Bounds on Fixed Input/Output Length Post-processing Functions for Biased Physical Random Number Generators

Kyohei Suzuki and Tetsu Iwata

Dept. of Computational Science and Engineering,
Nagoya University
Furo-cho, Chikusa-ku, Nagoya, 464-8603, Japan
kyouhe_s@nuee.nagoya-u.ac.jp, iwata@cse.nagoya-u.ac.jp

Abstract. Post-processing functions are used to reduce the imperfectness of physical random number generators. At FSE '07, Dichtl considered the case where the physical random number generator outputs independent bits that have a constant bias, and the post-processing function has fixed input and output lengths. In this paper, we first present a number of bounds on $\deg(n, m)$, which is a measure of the reduction of biases with n-bit input and m-bit output post-processing functions. We next show the exact values of $\deg(n, m)$ for a large class of (n, m) such that $1 \leq m \leq n \leq 16$, by using the bounds on $\deg(n, m)$ and a computer simulation. We finally discuss how we have derived these numerical values.

Keywords: physical random number generator, bias, post-processing, entropy extractor.

1 Introduction

Background. Cryptographic schemes are designed assuming that unbiased and independent bits are available. However, when we implement them in practice, the physical sources of randomness to which we have access are not perfect, and may contain biases and correlations. For example, we may use system clocks, keyboard or mouse movements, radioactive sources, or quantum mechanical sources (see [9,13] and [20, Chap. 17] for other examples), but they usually do not produce perfect random bits. Many cryptographic schemes rely on sequences of unbiased bits. It is therefore important to be able to extract unbiased bits from an imperfect physical source, and a natural approach to the problem is to apply a *post-processing function* (also called an entropy extractor, or a corrector), a function that transforms a weak random source into an almost perfect random source. This classical problem was extensively studied in the past [1,2,5,6,7,8,12,16,17,18,19,21,22,23,24,25,26],

At FSE '07, Dichtl studied the particular source of randomness, where the output bits of physical source are independent and have a constant (but unknown) bias [10]. That is, if $x_1, x_2, \ldots$ are the output bits of the physical source,

R. Avanzi, L. Keliher, and F. Sica (Eds.): SAC 2008, LNCS 5381, pp. 373–386, 2009.

then $\Pr(x_i = 1) = 1/2 + \epsilon$ holds for some ϵ. This setting may be of practical interest as some of the above sources of randomness, such as radioactive sources and quantum mechanical sources, may output data that are independent but biased. Also, post-processing functions studies in [10] have fixed input and output lengths, and this may be important in a real system as they have a fixed input/output ratio and latency, while, for example, a well known von Neumann's method [26] does not have this property.

Dichtl's result [10]. Dichtl proposed five post-processing functions, called XOR, H, H2, H3 and Solution S. These functions take 16-bit input and produce 8-bit output. For the first four functions, the 16-bit input x is divided into two 8-bit sequences a and b as $x = (a, b)$, and the output y is given by

$$\begin{cases} \text{XOR} : y = a \oplus b \\ \text{H} : y = a \oplus (a \ll 1) \oplus b \\ \text{H2} : y = a \oplus (a \ll 1) \oplus (a \ll 2) \oplus b \\ \text{H3} : y = a \oplus (a \ll 1) \oplus (a \ll 2) \oplus (a \ll 4) \oplus b \end{cases}$$

where $a \ll i$ is the i-bit cyclic left shift of a. Since our input bits have a constant bias, its probability is a polynomial in ϵ, i.e., if x is n bits and its Hamming weight is w, then $\Pr(x) = (1/2 - \epsilon)^{n-w}(1/2 + \epsilon)^w$. Now for post-processing function F, the output probability, $\Pr(y)$, is the sum of input probabilities of x such that $y = F(x)$, which is also a polynomial in ϵ whose degree is at most n. Dichtl proposed to measure the effectiveness of reduction of bias by the "lowest degree of ϵ with non-zero coefficient." Dichtl shows that, for XOR, the coefficient of ϵ in $\Pr(y)$ is zero for any y. Similarly, ϵ and ϵ^2 are zero for H, $\epsilon, \epsilon^2, \epsilon^3$ are zero for H2, and the coefficients of $\epsilon, \ldots, \epsilon^4$ are all zero for H3, thus the lowest degree with non-zero coefficient is 2, 3, 4, and 5 for XOR, H, H2, and H3, respectively. Since the lowest degree of raw input x is 1, they all reduce the bias compared to the raw input, and H3 reduces the bias the most effective way among these four functions.

Solution S is a special type of post-processing functions, where for any input x, x and its complement have the same output value. Solution S is derived by solving the system of linear equations, and the above property reduces the search space since for any y and odd i, the coefficients of ϵ^i in $\Pr(y)$ is zero. In particular, Dichtl shows that, for any y, the coefficients of $\epsilon, \ldots, \epsilon^5$ in $\Pr(y)$ are all zero, thereby reducing more bias than the previous four functions.

We note that post-processing functions in [10] are deterministic, while foundational works [16,21,22,23,24,25] assume a small amount of true randomness, and the works on deterministic extractors [1,2,6,18] directly capture the min-entropy, which is known to be an appropriate notion for random number generation to evaluate the randomness quantity of a binary sequence [3]. As in [10], in this paper, we consider the deterministic functions and use the lowest degree of the polynomial for evaluating the reduction of biases.

Lacharme's result [14]. Lacharme shows that the problem is closely related to the coding theory, i.e., if there exists an $[n, m, d]$ linear code, then there exists an

n-bit input and m-bit output post-processing function such that the coefficients of $\epsilon, \epsilon^2, \ldots, \epsilon^{d-1}$ are all zero. This is a natural generalization of [10], as H, H2 and H3 respectively correspond to generator matrices of [16,8,3], [16,8,4] and [16,8,5] linear codes. Also, a table of linear codes [11] can be used to construct linear post-processing functions. Then Lacharme proposes to use a resilient function as the post-processing function, and shows that, for an (n, m, t)-resilient function, the coefficients of $\epsilon, \epsilon^2, \ldots, \epsilon^t$ are all zero. Finally, Lacharme studies the relation between the bias and the min-entropy.

Our contributions. We first re-formalize the problem explicitly separating the general post-processing functions and the "Solution S type" post-processing functions. For any n-bit input and m-bit output post-processing function F, we let $\mathrm{mindeg}(F)$ be the minimum degree of ϵ with non-zero coefficient in the output probability, where the minimum is taken over all the output value. We then define $\deg(n, m)$ and $\deg^{\mathrm{s}}(n, m)$ to be the maximum of $\mathrm{mindeg}(F)$, where the maximum is taken over all n-bit input and m-bit output post-processing functions for $\deg(n, m)$, and over all "Solution S type" post-processing functions for $\deg^{\mathrm{s}}(n, m)$. In our terminology, Dichtl shows $\deg^{\mathrm{s}}(16, 8) = 6$ and derives the concrete truth table of F achieving $\mathrm{mindeg}(F) = 6$, and Lacharme shows that, if there exists an $[n, m, d]$ linear code, then $\deg(n, m) \geq d$, and if there exists a (n, m, t)-resilient function, then $\deg(n, m) \geq t + 1$.

We then present a number of bounds on $\deg(n, m)$ and $\deg^{\mathrm{s}}(n, m)$. Our bounds are elementary ones and we see that proving these bounds are important in understanding the basic properties of post-processing functions. Indeed, it turns out that they are actually useful in deriving the exact values of $\deg(n, m)$ and $\deg^{\mathrm{s}}(n, m)$. By using a computer simulation and the bounds we have derived, we next present the exact values of $\deg(n, m)$ for $1 \leq m \leq n \leq 16$, and $\deg^{\mathrm{s}}(n, m)$ for $1 \leq m < n \leq 16$. Out of 136 values of (n, m) for $\deg(n, m)$, we have determined 123 values, and out of 120 values for $\deg^{\mathrm{s}}(n, m)$, 115 values are determined. While the exact values for the remaining (n, m) are open, we derive both the upper and lower bounds. We finally discuss in detail how we have derived these numerical values. Our results can be seen as the generalization of [10] from $n = 16$ and $m = 8$ to $1 \leq m < n \leq 16$, and proving the optimality and non-optimality of the results in [14] for $1 \leq m \leq n \leq 16$.

2 Preliminaries

For a positive integer n, $\{0, 1\}^n$ is the set of all n-bit strings. For any set S, $\#S$ is the cardinality of S. An n-bit input and m-bit output post-processing function is a vector output Boolean function $F : \{0, 1\}^n \mapsto \{0, 1\}^m$. Let T_F be its truth table, i.e., $T_F = (F(00\cdots00), F(00\cdots01), \ldots, F(11\cdots11))^t$, which is the transposed vector of $(F(00\cdots00), F(00\cdots01), \ldots, F(11\cdots11))$. We say that F is balanced if each $y \in \{0, 1\}^m$ appears 2^{n-m} times in T_F. A balanced n-bit input and m-bit output post-processing function is denoted (n, m)-PP, and let (n, m)-$\mathcal{PP}$ be the set of all (n, m)-PPs. As in [10], we only consider (n, m)-PPs.

Let $x \in \{0,1\}^n$ be the input of an (n,m)-PP and $y \in \{0,1\}^m$ be its output. Throughout this paper, we assume that each bit of x has a constant (but unknown) bias ϵ, i.e., if $x = (x_1, x_2, \ldots, x_n) \in \{0,1\}^n$ is the input, then $\Pr(x_i = 1) = 1/2 + \epsilon$ for $1 \leq i \leq n$. The Hamming weight of $x = (x_1, x_2, \ldots, x_n)$ is denoted $w(x)$, which is $\#\{i \mid x_i = 1\}$. The probability of input, $\Pr(x)$, depends only on $w(x)$ and is given by

$$\Pr(x) = \left(\frac{1}{2} - \epsilon\right)^{n-w(x)} \left(\frac{1}{2} + \epsilon\right)^{w(x)}. \tag{1}$$

Therefore, $\Pr(x)$ is a polynomial in ϵ. Since $0 \leq w(x) \leq n$, there are $(n+1)$ possibilities for the value of $\Pr(x)$. If $w(x) = w$, the corresponding probability is denoted p_w, and hence $p_w = (1/2 - \epsilon)^{n-w} (1/2 + \epsilon)^w$.

For any $y \in \{0,1\}^m$, $F^{-1}(y)$ is the *preimage* of y, and is defined as the set of x such that $y = F(x)$, i.e., $F^{-1}(y) = \{x \mid y = F(x)\}$. The probability of output, $\Pr(y)$, is the sum of probabilities of 2^{n-m} n-bit inputs belonging to its preimage. That is,

$$\Pr(y) = \sum_{x \in F^{-1}(y)} \Pr(x). \tag{2}$$

Since $\Pr(x)$ is a polynomial in ϵ given by (1), $\Pr(y)$ is also a polynomial in ϵ whose degree is at most n. Therefore, $\Pr(y) = a_0 + a_1\epsilon + a_2\epsilon^2 + \cdots + a_n\epsilon^n$.

Now we define $\mathrm{mindeg}(\Pr(y))$ as follows.

Definition 1. *For all $y \in \{0,1\}^m$, define*

$$\mathrm{mindeg}(\Pr(y)) \stackrel{\mathrm{def}}{=} \min\{k \mid 1 \leq k \leq n, a_k \neq 0\}.$$

For given y, $\mathrm{mindeg}(\Pr(y))$ is the minimum degree other than the constant term.

Next, we define $\mathrm{mindeg}(T_F)$ as follows.

Definition 2. *For all $F \in (n,m)$-$\mathcal{PP}$, define*

$$\mathrm{mindeg}(T_F) \stackrel{\mathrm{def}}{=} \min\{\mathrm{mindeg}(\Pr(y)) \mid y \in \{0,1\}^m\}.$$

For given $F \in (n,m)$-$\mathcal{PP}$, $\mathrm{mindeg}(T_F)$ is the minimum of $\mathrm{mindeg}(\Pr(y))$, where y runs all the possible values.

Then, we define $\deg(n,m)$ as follows.

Definition 3. *For all $n \geq m \geq 1$, define*

$$\deg(n,m) \stackrel{\mathrm{def}}{=} \max\{\mathrm{mindeg}(T_F) \mid F \in (n,m)\text{-}\mathcal{PP}\}.$$

For given (n,m), $\deg(n,m)$ is the maximum value of $\mathrm{mindeg}(T_F)$, where the maximum is taken over all the possible $F \in (n,m)$-$\mathcal{PP}$. It is easy to see that $\#(n,m)\text{-}\mathcal{PP} = (2^n)!/\{(2^{n-m})!\}^{2^m}$, i.e., $F \in (n,m)$-$\mathcal{PP}$ satisfying $\mathrm{mindeg}(T_F) = \deg(n,m)$ reduces the bias most effective way among this number of possible (n,m)-PPs.

Solution S is a special type of an (n, m)-PP, where any input x and its complement, $\bar{x}$, have the same output [10]. This implies that x and $\bar{x}$ belong to the same preimage of some y. Therefore, for any $y \in \{0,1\}^m$ and for all odd i, the coefficient of ϵ^i in $\Pr(y)$ is zero since the coefficient of ϵ^i in $(\Pr(x) + \Pr(\bar{x}))$ is zero. An (n, m)-PP is said to be an (n, m)-SPP (Solution S type PP) if for any x, x and $\bar{x}$ have the same output, and let (n, m)-$\mathcal{SPP}$ be the set of all (n, m)-SPPs.

Now we define $\deg^{\mathrm{s}}(n, m)$ as follows.

Definition 4. *For all $n > m \geq 1$, define*

$$\deg^{\mathrm{s}}(n, m) \stackrel{\mathrm{def}}{=} \max\{\mathrm{mindeg}(T_F) \mid F \in (n, m)\text{-}\mathcal{SPP}\}.$$

Note that $\#(n, m)$-$\mathcal{SPP} = (2^{n-1})!/\{(2^{n-m-1})!\}^{2^m}$, and the maximum is taken over all these (n, m)-SPPs. Also, for any $n > m \geq 1$, we have $\deg(n, m) \geq \deg^{\mathrm{s}}(n, m)$ since (n, m)-$\mathcal{SPP} \subset (n, m)$-$\mathcal{PP}$. We do not consider the case $n = m$ since any $F \in (n, n)$-$\mathcal{PP}$ is a permutation over $\{0,1\}^n$, and thus two distinct inputs cannot have the same output. With the similar reasoning, we do not consider the case $n = 1$.

3 Bounds on $\deg(n, m)$ and $\deg^{\mathrm{s}}(n, m)$

In this section, we present bounds on $\deg(n, m)$ and $\deg^{\mathrm{s}}(n, m)$ with their proofs.

3.1 Bounds on $\deg(n, m)$

We show six bounds on $\deg(n, m)$.

Theorem 1. *For all $n \geq 1$, $\deg(n, n) = 1$.*

Proof. Any (n, n)-PP is a permutation over $\{0,1\}^n$. Therefore, for all $0 \leq w \leq n$, there always exists some $y \in \{0,1\}^n$ such that $\Pr(y) = p_w$. Now since $\mathrm{mindeg}(p_w) \neq 0$ for $0 \leq w \leq n$, $\mathrm{mindeg}(\Pr(y)) \neq 0$ for any $y \in \{0,1\}^n$. This implies $\mathrm{mindeg}(T_F) \geq 1$ for any $F \in (n, n)$-$\mathcal{PP}$.

On the other hand, $\mathrm{mindeg}(p_w) = 1$ holds for some $0 \leq w \leq n$, cf., $w = 0$. Thus, there always exists some $y \in \{0,1\}^n$ such that $\mathrm{mindeg}(\Pr(y)) = 1$. This implies $\mathrm{mindeg}(T_F) \leq 1$ for any $F \in (n, n)$-$\mathcal{PP}$, and hence $\deg(n, n) = 1$. □

Theorem 2. *For all $n \geq 2$, $\deg(n, n-1) = 2$.*

Proof. Constructing an $(n, n-1)$-PP corresponds to dividing 2^n n-bit inputs into 2^{n-1} preimages, where each preimage consists of two inputs. Our proof proceeds in two steps. First, we derive a necessary and sufficient condition that, for each $y \in \{0,1\}^{n-1}$, the coefficient of ϵ in $\Pr(y)$ is zero. Then we show that when the condition is satisfied, the coefficient of ϵ^2 in $\Pr(y)$ is non-zero for some y.

Now $\Pr(x)$ in (1) can be written as

$$\Pr(x) = \left\{ \sum_{i=0}^{n-w} \binom{n-w}{i} \left(\frac{1}{2}\right)^{n-w-i} (-\epsilon)^i \right\} \left\{ \sum_{j=0}^{w} \binom{w}{j} \left(\frac{1}{2}\right)^{w-j} \epsilon^j \right\},$$

where $w(x) = w$. Therefore, the coefficient of ϵ in $\Pr(x)$ is

$$\sum_{i+j=1} \binom{n-w}{i}(-1)^i \binom{w}{j} \left(\frac{1}{2}\right)^{n-i-j} = \frac{2w-n}{2^{n-1}}.$$

Fix any $y \in \{0,1\}^{n-1}$, and suppose that its preimage consists of x_1 and x_2, where $w(x_1) = w_1$ and $w(x_2) = w_2$. Since $\Pr(y) = \Pr(x_1) + \Pr(x_2)$, the coefficient of ϵ in $\Pr(y)$ is zero if and only if

$$\frac{2w_1 - n}{2^{n-1}} + \frac{2w_2 - n}{2^{n-1}} = 0,$$

which is equivalent to $w_2 = n - w_1$. Therefore, the necessary and sufficient condition is to form a preimage with two inputs of weight w and $n - w$. Now, since $\#\{x \mid w(x) = w\} = \#\{x \mid w(x) = n - w\}$ holds for any $0 \le w \le n$, it is possible to satisfy the above condition to construct $F \in (n, n-1)$-$\mathcal{PP}$ satisfying $\mathrm{mindeg}(T_F) \ge 2$.

Next, consider some $F \in (n, n-1)$-$\mathcal{PP}$ satisfying $\mathrm{mindeg}(T_F) \ge 2$. We show that the coefficient of ϵ^2 in $\Pr(y)$ is non-zero for some y. If $w(x) = w$, then the coefficient of ϵ^2 in $\Pr(x)$ is

$$\sum_{i+j=2} \binom{n-w}{i}(-1)^i \binom{w}{j} \left(\frac{1}{2}\right)^{n-i-j} = \frac{(2w-n)^2 - n}{2^{n-1}}.$$

Similarly, if $w(x') = n - w$, we see that ϵ^2 in $\Pr(x')$ has the same coefficient. Therefore, if the preimage is formed with two inputs of weight w and $n - w$, the coefficient of ϵ^2 in $\Pr(y)$ is

$$\frac{(2w-n)^2 - n}{2^{n-1}} + \frac{(2w-n)^2 - n}{2^{n-1}} = \frac{(2w-n)^2 - n}{2^{n-2}}.$$

So we need $(2w-n)^2 - n = 0$ to eliminate ϵ^2, which is equivalent to $w = (n \pm \sqrt{n})/2$. Now it is clear that we always have some w such that $0 \le w \le n$ and $w \ne (n \pm \sqrt{n})/2$ (since $n \ge 2$, we have at least three choices of w, and the right hand side takes at most two values). This implies the coefficient of ϵ^2 in $\Pr(y)$ is non-zero for some y. Therefore, for any $F \in (n, n-1)$-$\mathcal{PP}$, $\mathrm{mindeg}(T_F) \le 2$, and hence $\deg(n, n-1) = 2$. □

Theorem 3. *For all $n > m \ge 1$, $\deg(n, m) \ge \deg(n, m+1)$.*

Proof. Suppose we have $F \in (n, m+1)$-$\mathcal{PP}$, where $\deg(n, m+1) = \mathrm{mindeg}(T_F)$. We construct $F' \in (n, m)$-$\mathcal{PP}$ such that $\mathrm{mindeg}(T_{F'}) \ge \mathrm{mindeg}(T_F)$.

Since the output length of F is $(m+1)$ bits, F has 2^{m+1} preimages, where each preimage has 2^{n-m-1} inputs. Now we divide the 2^{m+1} preimages into 2^m pairs of preimages, and regard the pair of preimages as a new preimage. We then have 2^m new preimages each of which consists of 2^{n-m} inputs, and let F' be the resulting (n, m)-PP.

Let $y_0 \in \{0,1\}^m$ be some output of F' and $\Pr(y_0)$ be its probability. By definition, $\Pr(y_0) = \Pr(y_1) + \Pr(y_2)$ for some outputs $y_1 \in \{0,1\}^{m+1}$ and $y_2 \in \{0,1\}^{m+1}$ of F. Without loss of generality, assume that $\mathrm{mindeg}(\Pr(y_1)) \leq \mathrm{mindeg}(\Pr(y_2))$. Then we have

$$\mathrm{mindeg}(\Pr(y_0)) = \mathrm{mindeg}(\Pr(y_1) + \Pr(y_2)) \geq \mathrm{mindeg}(\Pr(y_1)).$$

Therefore, for any $y' \in \{0,1\}^m$ of F', we have some $y \in \{0,1\}^{m+1}$ of F satisfying $\mathrm{mindeg}(\Pr(y')) \geq \mathrm{mindeg}(\Pr(y))$. This implies $\min\{\mathrm{mindeg}(\Pr(y')) \mid y' \in \{0,1\}^m\} \geq \min\{\mathrm{mindeg}(\Pr(y)) \mid y \in \{0,1\}^{m+1}\}$, and the result follows. □

Theorem 4. *For all $n \geq m \geq 1$, $\deg(n,m) \leq \deg(n+1,m)$.*

Proof. Suppose that we have $F \in (n,m)$-$\mathcal{PP}$ such that $\mathrm{mindeg}(T_F) = \deg(n,m)$. We construct $F' \in (n+1,m)$-$\mathcal{PP}$ satisfying $\mathrm{mindeg}(T_F) = \mathrm{mindeg}(T_{F'})$.

For an input $x' \in \{0,1\}^{n+1}$, the output of F' is $F'(x') = F(x)$, where x is the least significant n bits of x', i.e., F' simply ignores the most significant bit of x'.

Now, for any $x \in \{0,1\}^n$, we have $\Pr(0||x) = (1/2 - \epsilon)\Pr(x)$ and $\Pr(1||x) = (1/2 + \epsilon)\Pr(x)$. Let $\Pr(y)$ be the probability that the output of F is y. Then the probability that F' outputs y is

$$\sum_{x \in F^{-1}(y)} \left(\frac{1}{2} - \epsilon\right) \Pr(x) + \sum_{x \in F^{-1}(y)} \left(\frac{1}{2} + \epsilon\right) \Pr(x) = \Pr(y).$$

Therefore we have $\mathrm{mindeg}(T_{F'}) = \mathrm{mindeg}(T_F)$. □

Theorem 5. *For all $n \geq m \geq 1$ and $k \geq 1$, $\deg(n,m) \leq \deg(kn,km)$.*

Proof. Suppose that we have $F \in (n,m)$-$\mathcal{PP}$ such that $\mathrm{mindeg}(T_F) = \deg(n,m)$. We construct $F' \in (kn,km)$-$\mathcal{PP}$ satisfying $\mathrm{mindeg}(T_F) \leq \mathrm{mindeg}(T_{F'})$.

For an input $x = (x_1, x_2, \ldots, x_k) \in (\{0,1\}^n)^k$ of F', the output is $y = (y_1, y_2, \ldots, y_k) \in (\{0,1\}^m)^k$, where $y_i = F(x_i)$ for $1 \leq i \leq k$.

Since $\Pr(y) = \Pr(y_1)\Pr(y_2)\cdots\Pr(y_k)$, we have

$$\mathrm{mindeg}(\Pr(y)) = \mathrm{mindeg}(\Pr(y_1)\Pr(y_2)\cdots\Pr(y_k)).$$

Now, $\mathrm{mindeg}(\Pr(y_1)\Pr(y_2)\cdots\Pr(y_k)) \geq \mathrm{mindeg}(\Pr(y_i))$ holds for any $1 \leq i \leq k$. Also, from the definition of $\mathrm{mindeg}(T_F)$, we have $\mathrm{mindeg}(\Pr(y_i)) \geq \mathrm{mindeg}(T_F)$ for any $1 \leq i \leq k$. Therefore, $\min\{\mathrm{mindeg}(\Pr(y)) \mid y \in \{0,1\}^{km}\} \geq \mathrm{mindeg}(T_F)$, and the result follows. □

Theorem 6. *For all $n \geq 1$, $\deg(n,1) = n$.*

We use the following Piling-up Lemma [15] to prove Theorem 6.

Lemma 1 (Piling-up Lemma). *Let $n \geq 1$ and $x_1, x_2, \ldots, x_n$ be independent random variables such that $\Pr(x_i = 1) = 1/2 + \epsilon_i$. Then*

$$\Pr(x_1 \oplus x_2 \oplus \cdots \oplus x_n = 1) = \frac{1}{2} + (-2)^{n-1} \prod_{1 \leq i \leq n} \epsilon_i.$$

Now Theorem 6 is proved directly from Lemma 1.

Proof (of Theorem 6). Consider $F(x_1, x_2, \ldots, x_n) = x_1 \oplus x_2 \oplus \cdots \oplus x_n$. We see that $F \in (n,1)$-$\mathcal{PP}$. Lemma 1 shows that the coefficients of $\epsilon, \epsilon^2, \ldots, \epsilon^{n-1}$ are all zero, and therefore, $\deg(n,1) \geq n$. On the other hand, by definition, we have $\deg(n,1) \leq n$. □

3.2 Bounds on $\deg^{\mathrm{s}}(n,m)$

Similarly to $\deg(n,m)$, we show six bounds on $\deg^{\mathrm{s}}(n,m)$.

Theorem 7. *For all $n \geq 2$, $\deg^{\mathrm{s}}(n, n-1) = 2$.*

Proof. This follows since $F \in (n, n-1)$-$\mathcal{PP}$ satisfying $\deg(T_F) \geq 2$ in the proof of Theorem 2 also satisfies $F \in (n, n-1)$-$\mathcal{SPP}$. □

Theorem 8. *For all $n - 1 > m \geq 1$, $\deg^{\mathrm{s}}(n,m) \geq \deg^{\mathrm{s}}(n, m+1)$.*

Proof. Similarly to the proof of Theorem 3, we can construct $F' \in (n,m)$-$\mathcal{SPP}$ such that $\mathrm{mindeg}(T_{F'}) \geq \mathrm{mindeg}(T_F)$ from any $F \in (n, m+1)$-$\mathcal{SPP}$. □

Theorem 9. *For all $n + 1 > m \geq 1$, $\deg^{\mathrm{s}}(n,m) \leq \deg^{\mathrm{s}}(n+1, m)$.*

Proof. A proof is similar to the proof of Theorem 4. For any $F \in (n,m)$-$\mathcal{SPP}$, there exists $F' \in (n+1, m)$-$\mathcal{SPP}$ satisfying $\mathrm{mindeg}(T_F) \leq \mathrm{mindeg}(T_{F'})$. □

Theorem 10. *For all $n > m \geq 1$ and $k \geq 1$, $\deg^{\mathrm{s}}(n,m) \leq \deg^{\mathrm{s}}(kn, km)$.*

Proof. Similarly to the proof of Theorem 5, for any $F \in (n,m)$-$\mathcal{SPP}$, there exists $F' \in (kn, km)$-$\mathcal{SPP}$ satisfying $\mathrm{mindeg}(T_F) \leq \mathrm{mindeg}(T_{F'})$. □

Theorem 11. *For all even $n \geq 2$, $\deg^{\mathrm{s}}(n,1) = n$.*

Proof. We see that $F(x_1, x_2, \ldots, x_n) = x_1 \oplus x_2 \oplus \cdots \oplus x_n \in (n,1)$-$\mathcal{SPP}$, and the rest of the proof is the same as Theorem 6. □

Theorem 12. *For all odd $n \geq 3$, $\deg^{\mathrm{s}}(n,1) = n - 1$.*

Proof. Since n is odd, $n-1$ is even and thus we have $\deg^{\mathrm{s}}(n-1,1) \geq n-1$ from Theorem 11. From Theorem 9, we have $\deg^{\mathrm{s}}(n-1,1) \leq \deg^{\mathrm{s}}(n,1)$, and therefore, $n - 1 \leq \deg^{\mathrm{s}}(n,1)$. On the other hand, we always have $\deg^{\mathrm{s}}(n,1) \leq n$, but since $\deg^{\mathrm{s}}(n,1)$ cannot be odd from the definition of $\deg^{\mathrm{s}}(n,1)$, we have $\deg^{\mathrm{s}}(n,1) \leq n-1$. □

4 Simulation Results

4.1 Values of $\deg(n,m)$ and $\deg^{\mathrm{s}}(n,m)$

We first present our simulation results in Table 1 and Table 2. Table 1 shows the values of $\deg(n,m)$ for $1 \leq m \leq n \leq 16$, and Table 2 shows $\deg^{\mathrm{s}}(n,m)$ for $1 \leq m < n \leq 16$. Table 3 is our environment for this simulation.

- In Table 1 and Table 2, if the entry is a^1, then the value is derived by our computer simulation discussed in the following sections.
- In Table 1, a^2 means that the upper bound is derived by our simulation (i.e., $\deg(n,m) \leq a$) and we apply Theorem 3 to derive the lower bound (i.e., $\deg(n,m) \geq a$).
- a^3 means that the upper bound is derived by our simulation and we apply Theorem 4 to derive the lower bound.
- a^4 means that the upper bound is derived by our simulation and the lower bound is taken from Lacharme's results [14].
- If the entry is a^1, b^1, then this means $\deg(n,m) = a$ or b for Table 1 and $\deg^{\mathrm{s}}(n,m) = a$ or b for Table 2, where both values are derived by our simulation. Similarly, if the entry is a^1, b^5, this means that the upper bound a is derived by our simulation and the lower bound b is taken from Lacharme's results [14]. The exact value for these entries remains as an open question.
- In Table 2, a^6 means that the upper bound is derived by our simulation and we apply Theorem 8 to derive the lower bound.
- a^7 means that the upper bound is derived by our simulation and we apply Theorem 9 to derive the lower bound.
- $\deg^{\mathrm{s}}(16,8) = 6$, which is denoted 6^8 in Table 2, is the value from [10].

The entry with underline shows that the bound is strictly better than the one given by the t-resilient functions in [14]. For example, $\deg(10,2) = 7$, but it is known that $(10,2,6)$-resilient function does not exist [4, Theorem 2, 3], and hence $F \in (10,2)$-PP such that $\mathrm{mindeg}(T_F) = 7$ cannot be a resilient function. For all the entries with underline, we have used the bound on t from [4].

Table 1 and Table 2 may be used to determine the values of n and m (and hence the input/output ratio) given the maximum bias that can be accepted for the application.

4.2 How to Derive $\deg(n,m)$

In this section, we discuss how we have derived numerical values of $\deg(n,m)$ in Table 1. We divide 2^n n-bit inputs into 2^m preimages, where each preimage has 2^{n-m} n-bit inputs. Consider some preimage, and let q_w be the number of x such that $w(x) = w$ in that preimage. Therefore, we require that

$$\sum_{w=0}^{n} q_w = 2^{n-m}. \tag{3}$$

If $w(x) = w$, then the coefficient of ϵ^l in $\Pr(x)$ is

$$\sum_{i+j=l} \binom{n-w}{i} (-1)^i \binom{w}{j} \left(\frac{1}{2}\right)^{n-i-j}.$$

Now consider the output probability of this preimage. The necessary and sufficient condition that the coefficients of $\epsilon, \epsilon^2, \ldots, \epsilon^e$ are all zero is;

$$\text{for } 1 \leq l \leq e,\ \sum_{w=0}^{n} \left\{ \sum_{i+j=l} \binom{n-w}{i} (-1)^i \binom{w}{j} \left(\frac{1}{2}\right)^{n-i-j} q_w \right\} = 0. \tag{4}$$

Table 1. The values of $\deg(n, m)$ for $1 \leq m \leq n \leq 16$

$n\backslash m$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
1	1^1															
2	2^1	1^1														
3	3^1	2^1	1^1													
4	4^1	2^1	2^1	1^1												
5	5^1	3^1	2^1	2^1	1^1											
6	6^1	4^1	3^1	2^1	2^1	1^1										
7	7^1	4^1	4^1	3^1	2^1	2^1	1^1									
8	8^1	5^1	4^1	4^1	2^1	2^1	2^1	1^1								
9	9^1	6^1	$\underline{5}^1$	4^1	3^1	2^1	2^1	2^1	1^1							
10	10^1	$\underline{7}^1$	$\underline{6}^1$	$\underline{5}^1$	4^1	3^1	2^1	2^1	2^1	1^1						
11	11^1	$\underline{8}^1$	6^2	$\underline{6}^1$	4^2	4^1	3^1	2^1	2^1	2^1	1^1					
12	12^1	8^1	$\underline{7}^1, 6^1$	6^3	$5^1, 4^1$	4^1	4^1	3^1	2^1	2^1	2^1	1^1				
13	13^1	$\underline{10}^1$	$\underline{8}^1$	$\underline{7}^1, 6^1$	6^1	$5^1, 4^1$	4^2	4^1	3^1	2^1	2^1	2^1	1^1			
14	14^1	$\underline{10}^1$	$\underline{9}^1, 8^1$	$\underline{8}^1, 7^5$	6^3	$6^1, 5^5$	$5^1, 4^1$	4^2	4^1	3^4	2^1	2^1	2^1	1^1		
15	15^1	$\underline{11}^1, 10^1$	$\underline{10}^1$	8^4	7^4	6^2	6^1	$5^1, 4^1$	4^2	4^1	3^4	2^1	2^1	2^1	1^1	
16	16^1	$\underline{12}^1$	$\underline{10}^3$	$\underline{9}^1, 8^5$	8^4	$7^1, 6^1$	$7^1, 6^1$	6^1	4^2	4^2	4^1	2^1	2^1	2^1	2^1	1^1

Table 2. The values of $\deg^{\mathrm{s}}(n, m)$ for $1 \leq m < n \leq 16$

$n\backslash m$	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
2	2^1														
3	2^1	2^1													
4	4^1	2^1	2^1												
5	4^1	2^1	2^1	2^1											
6	6^1	4^1	2^1	2^1	2^1										
7	6^1	4^1	4^1	2^1	2^1	2^1									
8	8^1	4^1	4^1	4^1	2^1	2^1	2^1								
9	8^1	6^1	4^1	4^1	2^1	2^1	2^1	2^1							
10	10^1	6^1	6^1	4^1	4^1	2^1	2^1	2^1	2^1						
11	10^1	8^1	6^1	6^1	4^1	4^1	2^1	2^1	2^1	2^1					
12	12^1	8^1	6^7	6^7	4^1	4^1	4^1	2^1	2^1	2^1	2^1				
13	12^1	10^1	8^1	6^6	6^1	4^6	4^6	4^1	2^1	2^1	2^1	2^1			
14	14^1	10^1	8^7	$8^1, 6^1$	6^7	$6^1, 4^1$	4^6	4^6	4^1	2^1	2^1	2^1	2^1		
15	14^1	10^1	10^1	$8^1, 6^1$	6^6	6^6	6^1	4^6	4^6	4^1	2^1	2^1	2^1	2^1	
16	16^1	12^1	10^7	$8^1, 6^1$	$8^1, 6^1$	6^6	6^6	6^8	4^6	4^6	4^1	2^1	2^1	2^1	2^1

Table 3. Environment for the simulation

Machine	Dell OPTIPLEX GX620
CPU	Pentium(R) 4 CPU 3.40GHz
OS	Microsoft Windows XP Professional SP2
Memory	4GB
Software	Wolfram Mathematica 6.0.1.0

The first step is to derive all the possible values of $\{q_0, q_1, \ldots, q_n\}$ that satisfy both (3) and (4). Suppose that we have d solutions, $\{q_0^{(1)}, q_1^{(1)}, \ldots, q_n^{(1)}\}$, $\{q_0^{(2)}, q_1^{(2)}, \ldots, q_n^{(2)}\}$, $\ldots$, $\{q_0^{(d)}, q_1^{(d)}, \ldots, q_n^{(d)}\}$. Let $Q_k = \{q_0^{(k)}, q_1^{(k)}, \ldots, q_n^{(k)}\}$ for $1 \le k \le d$. If we construct a preimage with $q_t^{(k)}$ inputs of weight t, then the resulting F is balanced (from (3)) and $\epsilon, \epsilon^2, \ldots, \epsilon^e$ are all eliminated (from (4)).

Now we have to construct 2^m preimages with the constraint that we have exactly $\binom{n}{w}$ inputs $x \in \{0,1\}^n$ such that $w(x) = w$. Therefore, the next step is to solve the following linear system;

$$\begin{bmatrix} z_1 & z_2 & \cdots & z_d \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_d \end{bmatrix} = \begin{bmatrix} \binom{n}{0} & \binom{n}{1} & \cdots & \binom{n}{n} \end{bmatrix}. \tag{5}$$

If there exists some $(z_1, z_2, \ldots, z_d)$ satisfying (5), this means it is possible to fulfill the above mentioned constraint, and thus we conclude $\deg(n, m) \ge e + 1$. Otherwise $\deg(n, m) \le e$.

See Appendix for an example to derive $\deg(4, 2)$.

4.3 How to Derive $\deg^{\mathrm{s}}(n, m)$

In (n, m)-SPPs, any input x and its complement, $\bar{x}$, have the same output. Thus we consider x and $\bar{x}$ as the pair $(x, \bar{x})$. Let $w(x, \bar{x}) = \min(w(x), w(\bar{x}))$, i.e.,$w(x, \bar{x})$ is the minimum value of $w(x)$ and $w(\bar{x})$. Now, in $F \in (n, m)$-$\mathcal{SPP}$, we have to divide 2^{n-1} input pairs into 2^m preimages, where each preimage has 2^{n-m-1} input pairs. Consider some preimage, and let $q_{w'}$ be the number of pairs $(x, \bar{x})$ such that $w(x, \bar{x}) = w'$ in that preimage. Then, we need

$$\sum_{w'=0}^{\lfloor n/2 \rfloor} q_{w'} = 2^{n-m-1}, \tag{6}$$

where $\lfloor n/2 \rfloor$ is maximum integer at most $n/2$.

If l is odd, then the coefficient of ϵ^l in $\Pr(x) + \Pr(\bar{x})$ is zero. Otherwise the coefficient is

$$2 \sum_{i+j=l} \binom{n-w'}{i} (-1)^i \binom{w'}{j} \left(\frac{1}{2}\right)^{n-i-j},$$

where $w(x, \bar{x}) = w'$. Consider the output probability of this preimage, and the coefficients of $\epsilon, \epsilon^2, \ldots, \epsilon^e$ are all zero iff;

$$\text{for even } 1 \le l \le e, \ \sum_{w'=0}^{\lfloor n/2 \rfloor} \left\{ 2 \sum_{i+j=l} \binom{n-w'}{i} (-1)^i \binom{w'}{j} \left(\frac{1}{2}\right)^{n-i-j} q_{w'} \right\} = 0. \tag{7}$$

Similarly to $\deg(n,m)$, we first solve a linear system of (6) and (7). Suppose that we have d' solutions, $Q_k = \{q_0^{(k)}, q_1^{(k)}, \ldots, q_{\lfloor n/2 \rfloor}^{(k)}\}$ for $1 \le k \le d'$, and consider the following linear system;

$$\begin{bmatrix} z_1 & z_2 & \cdots & z_{d'} \end{bmatrix} \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_{d'} \end{bmatrix} = \begin{bmatrix} \binom{n}{0} & \binom{n}{1} & \cdots & \binom{n}{\lfloor n/2 \rfloor} \end{bmatrix}. \tag{8}$$

If some $(z_1, z_2, \ldots, z_{d'})$ satisfies (8), then $\deg^{\mathrm{s}}(n,m) \ge e+2$. Otherwise we conclude that $\deg^{\mathrm{s}}(n,m) \le e$.

5 Summary of Results

In this paper, we have generalized the work in [10] in various ways. We first redefined $\deg(n,m)$ and $\deg^{\mathrm{s}}(n,m)$, and then presented twelve bounds on them. We believe that these bounds are important in understanding the basic properties of post-processing functions, and some of them are useful in deriving the exact values of $\deg(n,m)$ and $\deg^{\mathrm{s}}(n,m)$. We derived the tables of $\deg(n,m)$ for $1 \le m \le n \le 16$, and $\deg^{\mathrm{s}}(n,m)$ for $1 \le m < n \le 16$, and discussed how we have derived these numerical values. Several values of $\deg(n,m)$ and $\deg^{\mathrm{s}}(n,m)$ are left as open questions.

Our results suggest that, for some n and m, the resilient function is not the optimal solution as a post-processing function, and it would be interesting to see systematic constructions of optimal functions.

Acknowledgements

The authors would like to thank anonymous reviewers of SAC 2008 for their extensive and useful comments that significantly improved this paper. Especially, several open problems posed in the earlier version were solved by the comments. We also would like to thank Markus Dichtl and Takeshi Koshiba for useful feedbacks.

References

1. Barak, B., Impagliazzo, R., Wigderson, A.: Extracting randomness using few independent sources. SIAM J. Comput. 36(4), 1095–1118 (2006)
2. Barak, B., Kindler, G., Shaltiel, R., Sudakov, B., Wigderson, A.: Simulating independence: New constructions of condensers, Ramsey graphs, dispersers, and extractors. In: 37th STOC, pp. 1–10 (2005)
3. Barker, E., Kelsey, J.: Recommendation for random number generation using deterministic random bit generators (revised). NIST Special Publication 800-90 (2007), `http://csrc.nist.gov/publications/PubsSPs.html`

4. Bierbrauer, J., Gopalakrishnan, K., Stinson, D.R.: Bounds for resilient functions and orthogonal arrays. In: Desmedt, Y.G. (ed.) CRYPTO 1994. LNCS, vol. 839, pp. 247–256. Springer, Heidelberg (1994)
5. Blum, M.: Independent unbiased coin flips from a correlated biased source: A finite Markov chain. Combinatorica 6(2), 97–108 (1986)
6. Bourgain, J.: More on the sum-product phenomenon in prime fields and its applications. International Journal of Number Theory 1, 1–32 (2005)
7. Chor, B., Friedman, J., Goldreich, O., Håstad, J., Rudich, S., Smolensky, R.: The bit extraction problem or t-resilient functions. In: 26th FOCS, pp. 396–407 (1985)
8. Chor, B., Goldreich, O.: Unbiased bits from sources of weak randomness and probabilistic communication complexity. SIAM J. Comput. 17(2), 230–261 (1988)
9. Davis, D., Ihaka, R., Fenstermacher, P.: Cryptographic randomness from air turbulence in disk drives. In: Desmedt, Y.G. (ed.) CRYPTO 1994. LNCS, vol. 839, pp. 114–120. Springer, Heidelberg (1994)
10. Dichtl, M.: Bad and good ways of post-processing biased physical random numbers. In: Biryukov, A. (ed.) FSE 2007. LNCS, vol. 4593, pp. 137–152. Springer, Heidelberg (2007)
11. Grassl, M.: Code tables: Bounds on the parameters of various types of codes (2008), `http://www.codetables.de/`
12. Juels, A., Jakobsson, M., Shriver, E., Hillyer, B.K.: How to turn loaded dice into fair coins. IEEE Trans. Inform. Theory 46(3), 911–921 (2000)
13. Lacy, J.B., Mitchell, D.P., Schell, W.M.: Cryptolib: Cryptography in software. In: Proc. 4th USENIX Symposium (1993)
14. Lacharme, P.: Post-processing functions for a biased physical random number generator. In: Nyberg, K. (ed.) FSE 2008. LNCS, vol. 5086, pp. 334–342. Springer, Heidelberg (2008)
15. Matsui, M.: Linear cryptanalysis method for DES cipher. In: Helleseth, T. (ed.) EUROCRYPT 1993. LNCS, vol. 765, pp. 386–397. Springer, Heidelberg (1994)
16. Nisan, N., Ta-Shma, A.: Extracting randomness: A survey and new constructions. JCSS 58(1), 148–173 (1999)
17. Peres, Y.: Iterating von Neumann's procedure for extracting random bits. The Annals of Statistics 20(3), 590–597 (1992)
18. Raz, R.: Extractors with weak random seeds. In: 37th STOC, pp. 11–20 (2005)
19. Santha, M., Vazirani, U.V.: Generating quasi-random sequences from semi-random sources. JCSS 33, 75–87 (1986)
20. Schneier, B.: Applied cryptography. John Wiley & Sons, Inc., Chichester (1996)
21. Shaltiel, R.: Recent developments in explicit constructions of extractors. Bulletin of the European Association for Theoretical Computer Science (EATCS), 77 (2002)
22. Shaltiel, R., Umans, C.: Simple extractors for all min-entropies and a new pseudorandom generator. JACM 52(2), 172–216 (2005)
23. Ta-Shma, A.: On extracting randomness from weak random sources. In: STOC 1996, pp. 276–285 (1996)
24. Ta-Shma, A., Umans, C., Zuckerman, D.: Loss-less condensers, unbalanced expanders, and extractors. Combinatorica 27(2), 213–240 (2007)
25. Ta-Shma, A., Zuckerman, D., Safra, S.: Extractors from Reed-Muller codes. JCSS 72(5), 786–812 (2006)
26. von Neumann, J.: Various techniques used in connection with random digits. Applied Mathematics Series, U.S. National Bureau of Standards, vol. 12, pp. 36–38 (1951)

A Deriving deg(4, 2)

We show a small example to derive $\deg(4, 2)$.

We divide sixteen 4-bit inputs into four preimages, where each preimage has four 4-bit inputs. Consider some preimage, and let q_w be the number of x such that $w(x) = w$ in that preimage. Therefore, we need

$$q_0 + q_1 + q_2 + q_3 + q_4 = 4. \tag{9}$$

Now the coefficient of ϵ in $\Pr(y)$ is zero iff

$$-\frac{1}{2}q_0 - \frac{1}{4}q_1 + \frac{1}{4}q_3 + \frac{1}{2}q_4 = 0, \tag{10}$$

which corresponds to $l = 1$ in (4). Similarly, the coefficient of ϵ^2 is zero iff

$$\frac{3}{2}q_0 - \frac{1}{2}q_2 + \frac{3}{2}q_4 = 0. \tag{11}$$

We have seven solutions that satisfy both (9) and (10), and consider the following linear system;

$$\begin{bmatrix} z_1 & z_2 & z_3 & z_4 & z_5 & z_6 & z_7 \end{bmatrix} \begin{bmatrix} 0&0&4&0&0 \\ 0&1&2&1&0 \\ 0&2&0&2&0 \\ 0&2&1&0&1 \\ 1&0&1&2&0 \\ 1&0&2&0&1 \\ 1&1&0&1&1 \end{bmatrix} = \begin{bmatrix} 1&4&6&4&1 \end{bmatrix}, \tag{12}$$

where the matrix in (12) corresponds to the seven solutions. Now since

$$(z_1, z_2, \ldots, z_7) = (1, 1, 1, 0, 0, 0, 1)$$

satisfies (12), we conclude that $\deg(4, 2) \geq 2$.

On the other hand we only have one solution satisfying (9), (10) and (11), which is $(q_0, \ldots, q_4) = (0, 2, 0, 2, 0)$. Now we consider the following linear system;

$$\begin{bmatrix} z_1 \end{bmatrix} \begin{bmatrix} 0&2&0&2&0 \end{bmatrix} = \begin{bmatrix} 1&4&6&4&1 \end{bmatrix}.$$

Since there is no solution for this system, we have $\deg(4, 2) \leq 2$, and therefore, $\deg(4, 2) = 2$.