Trends in Mathematics
Information Center for Mathematical Sciences
Volume 9, Number 1, June, 2006, Pages 115–126

PARALLEL AND SEQUENTIAL COMPOSITIONS OF PSEUDORANDOM PERMUTATIONS

EONKYUNG LEE

1. Introduction

Pseudorandom functions, introduced by Goldreich, Goldwasser, and Micali [2], are one of the fundamental primitives for cryptographic protocol design. Most importantly, they provide a basis for private-key cryptography. A random function is called a *pseudorandom function (PRF)* if it is indistinguishable in practice from the uniform random function (a.k.a. the truly random function).

Pseudorandom permutations (PRPs) are a special case of PRFs, where the random function takes only permutations as values. There are a lot of block ciphers which have been intensively analyzed and widely trusted today. These block ciphers can be regarded as PRPs.

In order to enhance security of PRPs or to obtain PRFs from PRPs, it is very popular to use sequential and parallel compositions. For PRPs $\pi_1, \ldots, \pi_m$, their parallel composition is defined by

$$(\pi_m \oplus \cdots \oplus \pi_1)(x) = \pi_m(x) \oplus \cdots \oplus \pi_1(x).$$

This operation is also called "XOR". The sequential composition is defined by

$$(\pi_m \circ \cdots \circ \pi_1)(x) = \pi_m\left(\pi_{m-1}\left(\cdots \pi_1(x) \cdots\right)\right).$$

This operation shall be called just "composition" from now on.

The effect of composition and XOR on the security of PRPs have been studied independently. Several works including [4, 1, 7] have been devoted to prove that the composition of PRPs is a more secure PRP. For the XOR, Lucks [5] showed that the XOR of PRPs is a more secure PRF.

This article studies how to combine these two operations in order to get a more secure PRF from a given number of PRPs. A preliminary version is in [3]. Focusing on mathematical analysis more, the present article provides proofs, explanations, and discussions fully.

We describe our results briefly. For random permutations $\pi_1, \ldots, \pi_m$ and for positive integers s and c with $sc = m$, we define a random function

$$\mathrm{SUM}^s - \mathrm{CMP}^c = (\pi_{sc} \circ \cdots \circ \pi_{(s-1)c+1}) \oplus \cdots \oplus (\pi_c \circ \cdots \circ \pi_1).$$

Note that when comparing the securities of two cryptographic schemes, people rely on the bounds of security known most recently. Latest results on the composition and XOR of random permutations are Vaudenay's in 1998 [7] and Luck's in 2000 [5],

respectively. Considering that no better result has appeared in both areas at least for the last six years, we based on their results bound the security of $\text{SUM}^s - \text{CMP}^c$ and then show the following.

(a) If π_i's are insecure, $\pi_m \circ \cdots \circ \pi_1$ is more secure than $\pi_m \oplus \cdots \oplus \pi_1$. If π_i's are secure, $\pi_m \oplus \cdots \oplus \pi_1$ is more secure than $\pi_m \circ \cdots \circ \pi_1$.
(b) In most cases, if m is a composite number and c is its second smallest factor, $\text{SUM}^s - \text{CMP}^c$ is more secure than $\text{SUM}^m - \text{CMP}^1$ regardless of the security of π_i's.
(c) In most cases, the optimal number of compositions for security of $\text{SUM}^s - \text{CMP}^c$ occurs between the second smallest factor of m and the smallest factor of m not less than $\sqrt{m}$ if π_i's are secure.

2. Preliminaries

Let $I_n = \{0,1\}^n$ be the set of all n-bit strings, and let **CN** be the set of all composite, positive integers.

2.1. Random Function.

Definition 1. A *random function* f from I_n to I_m is a random variable which takes as values functions from I_n to I_m. If f takes only permutations with $m = n$, it is called a *random permutation* on I_n.

Definition 2. If a random function (resp. random permutation) has the uniform distribution over all functions from I_n to I_m (resp. over all permutations on I_n), it is called the *uniform random function (URF)* (resp. *uniform random permutation (URP)*) and denoted by $\text{URF}_{n\to m}$ (resp. URP_n). URF_n means $\text{URF}_{n\to n}$.

For a security model for random functions, we consider an adaptive version of the Luby-Rackoff model, in which the number of adversary's queries to an oracle is bounded.

Definition 3. Given two random functions f and f', let an oracle $\mathcal{O}$ simulate either f or f'. A *q-limited distinguisher* for f and f' is a computationally unbounded Turing machine $\mathcal{D}^{\mathcal{O}}$ that outputs either 0 or 1 after a limited number q of interactive queries to $\mathcal{O}$.

The distinguishability between two random functions, f and f', is quantified by the *maximal advantage* over all q-limited distinguishers $\mathcal{D}$ as follows:

$$\mathsf{Adv}^q(f, f') = \max_{\mathcal{D}} \left| \Pr[\mathcal{D}^f = 1] - \Pr[\mathcal{D}^{f'} = 1] \right|.$$

2.2. Decorrelation Theory.

The decorrelation theory is a set of mathematical tools which aims at studying and defining the security of block ciphers in the Luby-Rackoff model. See [9] for details.

Definition 4. Given a random function f from I_n to I_m and an integer d, we define the *d-wise distribution matrix* $[f]^d$ of f as an $I_n^d \times I_m^d$-matrix where the

(x, y)-entry of $[f]^d$ corresponding to the multi-points $x = (x_1, \ldots, x_d) \in I_n^d$ and $y = (y_1, \ldots, y_d) \in I_m^d$ is defined as

$$[f]^d_{x,y} = \Pr[f(x_i) = y_i \text{ for all } 1 \le i \le d].$$

Definition 5. Given two random functions f and f' from I_n to I_m, a positive integer d, and a matrix norm $\|\cdot\|$ over the $I_n^d \times I_m^d$-matrix space $\mathbf{R}^{I_n^d \times I_m^d}$, we define the *$d$-wise decorrelation $\|\cdot\|$-distance* between f and f' as

$$\mathsf{Dec}^d_{\|\cdot\|}(f, f') = \|[f]^d - [f']^d\|.$$

Here, if f' is the URF, the distance is denoted by $\mathsf{DecF}^d_{\|\cdot\|}(f)$ and called *d-wise decorrelation bias of function f*. Similarly, if f is a random permutation and f' is the URP, the distance is denoted by $\mathsf{DecP}^d_{\|\cdot\|}(f)$ and called *d-wise decorrelation bias of permutation f*.

By defining a new matrix norm $\|\cdot\|_a$, Vaudeny linked the decorrelation distance between two random functions to the maximal advantage of distinguisher.

Lemma 1 ([8])**.** *For any random functions f and f', and any positive integer d, we have*

$$\mathsf{Dec}^d_{\|\cdot\|_a}(f, f') = 2 \cdot \mathsf{Adv}^d(f, f').$$

From now on, this paper will use only $\|\cdot\|_a$ as a matrix norm associated with decorrelation distance. Thus, we will simply write Dec^d, DecF^d, and DecP^d instead of $\mathsf{Dec}^d_{\|\cdot\|_a}$, $\mathsf{DecF}^d_{\|\cdot\|_a}$, and $\mathsf{DecP}^d_{\|\cdot\|_a}$, respectively.

3. A Random Function and Its Security

When combining XOR and composition operations, we can think in two ways: XOR-ing after composing and composing after XOR-ing. Both ways produce random functions from random permutations, but we will consider only the former because composing random functions usually diminishes security.

Definition 6. For positive integers c and s, and for i.i.d. random permutations $\pi_1, \ldots, \pi_{sc}$ on I_n, we define a random function $\textsc{sum}^s - \textsc{cmp}^c$ from I_n to I_n as follows:

$$\textsc{sum}^s - \textsc{cmp}^c = (\pi_{sc} \circ \cdots \circ \pi_{(s-1)c+1}) \oplus \cdots \oplus (\pi_c \circ \cdots \circ \pi_1). \tag{1}$$

In order to get the security of $\textsc{sum}^s - \textsc{cmp}^c$, we use the following results of Vaudeny and Lucks.

Lemma 2 ([7])**.** *For any i.i.d. random permutations $\pi_1, \ldots, \pi_c$,*

$$\mathsf{DecP}^d(\pi_c \circ \cdots \circ \pi_1) \le \mathsf{DecP}^d(\pi_1)^c.$$

Lemma 3 ([5])**.** *Let $\pi_1^*, \ldots, \pi_s^*$ be independent URPs on I_n. For any $d \le 2^{n-1}/s$,*

$$\mathsf{DecF}^d(\pi_s^* \oplus \cdots \oplus \pi_1^*) \le \frac{2}{2^{s(n-1)}} \sum_{0 \le i < d} i^s.$$

For more feasible handling, we find a simpler form of alternative to the above boundary formula:

$$\underbrace{\frac{2(d-1)^{s+1}}{(s+1)2^{s(n-1)}}}_{\alpha} \le \underbrace{\frac{2}{2^{s(n-1)}}\sum_{0\le i<d} i^s}_{\delta} \le \underbrace{\frac{2d^{s+1}}{(s+1)2^{s(n-1)}}}_{\beta}.$$

Note that what is important in our analysis is not the value δ itself but the difference between δ's corresponding to distinct s's, that α, β, and δ are all decreasing in s in almost the same shapes, and that α and β only differ in d—the number of queries—by one, and so yield almost the same results in our analysis. So, it hardly affects the result to use β instead of δ as the bound of $\mathsf{DecF}^d(\pi_s^* \oplus \cdots \oplus \pi_1^*)$.

In order to get the security of $\pi_s \oplus \cdots \oplus \pi_1$ when not every π_i is uniform, the following lemma is used. It is obtained by a standard technique (see [8, 6] for its decorrelation version), also mentioned implicitly by Lucks [5].

Lemma 4. *For independent random permutations $\pi_1, \ldots, \pi_s, \pi'_1, \ldots, \pi'_s$ on I_n,*

$$\mathsf{Dec}^d(\pi_s \oplus \ldots \oplus \pi_1, \pi'_s \oplus \ldots \oplus \pi'_1) \le \sum_{i=1}^{s} \mathsf{Dec}^d(\pi_i, \pi'_i).$$

From the above three lemmas, we have the following result.

Theorem 5. *For positive integers c and s, let $\pi_1, \ldots, \pi_{sc}$ be i.i.d. random permutations on I_n. Using them, define $\mathrm{SUM}^s - \mathrm{CMP}^c$ as in (1). Then, for any $d \le 2^{n-1}/s$,*

$$\mathsf{DecF}^d(\mathrm{SUM}^s - \mathrm{CMP}^c) \le s\left(\mathsf{DecP}^d(\pi_1)\right)^c + \frac{2d}{s+1}\left(\frac{d}{2^{n-1}}\right)^s. \tag{2}$$

Proof. Let $\psi_1^*, \ldots, \psi_s^*$ be independent URPs on I_n. For every $1 \le i \le s$, let $\psi_i = \pi_{(i-1)c+c} \circ \cdots \circ \pi_{(i-1)c+1}$. Then

$$\begin{aligned}
\mathsf{DecF}^d(\mathrm{SUM}^s - \mathrm{CMP}^c) &= \|[\psi_s \oplus \cdots \oplus \psi_1]^d - [\mathrm{URF}_n]^d\|_a \\
&\le \|[\oplus_{i=1}^s \psi_i]^d - [\oplus_{i=1}^s \psi_i^*]^d\|_a + \|[\oplus_{i=1}^s \psi_i^*]^d - [\mathrm{URF}_n]^d\|_a \\
&\le s\|[\psi_1]^d - [\mathrm{URP}_n]^d\|_a + \frac{2d}{s+1}\left(\frac{d}{2^{n-1}}\right)^s \\
&\le s\|[\pi_1]^d - [\mathrm{URP}_n]^d\|_a^c + \frac{2d}{s+1}\left(\frac{d}{2^{n-1}}\right)^s \\
&= s\left(\mathsf{DecP}^d(\pi_1)\right)^c + \frac{2d}{s+1}\left(\frac{d}{2^{n-1}}\right)^s.
\end{aligned}$$

□

4. Trade-off between Composition and XOR

Let $m = sc$ be the number of i.i.d. random permutations, and ε the d-wise decorrelation bias of them. Let $\mathsf{UB}-\mathsf{DecF}^d(\mathrm{SUM}^s - \mathrm{CMP}^c)$ denote the upper bound of $\mathsf{DecF}^d(\mathrm{SUM}^s - \mathrm{CMP}^c)$ in (2). Then, it is expressed in terms of $(n, d, \varepsilon, m, c)$ as follows:

$$\mathsf{UB}-\mathsf{DecF}^d(\mathrm{SUM}^s - \mathrm{CMP}^c) = \frac{m}{c}\varepsilon^c + \frac{2cd}{c+m}\left(\frac{d}{2^{n-1}}\right)^{\frac{m}{c}}. \tag{3}$$

As c increases, the first and second terms in the right hand side of (3) behave in the opposite directions: $\frac{m}{c}\varepsilon^c$ decreases to ε^m and $\frac{2cd}{c+m}\left(\frac{d}{2^{n-1}}\right)^{m/c}$ increases to $\frac{d^2}{2^{n-1}}$.

Define a function f of several variables as

$$f(n,d,\varepsilon,m,x)=\frac{m}{x}\varepsilon^x+\frac{2dx}{x+m}\left(\frac{d}{2^{n-1}}\right)^{\frac{m}{x}}.$$

As a bound of $\mathsf{DecF}^d(\text{SUM}^s-\text{CMP}^c)$, f has meaning for integer variables n,d,m,x. However, we regard f as a function of five real variables from now on. All of our results on reals can be directly applied to integers except in some cases (e.g. Theorem 9). In such cases we argue considering that they are integers, especially that x is a factor of the integer m. For (3) to hold, d must be chosen to be at most $\frac{2^{n-1}x}{m}$ as referred to in Theorem 5. Our results for $d\le d_0$ for some d_0's can all be directly applied to this situation by replacing $d\le d_0$ with $d\le \min\left\{\frac{2^{n-1}x}{m},\ d_0\right\}$.

The following lemma shows that f has a nice property in some domain of interest.

Lemma 6. *For any $n\in[4,\infty)$, $d\in[1,2^{n-4}]$, $\varepsilon\in(0,1)$, and $m\in[1,\infty)$, $f(n,d,\varepsilon,m,x)$ is a strictly convex function in $x\in[1,m]$.*

Proof. Fix $n\in[4,\infty)$, $d\in[1,2^{n-4}]$, $\varepsilon\in(0,1)$, and $m\in[1,\infty)$. Let

$$f_1(n,d,\varepsilon,m,x)=x^{-1}\varepsilon^x \quad\text{and}\quad f_2(n,d,\varepsilon,m,x)=x(x+m)^{-1}t^{-x^{-1}},$$

where $t=\left(\frac{2^{n-1}}{d}\right)^m$. Then, $f(n,d,\varepsilon,m,x)=m\cdot f_1(n,d,\varepsilon,m,x)+2d\cdot f_2(n,d,\varepsilon,m,x)$. Since $f(n,d,\varepsilon,m,x)$ is continuous in $x\in[1,m]$, it suffices to show that $\frac{\partial^2 f_1}{\partial x^2}(n,d,\varepsilon,m,x)>0$ and $\frac{\partial^2 f_2}{\partial x^2}(n,d,\varepsilon,m,x)>0$ for all $x\in(1,m)$.

Case of f_1: Clearly, for all $x\in(1,m)$

$$\frac{\partial^2 f_1}{\partial x^2}(n,d,\varepsilon,m,x)=2x^{-3}\varepsilon^x-2x^{-2}\varepsilon^x\ln\varepsilon+x^{-1}\varepsilon^x(\ln\varepsilon)^2>0.$$

Case of f_2: $\frac{\partial^2 f_2}{\partial x^2}(n,d,\varepsilon,m,x)=x^{-3}(x+m)^{-3}t^{-x^{-1}}k(x)$, where

$$k(x)=-2(m+\ln t)x^3+\ln t(-2m+\ln t)x^2+2m(\ln t)^2x+m^2(\ln t)^2.$$

In a moment let's consider $k(x)$ on $\mathbf{R}$. Then, $k'(x)=-6(m+\ln t)(x-x_1)(x-x_2)$, where

$$x_1=\frac{-m\ln t}{m+\ln t}<0 \quad\text{and}\quad x_2=\frac{\ln t}{3}>0.$$

Since $k(0)=m^2(\ln t)^2>0$ and $k(\frac{\ln t}{2})=m(\ln t)^2(m+\frac{\ln t}{4})>0$ for $\frac{\ln t}{2}>0$, we have $k(x)>0$ for all $0\le x\le\frac{\ln t}{2}$. On the other hand, $d\le 2^{n-4}$ implies $m\le\frac{\ln t}{2}$ by the definition of t, so $k(x)>0$ for all $x\in(1,m)$. Therefore, $\frac{\partial^2 f_2}{\partial x^2}(n,d,\varepsilon,m,x)>0$ for all $x\in(1,m)$. □

4.1. **Composition versus XOR.** First, we compare $\textsc{sum}^1 - \textsc{cmp}^m$ and $\textsc{sum}^m - \textsc{cmp}^1$. The following theorem gives $\varepsilon_0 \in (0,1]$, determined by (n,d,m), such that $\mathsf{DecP}^d(\pi_1) < \varepsilon_0$ if and only if $\mathsf{UB}-\mathsf{DecF}^d(\textsc{sum}^m - \textsc{cmp}^1) < \mathsf{UB}-\mathsf{DecF}^d(\textsc{sum}^1 - \textsc{cmp}^m)$.

Theorem 7. *Define α_0 and ε_0 as*

$$\alpha_0(n,d,m) = \tfrac{d^2}{2^{n-1}} - \tfrac{2d}{m+1}\left(\tfrac{d}{2^{n-1}}\right)^m;$$

$$\varepsilon_0(n,d,m) = \begin{cases} 1 & \text{if } \alpha_0(n,d,m) \geq m-1, \\ \text{Root of } x^m - mx + \alpha_0(n,d,m) = 0 \text{ in } (0,1) & \text{otherwise.} \end{cases}$$

For any $n \in [4,\infty)$, $d \in [1, 2^{n-4}]$, and $m \in [1,n]$,

$$\begin{cases} f(n,d,\varepsilon,m,1) < f(n,d,\varepsilon,m,m) & \text{for all} \quad 0 < \varepsilon < \varepsilon_0(n,d,m), \\ f(n,d,\varepsilon,m,1) = f(n,d,\varepsilon,m,m) & \text{for} \quad \varepsilon = \varepsilon_0(n,d,m), \\ f(n,d,\varepsilon,m,1) > f(n,d,\varepsilon,m,m) & \text{for all} \quad \varepsilon_0(n,d,m) < \varepsilon < 1. \end{cases}$$

Proof. Fix $n \in [4,\infty)$, $d \in [1, 2^{n-4}]$, and $m \in [1,n]$. Put $\alpha_0 = \alpha_0(n,d,m)$ and $\varepsilon_0 = \varepsilon_0(n,d,m)$. Define a function k as

$$k(\varepsilon) = f(n,d,\varepsilon,m,m) - f(n,d,\varepsilon,m,1) = \varepsilon^m - m\varepsilon + \alpha_0.$$

In a moment let's consider $k(x)$ on $[0,1]$. Then, $k(x)$ is a decreasing function on $[0,1]$, $k(0) = \alpha_0 > 0$, and $k(1) = 1 - m + \alpha_0$.

Case 1. $1 - m + \alpha_0 \geq 0$: $k(\varepsilon) > 0$ for all $\varepsilon \in (0,1) = (0, \varepsilon_0)$.

Case 2. $1 - m + \alpha_0 < 0$: There exists a unique $\varepsilon_0' \in (0,1)$ such that $k(\varepsilon_0') = 0$. Then,

$$\begin{cases} k(\varepsilon) > 0 & \text{for all} \quad \varepsilon \in (0, \varepsilon_0') = (0, \varepsilon_0), \\ k(\varepsilon) = 0 & \text{for} \quad \varepsilon = \varepsilon_0' = \varepsilon_0, \\ k(\varepsilon) < 0 & \text{for all} \quad \varepsilon \in (\varepsilon_0', 1) = (\varepsilon_0, 1). \end{cases}$$

From Cases 1 and 2, the conclusion follows. □

It is clear that $x^m - mx + \alpha_0 = 0$ has exactly one root in $(0,1)$ whenever $\alpha_0 < m-1$. If $d \geq \sqrt{(m+1)2^{n-1}}$, then $\alpha_0 \geq m-1$. If $d \leq 2^{n/2}$ and $m \geq 3$, then $\alpha_0 < m-1$. For example, $\varepsilon_0(128, 2^{40}, 40) \approx 2^{-52.4}$ and $\varepsilon_0(128, 2^{60}, 60) \approx 2^{-12.9}$.

When m is a prime number, the only comparable forms are $\textsc{sum}^1 - \textsc{cmp}^m$ and $\textsc{sum}^m - \textsc{cmp}^1$. From now on, we focus on composite numbers m. For any of such m's, there exists at least one factor (other than the trivial factor 1) of m not greater than $\sqrt{m}$. The following theorem shows that, in most cases of (n,d,m), $\mathsf{UB}-\mathsf{DecF}^d(\textsc{sum}^s - \textsc{cmp}^c) < \mathsf{UB}-\mathsf{DecF}^d(\textsc{sum}^m - \textsc{cmp}^1)$ for all factors c of m such that $1 < c \leq \sqrt{m}$ regardless of the value of $\mathsf{DecP}^d(\pi_1)$.

Theorem 8. *For any $n \in [16,\infty)$, $d \in [1, 2^{n/2}]$ (resp. $d \in (2^{n/2}, 2^{3n/4}]$), $\varepsilon \in [2^{-n}, 2^{-2}]$, and $m \in [9,n]$ (resp. $m \in [49,n]$),*

$$f(n,d,\varepsilon,m,x) < f(n,d,\varepsilon,m,1) \quad \text{for all} \quad x \in (1, \sqrt{m}]. \tag{4}$$

Proof. For $n \in [16,\infty)$, $d \in [1, 2^{n-4}]$, $\varepsilon \in [2^{-n}, 2^{-2}]$, and $m \in [4, n]$, define a function g as

$$\begin{aligned} g(n,d,\varepsilon,m) &= f(n,d,\varepsilon,m,1) - f(n,d,\varepsilon,m,\sqrt{m}) \\ &= m\varepsilon + \frac{2d}{m+1}\left(\frac{d}{2^{n-1}}\right)^m - \sqrt{m}\varepsilon^{\sqrt{m}} - \frac{2d}{\sqrt{m}+1}\left(\frac{d}{2^{n-1}}\right)^{\sqrt{m}}. \end{aligned}$$

We will show that $g(n,d,\varepsilon,m) > 0$ for all $n \in [16,\infty)$, $d \in [1, 2^{n/2}]$ (resp. $d \in (2^{n/2}, 2^{3n/4}]$), $\varepsilon \in [2^{-n}, 2^{-2}]$, and $m \in [9, n]$ (resp. $m \in [49, n]$). Then, to combine this with Lemma 6 gives the desired results. In order to see how g behaves, we check its slope for d, ε, m.

Claim. For all $n \in [16,\infty)$, $d \in [1, 2^{n-4}]$, $\varepsilon \in [2^{-n}, 2^{-2}]$, and $m \in [4, n]$, the following hold:

$$\text{(a)}\ \frac{\partial g}{\partial d}(n,d,\varepsilon,m) < 0; \quad \text{(b)}\ \frac{\partial g}{\partial \varepsilon}(n,d,\varepsilon,m) > 0; \quad \text{(c)}\ \frac{\partial g}{\partial m}(n,d,\varepsilon,m) > 0.$$

Proof of Claim: Since (a) and (b) are easy, we prove only (c). By straightforward calculation,

$$\begin{aligned} \frac{\partial g}{\partial m}(n,d,\varepsilon,m) &= \varepsilon - \frac{\varepsilon^{\sqrt{m}}}{2\sqrt{m}}\left(1 + \sqrt{m}\ln\varepsilon\right) \\ &\quad + \frac{2d}{m+1}\left(\frac{d}{2^{n-1}}\right)^m \ln\frac{d}{2^{n-1}} - \frac{d}{\sqrt{m}(\sqrt{m}+1)}\left(\frac{d}{2^{n-1}}\right)^{\sqrt{m}} \ln\frac{d}{2^{n-1}} \\ &\quad + \frac{d}{\sqrt{m}(\sqrt{m}+1)^2}\left(\frac{d}{2^{n-1}}\right)^{\sqrt{m}} - \frac{2d}{(m+1)^2}\left(\frac{d}{2^{n-1}}\right)^m. \end{aligned}$$

Since $1 + \sqrt{m}\ln\varepsilon \le -1$ for all $\varepsilon \le 2^{-2}$, we get

$$\varepsilon - \frac{\varepsilon^{\sqrt{m}}}{2\sqrt{m}}\left(1 + \sqrt{m}\ln\varepsilon\right) > 0. \tag{5}$$

Let $t = d/2^{n-1}$. Since $\frac{m+\sqrt{m}}{m+1} \le 2$ and $2^{-1}t^{\sqrt{m}-m} \ge 2^{2(n-1)-1} > 2$ for all $n \ge 16$, we get $\frac{t^{\sqrt{m}-m}}{2} > \frac{m+\sqrt{m}}{m+1}$, which is equivalent to

$$\frac{2dt^m}{(m+1)}\ln t - \frac{dt^{\sqrt{m}}}{\sqrt{m}(\sqrt{m}+1)}\ln t > 0. \tag{6}$$

On the other hand, since $\frac{t^{\sqrt{m}-m}}{2} > 2$ and $\frac{\sqrt{m}(\sqrt{m}+1)^2}{(m+1)^2} = \frac{m+\sqrt{m}}{m+1}\frac{\sqrt{m}+1}{m+1} \le 2$, we get $\frac{t^{\sqrt{m}-m}}{2} > \frac{\sqrt{m}(\sqrt{m}+1)^2}{(m+1)^2}$, which is equivalent to

$$\frac{dt^{\sqrt{m}}}{\sqrt{m}(\sqrt{m}+1)^2} - \frac{2dt^m}{(m+1)^2} > 0. \tag{7}$$

The inequality (c) follows from (5), (6), and (7). *End of proof of Claim.*

Since $1 < 2^{n/2} < 2^{3n/4} \le 2^{n-4}$ for all $n \ge 16$, the inequality (4) is obtained by the following due to (a), (b), and (c):

$$g(n, 2^{n/2}, 2^{-n}, 9) > 0 \quad \text{and} \quad g(n, 2^{3n/4}, 2^{-n}, 49) > 0 \quad \text{for all } n \ge 16.$$

$g(n, 2^{n/2}, 2^{-n}, 9) = 2^{-4n}\left(5 \cdot 2^{3n} - 3 \cdot 2^n + \frac{2^9}{5}\right)$, and $5x^3 - 3x + \frac{2^9}{5} > 0$ for all $x \geq 0$. Hence, $g(n, 2^{n/2}, 2^{-n}, 9) > 0$ for all $n \geq 16$.

$g(n, 2^{3n/4}, 2^{-n}, 49) = 2^{-23n/2}\left(17 \cdot 2^{21n/2} - 7 \cdot 2^{9n/2} + \frac{2^{49}}{25}\right)$, and $17x^{21} - 7x^9 + \frac{2^{49}}{25} > 0$ for all $x \geq 0$. Hence, $g(n, 2^{3n/4}, 2^{-n}, 49) > 0$ for all $n \geq 16$. □

Remark. In the above theorem, $m \geq 9$ for $d \leq 2^{n/2}$ and $m \geq 49$ for $d \leq 2^{3n/4}$ were drawn from the following observation.

$$g(n, 2^{kn}, 2^{-n}, m) \quad = \quad 2^{-mn}\left(m \cdot 2^{a_1 n} + \frac{2^{a_2 n + b_2}}{m+1} - \frac{2^{a_3 n + b_3}}{\sqrt{m}+1} - \sqrt{m}2^{a_4 n}\right),$$

where $a_1 = m-1$, $a_2 = k(m+1)$, $b_2 = m+1$, $a_3 = m - \sqrt{m} + k\sqrt{m} + k$, $b_3 = \sqrt{m}+1$, and $a_4 = m - \sqrt{m}$. Since $a_1 > a_2$ and $a_3 > a_4$ for $k = \frac{1}{2}$ and $\frac{3}{4}$, a necessary (but not sufficient) condition for $g(n, 2^{kn}, 2^{-n}, m) > 0$ for all n is $a_1 \geq a_3$, which holds if and only if $\sqrt{m} \geq \frac{1+k}{1-k}$.

Theorem 8 does not always hold for other m's, i.e. $4 \leq m \leq 8$ for $1 \leq d \leq 2^{n/2}$ nor $4 \leq m \leq 48$ for $2^{n/2} < d \leq 2^{3n/4}$. For these m's, only XOR-ing is sometimes better than accompanying composition. For example, $\mathsf{UB}-\mathsf{DecF}^d(\text{SUM}^4-\text{CMP}^1) < \mathsf{UB}-\mathsf{DecF}^d(\text{SUM}^2-\text{CMP}^2)$ when $(n, d, \varepsilon) = (128, 2^{60}, 2^{-77})$, and $\mathsf{UB}-\mathsf{DecF}^d(\text{SUM}^8-\text{CMP}^1) < \mathsf{UB}-\mathsf{DecF}^d(\text{SUM}^4-\text{CMP}^2)$ when $(n, d, \varepsilon) = (128, 2^{90}, 2^{-65})$.

4.2. **Optimal Number of Compositions.** Theorem 8 says that composition helps XOR to lower $\mathsf{UB}-\mathsf{DecF}^d(\text{SUM}^s - \text{CMP}^c)$ either when $1 \leq d \leq 2^{n/2}$ and $9 \leq m \leq n$ or when $2^{n/2} < d \leq 2^{3n/4}$ and $49 \leq m \leq n$. Then, what is the number of compositions to obtain the minimum value for these (d, m)'s? This number occurs at every factor of m between the second smallest one and m. For example, the minimum occurs at $c = 2$ when $(n, d, \varepsilon, m) = (128, 2^{63}, 2^{-100}, 12)$, and at $c = 12$ when $(n, d, \varepsilon, m) = (128, 2^{15}, 2^{-15}, 12)$. This section analyzes concretely how the optimal number of compositions is related to (n, d, ε, m), from which the optimal number of XORs follows immediately due to $m = sc$.

Notation. For a positive integer m, let $\text{FAC}(m)$ denote the set of all factors of m, and let m_2 be the second smallest factor of m, m_ℓ the greatest factor of m not greater than $\sqrt{m}$, and m_u the smallest factor of m not less than $\sqrt{m}$. Namely, $m_2 = \min(1, m] \cap \text{FAC}(m)$, $m_\ell = \max[1, \sqrt{m}] \cap \text{FAC}(m)$, and $m_u = \min[\sqrt{m}, m] \cap \text{FAC}(m)$.

Given (n, d, ε, m), the minimum of $f(n, d, \varepsilon, m, x)$ occurs at a single point $x \in [1, m]$ because of Lemma 6, but can occur at more than one point $x \in \text{FAC}(m)$. Thus, we define C_0 as the set of all factors of m where f has the minimum:

$$C_0 = \{c_0 \in \text{FAC}(m) \mid f(n, d, \varepsilon, m, c_0) \leq f(n, d, \varepsilon, m, c) \text{ for all } c \in \text{FAC}(m)\}.$$

C_0 is determined by (n, d, ε, m), and the number of its elements is either one or two. The following theorem finds the value, ε_1, of $\mathsf{DecP}^d(\pi_1)$ which is used to determine whether C_0 is inside $[1, m_u]$ or inside $[m_\ell, m]$.

Theorem 9. *Define* α_1 *and* ε_1 *as*

$$\alpha_1(n,d,m) = \tfrac{2d}{(\sqrt{m}+1)^2}\left(\tfrac{d}{2^{n-1}}\right)^{\sqrt{m}}\left(1+(\sqrt{m}+1)\ln\tfrac{2^{n-1}}{d}\right);$$

$$\varepsilon_1(n,d,m) = \begin{cases} 1 & \text{if } \alpha_1(n,d,m) \geq 1, \\ \text{Root of } x^{\sqrt{m}}(\sqrt{m}\ln x - 1) + \alpha_1(n,d,m) = 0 \text{ in } (0,1) & \text{otherwise.} \end{cases}$$

For any $n \in [4,\infty)$, $d \in [1, 2^{n-4}]$, *and* $m \in [1,n] \cap \mathbf{CN}$, *we have*

$$C_0 \subset \begin{cases} [1, m_u] & \text{for all} \quad 0 < \varepsilon \leq \varepsilon_1(n,d,m), \\ [m_\ell, m] & \text{for all} \quad \varepsilon_1(n,d,m) < \varepsilon < 1. \end{cases}$$

Proof. Fix $n \in [4,\infty)$, $d \in [1, 2^{n-4}]$, and $m \in [1,n] \cap \mathbf{CN}$. Put $\alpha_1 = \alpha_1(n,d,m)$ and $\varepsilon_1 = \varepsilon_1(n,d,m)$. Define a function $k(\varepsilon)$ on $(0,1)$ as

$$k(\varepsilon) = \frac{\partial f}{\partial x}(n,d,\varepsilon,m,\sqrt{m}) = \varepsilon^{\sqrt{m}}(\sqrt{m}\ln\varepsilon - 1) + \alpha_1.$$

Then, $k(\varepsilon)$ is decreasing on $(0,1)$, $\lim_{\varepsilon\to 0} k(\varepsilon) = \alpha_1 > 0$, and $\lim_{\varepsilon\to 1} k(\varepsilon) = \alpha_1 - 1$.

If $\alpha_1 - 1 \geq 0$, then $k(\varepsilon) > 0$ for all $\varepsilon \in (0,1) = (0,\varepsilon_1)$. Otherwise, there exists uniquely $\varepsilon_1' \in (0,1)$ such that

$$\begin{cases} k(\varepsilon) > 0 & \text{for all} \quad \varepsilon \in (0,\varepsilon_1') = (0,\varepsilon_1), \\ k(\varepsilon) = 0 & \text{for} \quad\;\; \varepsilon = \varepsilon_1' = \varepsilon_1, \\ k(\varepsilon) < 0 & \text{for all} \quad \varepsilon \in (\varepsilon_1', 1) = (\varepsilon_1, 1). \end{cases}$$

Therefore, $\frac{\partial f}{\partial x}(n,d,\varepsilon,m,\sqrt{m}) \geq 0$ if and only if $0 < \varepsilon \leq \varepsilon_1$.

Case 1. $0 < \varepsilon \leq \varepsilon_1$

Since $f(n,d,\varepsilon,m,x)$ is strictly convex in x and $\frac{\partial f}{\partial x}(n,d,\varepsilon,m,\sqrt{m}) \geq 0$, we have

$$f(n,d,\varepsilon,m,m_u) < f(n,d,\varepsilon,m,c) \quad \text{for all} \quad c \in (m_u, m] \cap \mathrm{FAC}(m).$$

Therefore, $C_0 \subset [1, m_u]$.

Case 2. $\varepsilon_1 < \varepsilon < 1$

Since $f(n,d,\varepsilon,m,x)$ is strictly convex in x and $\frac{\partial f}{\partial x}(n,d,\varepsilon,m,\sqrt{m}) < 0$, we have

$$f(n,d,\varepsilon,m,m_\ell) < f(n,d,\varepsilon,m,c) \quad \text{for all} \quad c \in [1, m_\ell) \cap \mathrm{FAC}(m).$$

Therefore, $C_0 \subset [m_\ell, m]$. □

For example, $\varepsilon_1(128, 2^{32}, 32) \approx 2^{-90.1}$, $\varepsilon_1(128, 2^{64}, 64) \approx 2^{-55.7}$, and $\varepsilon_1(128, 2^{96}, 96) \approx 2^{-21.8}$.

Note that C_0 is composed of a single element, say c_0, in general. Recall $f(n,d,\varepsilon,m,x) = \frac{m}{x}\varepsilon^x + \frac{2dx}{x+m}\left(\frac{d}{2^{n-1}}\right)^{\frac{m}{x}}$. Let $x_0 \in [1,m]$ be the point where $f(n,d,\varepsilon,m,\cdot)$ has the minimum. The value of $f(n,d,\varepsilon,m,x)$ at $x \in [1,x_0]$ (resp. at $x \in [x_0,m]$) depends mainly on $\frac{m}{x}\varepsilon^x$ (resp. on $\frac{2dx}{x+m}\left(\frac{d}{2^{n-1}}\right)^{\frac{m}{x}}$). At every $x \in [1,m]$, $\frac{m}{x}\varepsilon^x$ is an increasing function in (ε, m), and $\frac{2dx}{x+m}\left(\frac{d}{2^{n-1}}\right)^{\frac{m}{x}}$ is an increasing function in d and a decreasing function in m. Therefore, c_0 tends to increase as d decreases, and m and ε increase, and to decrease as d increases, and m and ε decrease.

Consider the case where we are given random permutations. In this case, $\varepsilon = \mathsf{DecP}^d(\pi_1)$ is an increasing function in d. This implies that c_0 should be observed when both d and ε move in the same direction. Therefore, we combine Theorem 8 with Theorem 9 for relatively small d's and ε's in the following corollary: in most cases, the optimal number of compositions occurs between m_2 and m_u for π_i's with $\mathsf{DecP}^d(\pi_i) \le \varepsilon_2$. Here, ε_2 is easier to compute than ε_1.

Corollary 10. *Define* $\varepsilon_2(n,d,m) = \min\left\{2^{-2},\ \frac{d}{2^{n-1}}\left(\frac{2d\left(1+(\sqrt{m}+1)\ln\frac{2^{n-1}}{d}\right)}{(n\sqrt{m}+1)(\sqrt{m}+1)^2}\right)^{\frac{1}{\sqrt{m}}}\right\}$. *For any* $n \in [16,\infty)$, $d \in [2, 2^{n/2}]$ *(resp.* $d \in (2^{n/2}, 2^{3n/4}]$*),* $m \in [9,n] \cap \mathbf{CN}$ *(resp.* $m \in [49,n] \cap \mathbf{CN}$*), and* $\varepsilon \in [2^{-n}, \varepsilon_2(n,d,m)]$*, we have* $C_0 \subset \left[m_2,\ m_u\right]$.

Proof. Recall α_1 and ε_1 from Theorem 9. We will show that $2^{-n} \le \varepsilon_2(n,d,m) \le \varepsilon_1(n,d,m)$ holds for all (n,d,m). Then, the conclusion follows from Theorems 8 and 9.

Fix $n \in [16,\infty)$ and $m \in [9,n]$. Note that

$$\frac{d}{2^{n-1}}\left(\frac{2d\left(1+(\sqrt{m}+1)\ln\frac{2^{n-1}}{d}\right)}{(n\sqrt{m}+1)(\sqrt{m}+1)^2}\right)^{\frac{1}{\sqrt{m}}} \ge 2^{-n} \quad \text{for all} \quad d \in [2, 2^{3n/4}]$$

$$\iff \quad \frac{2}{2^{n-1}}\left(\frac{4\left(1+(\sqrt{m}+1)\ln\frac{2^{n-1}}{2}\right)}{(n\sqrt{m}+1)(\sqrt{m}+1)^2}\right)^{\frac{1}{\sqrt{m}}} \ge 2^{-n}$$

$$\impliedby \quad 2^{2\sqrt{m}+2}\sqrt{m} + (\sqrt{m}+1)^2 \le (\sqrt{m}+1)\left(2^{2\sqrt{m}+1} - \sqrt{m}(\sqrt{m}+1)\right)m.$$

The last inequality above holds because for all $x \ge 2$

$$(x+1)\left(2^{2x+1} - x(x+1)\right)x^2 - \left(2^{2x+2}x + (x+1)^2\right) = 2^{2x+1}x(x+2)(x-1) - (x^3+1)(x+1)^2 > 0.$$

Therefore, we have $\varepsilon_2(n,d,m) \ge 2^{-n}$ for all $d \in [2, 2^{3n/4}]$.

Fix $d \in [2, 2^{3n/4}]$. Put $\alpha_1 = \alpha_1(n,d,m)$, $\varepsilon_1 = \varepsilon_1(n,d,m)$, and $\varepsilon_2 = \varepsilon_2(n,d,m)$. Choose $\varepsilon \in [2^{-n}, \varepsilon_2]$. We will show that $\varepsilon \in [2^{-n}, \varepsilon_1]$. Since $\varepsilon \ge 2^{-n}$, we have

$$\frac{\partial f}{\partial x}(n,d,\varepsilon,m,\sqrt{m}) = -\varepsilon^{\sqrt{m}}(1 - \sqrt{m}\ln\varepsilon) + \alpha_1 \ge -(n\sqrt{m}+1)\varepsilon^{\sqrt{m}} + \alpha_1.$$

Note that $-(n\sqrt{m}+1)x^{\sqrt{m}} + \alpha_1 \ge 0$ for all $0 \le x \le \left(\frac{\alpha_1}{n\sqrt{m}+1}\right)^{\frac{1}{\sqrt{m}}}$, and that $2^{-n} \le \varepsilon \le \varepsilon_2 = \min\left\{2^{-2}, \left(\frac{\alpha_1}{n\sqrt{m}+1}\right)^{\frac{1}{\sqrt{m}}}\right\}$. Hence, $\frac{\partial f}{\partial x}(n,d,\varepsilon,m,\sqrt{m}) \ge 0$. Since $\frac{\partial f}{\partial x}(n,d,\omega,m,\sqrt{m})$ is a decreasing function in $\omega \in (0,1)$, $\varepsilon \in [2^{-n}, \varepsilon_1]$ holds by the definition of ε_1, from which $\varepsilon_2 \le \varepsilon_1$ follows. □

For example, $\varepsilon_2(128, 2^{32}, 32) \approx 2^{-90.3}$, $\varepsilon_2(128, 2^{64}, 64) \approx 2^{-55.8}$, and $\varepsilon_2(128, 2^{96}, 96) \approx 2^{-22.1}$.

Tightness of the range of C_0. In Corollary 10, both the upper and lower bounds of C_0 are very tight. As seen below Theorem 9, we should select (n,d,ε,m) among large d's, and small m's and ε's for $m_2 \in C_0$. Actually, $C_0 = \{2\}$ for

$(n,d,\varepsilon,m) = (128, 2^{63}, 2^{-100}, 12)$ and for $(n,d,\varepsilon,m) = (128, 2^{63}, 2^{-99}, 16)$, from which the tightness of the lower bound follows. For $m_u \in C_0$, we should select (n,d,ε,m) among small d's, and large m's and ε's. Actually, $C_0 = \{16\}$ for $(n,d,\varepsilon,m) = (128, 29, \varepsilon_2(128,29,128), 128) \approx (128, 29, 2^{-122}, 128)$, and $C_0 = \{12\}$ for $(n,d,\varepsilon,m) = (128, 1, \varepsilon_2(128,1,120), 120) \approx (128, 1, 2^{-128}, 120)$, from which the tightness of the upper bound follows.

4.3. **Discussion of the Security Bound of** $\text{SUM}^s - \text{CMP}^c$**.** The upper bound of $\mathsf{DecF}^d(\text{SUM}^s - \text{CMP}^c)$ may be updated to be lowered in future. An ideal way to obtain its tight bound is probably to compute it directly from $\mathsf{DecP}^d(\pi_1)$ without any intermediate step (i.e. without Lemmas 2, 3, 4), which seems quite difficult.

Another approach is to tighten the bounds in these lemmas. In particular, the following bound from Lemma 4 looks loose since it is a kind of corollary of some general theorem (in [8, 6]) unlike the other two lemmas.

$$\delta \stackrel{def}{=} \mathsf{Dec}^d(\psi_s \oplus \cdots \oplus \psi_1, \psi_s^* \oplus \cdots \oplus \psi_1^*) \le s \cdot \mathsf{DecP}^d(\psi_1),$$

where $\psi_1^*, \ldots, \psi_s^*$ are independent URPs on I_n, $\psi_i = \pi_{(i-1)c+c} \circ \cdots \circ \pi_{(i-1)c+1}$ for $1 \le i \le s$, and $\pi_1, \ldots, \pi_{sc}$ are i.i.d. random permutations on I_n. We can think about a lot of candidates for the lower upper bound of δ. Let's consider two basic types: $\mathsf{DecP}^d(\psi_1)^{O(s)}$ and $O\left(\mathsf{DecP}^d(\psi_1)\right)$.

Case 1. $\delta = \mathsf{DecP}^d(\psi_1)^{O(s)}$: Then the following holds.

$$\mathsf{DecF}^d(\text{SUM}^s - \text{CMP}^c) \le \mathsf{DecP}^d(\pi_1)^{O(m)} + \frac{2d}{s+1}\left(\frac{d}{2^{n-1}}\right)^s.$$

In this case, there is no trade-off between composition and XOR given $m = sc$. Namely, for any (n,d,m) and for any π_i, $\mathsf{UB} - \mathsf{DecF}^d(\text{SUM}^m - \text{CMP}^1) < \mathsf{UB} - \mathsf{DecF}^d(\text{SUM}^s - \text{CMP}^c)$ for all $c > 2$.

Case 2. $\delta = O\left(\mathsf{DecP}^d(\psi_1)\right)$: In particular we focus on the case of $\delta \le \mathsf{DecP}^d(\psi_1)$. Then the following holds.

$$\mathsf{DecF}^d(\text{SUM}^s - \text{CMP}^c) \le \mathsf{DecP}^d(\pi_1)^c + \frac{2d}{s+1}\left(\frac{d}{2^{n-1}}\right)^s.$$

In this case, we get almost the same results as those in Sections 4.1 and 4.2. Here, we show the one corresponding to Corollary 10: in most cases, the optimal number of compositions occurs between m_2 and m_u for secure π_i's with $\mathsf{DecP}^d(\pi_1) \le \varepsilon_3$ for some $\varepsilon_3 > \varepsilon_2$. The proof is similar, and so omitted.

Proposition 11. *Define f', C_0', and ε_3 as*

$$\begin{aligned}
f'(n,d,\varepsilon,m,x) &= \varepsilon^x + \tfrac{2dx}{x+m}\left(\tfrac{d}{2^{n-1}}\right)^{\frac{m}{x}};\\
C_0'(n,d,\varepsilon,m) &= \{c_0 \in \text{FAC}(m) \mid f'(n,d,\varepsilon,m,c_0) \le f'(n,d,\varepsilon,m,c) \text{ for all } c \in \text{FAC}(m)\};\\
\varepsilon_3(n,d,m) &= \min\left\{2^{-2},\ \tfrac{d}{2^{n-1}}\left(\tfrac{2d\left(1+(1+\sqrt{m})\ln\frac{2^{n-1}}{d}\right)}{n(1+\sqrt{m})^2}\right)^{\frac{1}{\sqrt{m}}}\right\}.
\end{aligned}$$

For any $n \in [16, \infty)$, $d \in [1, 2^{n/2}]$ *(resp.* $d \in (2^{n/2}, 2^{3n/4}]$*),* $m \in [9, n] \cap \mathbf{CN}$ *(resp.* $m \in [49, n] \cap \mathbf{CN}$*), and* $\varepsilon \in [2^{-n}, \varepsilon_3(n, d, m)]$*, we have* $C_0'(n, d, \varepsilon, m) \subset [m_2, m_u]$.

This result is the same as Corollary 10 except that it holds for a little lager ranges of d and ε.

References

[1] W. Aiello, M. Bellare, G. Di Crescenzo, and R. Venkatesan. Security Amplification by Composition: The Case of Doubly-Iterated, Ideal Ciphers. *CRYPTO '98*, LNCS **1462** (1998) 390–407.

[2] O. Goldreich, S. Goldwasser, and S. Micali. How to Construct Random Functions. *J. ACM* **33** (1986) 792–807.

[3] E. Lee. On the Trade-off between Composition and XOR of Random Permutations. *The Journal of the Korean Institute of Communication Sciences* **31** (2006) 286–292.

[4] M. Luby and C. Rackoff. Pseudorandom Permutation Generators and Cryptographic Composition. *ACM Symposium on Theory of Computing* (1986) 356–363.

[5] S. Lucks. The Sum of PRPs is a Secure PRF. *EUROCRYPT '00*, LNCS **1807** (2000) 470–484.

[6] S. Moriai and S. Vaudenay. On the Pseudorandomness of Top-Level Schemes of Block Ciphers. *ASIACRYPT '00*, LNCS **1976** (2000) 289–302.

[7] S. Vaudenay. Provable Security for Block Ciphers by Decorrelation. *STACS '98*, LNCS **1373** (1998) 249–275.

[8] S. Vaudenay. Adaptive-Attack Norm for Decorrelation and Super-Pseudorandomness. *SAC '99*, LNCS **1758** (2000) 49–61.

[9] S. Vaudenay. Decorrelation: A Theory for Block Cipher Security. *J. Cryptology* **16** (2003) 249–286.

Department of Applied Mathematics, Sejong University, Seoul, 143-747, Korea
E-mail address: eonkyung@sejong.ac.kr