe, editor, *Advances in Cryptology - ASIACRYPT 2010*, volume 6477 of *Lecture Notes in Computer Science*, pages 268–282. Springer, 2010.

[9] Giovanni Di Crescenzo and Russell Impagliazzo. Security-preserving hardness-amplification for any regular one-way function. In *Proc. 31st ACM Symp. on Theory of Computing*, pages 169–178, New York, 1999. Association for Computing Machinery.

[10] Yevgeniy Dodis, Russell Impagliazzo, Ragesh Jaiswal, and Valentine Kabanets. Security amplification for interactive cryptographic primitives. In Omer Reingold, editor, *6th Theory of Cryptography Conference, TCC 2009*, volume 5444 of *Lecture Notes in Computer Science*, pages 128–145. Springer, 2009.

[11] Cynthia Dwork, Moni Naor, and Omer Reingold. Immunizing encryption schemes from decryption errors. In Christian Cachin and Jan L. Camenisch, editors, *Advances in Cryptology - EUROCRYPT 2004*, volume 3027 of *Lecture Notes in Computer Science*, pages 342–360. Springer Berlin Heidelberg, 2004.

[12] Ofer Gabber and Zvi Galil. Explicit constructions of linear-sized superconcentrators. *Journal of Computer and System Sciences*, 22:407 – 420, 1981.
[13] Oded Goldreich. *Foundations of Cryptography: Basic Tools*, volume 1. Cambridge, New York, 2001.

[14] Oded Goldreich. On security preserving reductions – revised terminology. In Oded Goldreich, editor, *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation*, pages 540–546. Springer, Berlin, Heidelberg, 2011.

[15] Oded Goldreich, Russell Impagliazzo, Leonid A. Levin, Ramarathnam Venkatesan, and David Zuckerman. Security preserving amplification of hardness. In *Proc. 31st IEEE Symp. on Foundations of Computer Science, FOCS 1990*, pages 318–326. IEEE Computer Society Press, 1990.

[16] Oded Goldreich, Leonid A. Levin, and Noam Nisan. On constructing 1-1 one-way functions. In Oded Goldreich, editor, *Studies in Complexity and Cryptography. Miscellanea on the Interplay between Randomness and Computation*, pages 13–25. Springer, Berlin, Heidelberg, 2011.

[17] Iftach Haitner. A parallel repetition theorem for any interactive argument. *54th Annual IEEE Symposium on Foundations of Computer Science, FOCS 2013*, pages 241–250, 2009.

[18] Iftach Haitner, Danny Harnik, and Omer Reingold. On the power of the randomized iterate. In Cynthia Dwork, editor, *Advances in Cryptology - CRYPTO 2006*, volume 4117 of *Lecture Notes in Computer Science*, pages 22–40. Springer Berlin Heidelberg, 2006.

[19] Shai Halevi and Tal Rabin. Degradation and amplification of computational hardness. In Ran Canetti, editor, *5th Theory of Cryptography Conference, TCC 2008*, volume 4948 of *Lecture Notes in Computer Science*, pages 626–643. Springer, 2008.

[20] Johan Håstad, Rafael Pass, Douglas Wikström, and Krzysztof Pietrzak. An efficient parallel repetition theorem. In Daniele Micciancio, editor, *7th Theory of Cryptography Conference, TCC 2010*, volume 5978 of *Lecture Notes in Computer Science*, pages 1–18. Springer, 2010.

[21] Roger A. Horn and Charles R. Johnson. *Matrix Analysis*. Cambridge Univ. Press, New York, 1985.

[22] Russell Impagliazzo, Ragesh Jaiswal, and Valentine Kabanets. Chernoff-type direct product theorems. *Journal of Cryptology*, 22:75–92, 2009.

[23] Shuji Jimbo and Akira Maruoka. Expanders obtained from affine transformations. *Combinatorica*, 7:343–355, 1987.

[24] Charanjit S. Jutla. Almost optimal bounds for direct product threshold theorem. In Daniele Micciancio, editor, *7th Theory of Cryptography Conference, TCC 2010*, volume 5978 of *Lecture Notes in Computer Science*, pages 37–51. Springer, 2010.

[25] Henry C. Lin, Luca Trevisan, and Hoeteck Wee. On hardness amplification of one-way functions. In Joe Kilian, editor, *Second Theory of Cryptography Conference - TCC 2005*, volume 3378 of *Lecture Notes in Computer Science*, pages 34–49. Springer, 2005.
[26] Michel Loève. *Probability Theory II*. Graduate Texts in Mathematics. Springer, New York, fourth edition, 1978.

[27] Michael Luby. *Pseudorandomness and Cryptographic Applications*. Princeton University Press, Princeton, NJ, USA, 1996.

[28] Michael Luby and Charles Rackoff. Pseudo-random permutation generators and cryptographic composition. In Juris Hartmanis, editor, *18th Annual ACM Symposium on Theory of Computing, STOC 1986*, pages 356–363. ACM, 1986.

[29] G. A. Margulis. Explicit constructions of concentrators. *Probl. Peredachi Inf.*, 9:71–80, 1973.

[30] Ueli M. Maurer and Stefano Tessaro. Computational indistinguishability amplification: Tight product theorems for system composition. In Shai Halevi, editor, *29th Annual International Conference Advances in Cryptology - CRYPTO 2009*, volume 5677 of *Lecture Notes in Computer Science*, pages 355–373. Springer, 2009.

[31] Ueli M. Maurer and Stefano Tessaro. A hardcore lemma for computational indistinguishability: Security amplification for arbitrarily weak PRGs with optimal stretch. In Daniele Micciancio, editor, *7th Theory of Cryptography Conference - TCC 2010*, volume 5978 of *Lecture Notes in Computer Science*, pages 237–254. Springer, 2010.

[32] Moni Naor and Omer Reingold. On the construction of pseudo-random permutations: Luby-Rackoff revisited. *Journal of Cryptology*, 12:29–66, 1997.

[33] Rafael Pass and Muthuramakrishnan Venkitasubramaniam. A parallel repetition theorem for constant-round Arthur-Merlin proofs. *ACM Trans. Comput. Theory*, 4:10:1–10:22, 2012.

[34] Omer Reingold, Salil Vadhan, and Avi Wigderson. Entropy waves, the zig-zag graph product, and new constant-degree expanders. *Ann. of Math. (2)*, 155:157–187, 2002.

[35] Herman Rubin and Jean E. Rubin. *Equivalents of the Axiom of Choice*. Studies in Logic and the Foundations of Mathematics. North-Holland, Amsterdam, 1970.

[36] Eugene Seneta. *Non-negative Matrices and Markov Chains*. Springer Series in Statistics. Springer, New York, second edition, 1981.

[37] Stefano Tessaro. Security amplification for the cascade of arbitrarily weak PRPs: Tight bounds via the interactive hardcore lemma. In Yuval Ishai, editor, *8th Theory of Cryptography Conference - TCC 2011*, volume 6597 of *Lecture Notes in Computer Science*, pages 37–54. Springer, 2011.

[38] Andrew Chi-Chih Yao. Theory and applications of trapdoor functions. In *Proc. 23rd Annual IEEE Symp. on Foundations of Computer Science, FOCS 1982*, pages 80–91. IEEE Computer Society Press, 1982.