ANGEWANDTE MATHEMATIK UND INFORMATIK
UNIVERSITÄT ZU KÖLN

Report No. 96.240

Amplifying the Security of One-Way Functions
—
A Proof of Yao's XOR-Lemma

by

Frank Damm
Franz-Peter Heider

September 1996

Frank Damm
Institut für Informatik
Universität zu Köln
Pohligstraße 1
D-50969 Köln
e-mail: fdamm @ informatik.uni-koeln.de

Franz-Peter Heider
Mathematisches Institut
Universität zu Köln
Weyertal 86-90
D-50931 Köln

Keywords: computational complexity, cryptography, theory of computation, one-way functions
AMS Classification: 94A60, 68P25, 68Q15, 68Q25

Amplifying the Security of One-Way Functions
—
A Proof of Yao's XOR-Lemma

by

Frank Damm and Franz-Peter Heider

Abstract

In this paper we give a consistent and simple proof for the XOR-Lemma which was hinted at by Yao in [3] and subsequently presented by him in lectures. It can be found in print in [2].

By the lemma we know that the security of any one-way function $f : X \longrightarrow \{0,1\}$ can be substantially amplified if the function is replaced by the XOR with itself, namely by $f \oplus f : X \times X \longrightarrow \{0,1\}, (x,y) \longrightarrow f(x) \oplus f(y)$.

Applications are in cryptography and complexity theory. However, the existence of one-way functions still remains an open problem.

1 Introduction

One-way functions can be verbally defined as function families $(f_n)_{n \in \mathrm{N}}$ that are computable in time polynomial in n and can not be successfully inverted by any probabilistic algorithm of polynomial running time. Although their existence is unproven, it is common cryptographic praxis to build the security of encryption systems, digital signatures and cryptographic hash functions on the hope of these being instances of one-way functions.

The main part of this paper contains the proof of Yao's XOR-Lemmma which gives a technique of amplifying the security of one-way functions. This technique was first mentioned by Yao in a half sentence of [3]. Furtheron, it was formalised in lectures by Yao (cited in [2]) and printed in the book by Kranakis [2]. This book contains a proof too, which however is hard to follow. The technical method of the intuitive part is quite different from that of the formal part in the proof, and several times arguments far stronger than necessary are employed. Considerung the length of the proof, some open assertions put into the exercises and the correctness of the lemma, it is not possible to definitely say the proof is wrong. Anyway, we give a proof now that is consistent and simple. The definition 1 makes the start of the formal part by giving us the notion of unapproximable predicates and friendship functions. Section 2 states the assertion of the XOR-Lemma and section 3 contains its proof.

Definition 1 *Given a family of permutations $\mathcal{F} = (f_n)_{n \in \mathrm{N}}$ defined on $\mathcal{X} = (X_n)_{n \in \mathrm{N}}$, $X_n \subseteq \{0,1\}^n$, hence $f_n(X_n) = X_n$. Let f_n be a polynomial time computable function. Let $\mathcal{B} = (B_n)_{n \in \mathrm{N}}$ be a family of predicates $B_n : X_n \longrightarrow \{0,1\}$.*

(i) *$\mathcal{B}$ is called* unapproximable *if and only if for every polynomial q, for every polynomial time computable probabilistic algorithm $\mathcal{C} = (C_n)_{n\in\mathbb{N}}$ there is an $n_0 \in \mathbb{N}$, such that for all $n \in \mathbb{N}$ with $n \geq n_0$*

$$Prob\,(x \in X_n : B_n(x) = C_n(x)) < \frac{1}{2} + \frac{1}{q(n)}$$

(ii) *$\mathcal{F}$ is called a* friendship function *for $\mathcal{B}$ if and only if $\mathcal{B}$ is unapproximable on $(f_n(X_n))_{n\in\mathbb{N}}$ and there are polynomial time computable algorithms for the calculation of $f_n(x)$ from input (n, x) and for the calculation of $B_n(f_n(x))$ from (n, x).*

$$\begin{tikzcd}
X_n \arrow[d, "\text{(polynomial) } f_n"'] \arrow[dr, "\substack{B_n \circ f_n \\ \text{(polynomial)}}"] & \\
X_n \arrow[r, "\substack{B_n \\ \text{(not polynomial)}}"'] & \{0,1\}
\end{tikzcd}
\qquad
\begin{tikzcd}
x \arrow[d, "f_n"'] \arrow[dr, "B_n \circ f_n"] & \\
f_n(x) \arrow[r, "B_n"'] & B_n(f_n(x))
\end{tikzcd}$$

Figure 1: Unapproximable predicate $\mathcal{B}$ with friendship function $\mathcal{F}$

2 Amplification of Security by XOR

Consider a family of predicates $\mathcal{B} := \{B_m \mid m \in \mathbb{N}\}$ defined on $\mathcal{X} = (X_n)_{n\in\mathbb{N}}$ with $(X_n) \subseteq \{0,1\}^n$: $B_n : X_n \longrightarrow \{0,1\}$. Further consider a family of friendship functions $\mathcal{F} = (f_n)_{n\in\mathbb{N}}$ to $\mathcal{B}$, f_n being a permutation on X_n.

Let $M \subseteq \mathbb{N}$ be infinite and for each $m \in M$

$$\varepsilon_m \in \mathbb{R}_+, \varepsilon_m < 1/2$$
$$\delta_m \in \mathbb{R}_+, \delta_m < 1$$

such that $\rho_1(m) := 1/\delta_m$ and $\rho_2(m) := 1/\varepsilon_m$ are polynomial in m.

Now we are prepared for the XOR-Lemma:

Theorem 2 *If there is a polynomial time computable algorithm $\mathcal{C} = \{C_m \mid m \in \mathbb{N}\}$ such that for each $m \in M$*

$$Prob\Big((x, y) \in X_m^2 : C_m(x, y) = B_m(x) \oplus B_m(y)\Big) \geq \frac{1}{2} + \varepsilon_m$$

then we can obtain a polynomial time computable algorithm $\mathcal{D} = \{D_m \mid m \in \mathbb{N}\}$ from $\mathcal{C}$ such that for all $m \in M$

$$Prob\,(x \in X_m : D_m(x) = B_m(x)) \geq \frac{1}{2} + (1 - \delta_m)\sqrt{\frac{\varepsilon_m}{2}}$$

3 Proof of the Theorem

The proof is given in several parts. First, the algorithm $\mathcal{D}$ is constructed, followed by a motivation for its construction. Subsequently, the technical proof for the high probability of calculating $\mathcal{B}$ correctly is presented in three more pieces. The first of these pieces investigates into the probabilities for $\mathcal{D} = \mathcal{B}$ per case of $\mathcal{D}$. The set $\mathcal{X}$ is approximated by a random sample of polynomial size and the extent of this approximation is proved in the second piece. The final piece combines the results obtained in the restriction to each case of $\mathcal{D}$.

3.1 Construction of the Algorithm $\mathcal{D}$

For all $m \in M$ let $t_m := \# X_m$ and

$$\begin{array}{rclcrcl} \eta_m & := & \left(1 - \frac{\delta_m}{2}\right)\sqrt{\frac{\varepsilon_m}{2}} & \qquad & \nu_m & := & \frac{\delta_m}{2}\sqrt{\frac{\varepsilon_m}{2}} \\ s_m & := & 2\left\lceil \frac{1}{\nu_m^2} \right\rceil^2 + 1 & & l_m & := & s_m^5 \end{array}$$

In the following, m will be a fixed value and not explicitely written. Therefore D, C, B, f, X, ε, δ, η, ν, s, l, t will denote $D_m, C_m, B_m, f_m, X_m, \varepsilon_m, \delta_m, \eta_m, \nu_m, s_m, l_m, t_m$.

η and ν are made to fulfill

$$0 < \eta < \frac{1}{2}, 0 < \nu < \frac{1}{4}, \eta + \nu = \sqrt{\frac{\varepsilon}{2}}$$

For its work, algorithm D takes as an input a two dimensional sample from $X \times X$. The size of the sample will be $l \times s$. Here s is polynomial in m and always odd. l is polynomial in m too, but much larger than s. η was chosen to get a distance of ν between $1/2 + \eta$ and the value $1/2 + (1-\delta)\sqrt{\varepsilon/2}$ from the assertion of the theorem:

$$\frac{1}{2} + (1-\delta)\sqrt{\frac{\varepsilon}{2}} = \frac{1}{2} + \eta - \nu$$

Algorithm D works in two steps. In the step 1 it does a precalculation and takes as an input the sample $(x_1, \ldots, x_l, y_1, \ldots, y_s)$ from X^{l+s} and the values of B at these points: $(B(x_1), \ldots, B(x_l), B(y_1), \ldots, B(y_s))$ To obtain these values, the friendship function is employed. Let $(z_1, \ldots, z_{l+s})$ be from X and calculate

$$\begin{array}{lcll} x_i := f(z_i) & \text{and} & B(x_i) := B(f(z_i)) & \text{for } i = 1, \ldots, l \\ y_i := f(z_{i+l}) & \text{and} & B(y_i) := B(f(z_{i+l})) & \text{for } i = 1, \ldots, s \end{array}$$

Since f is a permutation, the random draw of $(z_1, \ldots, z_{l+s})$ from X is equivalent to a random draw of $(x_1, \ldots, x_l, y_1, \ldots, y_s)$ from X.

In the step 2 of D the main calculation is executed, based on the result of step 1. Step 2 takes as an input the value $x \in X$ and produces a value from $\{0, 1\}$ as an output.

	D
Step 1	
sample input	$(x_1,\dots,x_l,y_1,\dots,y_s)\in X^{l+s}$, $(B(x_1),\dots,B(x_l),B(y_1),\dots,B(y_s))\in\{0,1\}^{l+s}$

calculation	for every $i\in\{1,\dots,l\}$:
	$k_y(x_i,B(x_i)) := \#\{j\in\{1,\dots,s\}: C(x_i,y_j)=B(x_i)\oplus B(y_j)\}$,
	case 1: $\exists\, x_0\in\{x_1,\dots,x_l\}: \lvert k_y(x_0,B(x_0))-\frac{s}{2}\rvert\ge s\eta$;
	if so: stop step 1
	case 2: otherwise. Continue for every $k\in\{0,\dots,s\}$:
	$\sigma(k) := \#\{x_i\in\{x_1,\dots,x_l\}: k_y(x_i,B(x_i))=k\}$

Step 2	
input	$x\in X$
calculation	if case 1 holds: $D(x) := \begin{cases} C(x_0,x)\oplus B(x_0)\oplus 1 & \text{für } k_y(x_0,B(x_0))<\frac{s}{2}\\ C(x_0,x)\oplus B(x_0) & \text{für } k_y(x_0,B(x_0))>\frac{s}{2}\end{cases}$ if case 2 holds: calculate $k_y(x,0) := \#\{j\in\{1,\dots,s\}: C(x,y_j)=B(y_j)\}$ subcase 1: $\sigma(k_y(x,0))>\sigma(s-k_y(x,0))$; let $D(x):=0$ subcase 2: $\sigma(k_y(x,0))<\sigma(s-k_y(x,0))$; let $D(x):=1$ subcase 3: $\sigma(k_y(x,0))=\sigma(s-k_y(x,0))$; let $D(x) := \begin{cases} 0 & \text{with probability } 1/2\\ 1 & \text{with probability } 1/2\end{cases}$
output	$D(x)$

D is a polynomial time computable algorithm, because $\rho_1(m)$ and $\rho_2(m)$ are polynomial in m and the friendship function f as well as the algorithm C are polynomial time computable.

3.2 Motivation of Algorithm D

D uses its knowledge of the values of B at $x_1,\dots,x_l$, $y_1,\dots,y_s$ to approximate the value $B(x)$ for an input x. Because of the assumption of the theorem, algorithm $C(.,.)$ is able to calculate the XOR of B with itself with probability of at least $\frac{1}{2}+\varepsilon$. The values $x_1,\dots,x_l$ are used as the first argument (x-direction) of $C(.,.)$ respectively $B(.)\oplus B(.)$, while the values $y_1,\dots,y_s$ are used as the second argument (y-direction).

Given an $x_i\in\{x_1,\dots,x_l\}$, $k_y(x_i,B(x_i))$ counts the number of y's in the sample such that $C(x_i,.)$ is successful in calculating $B(x_i)\oplus B(.)$. Figures 2, 3, 4 show $k_y(x_i,B(x_i))$ versus $x_i\in\{x_1,\dots,x_l\}$. In these figures, the values in x-direction are permuted to obtain k_y in increasing order. Figures 2 and 3 correspond to case 1 of D, while figure 4 corresponds to case 2 (subcase 2).

Case 1 yields an x_0 in the sample with

$$\left|k_y(x_0,B(x_0))-\frac{s}{2}\right|\ge s\eta$$

In this case there are two possibilities. Considering x_0 and $y_1, \ldots, y_s$ either $C(x_0, y_j) = B(x_0) \oplus B(y_j)$ holds for much fewer than half of the y-values, or it holds for much more than half of them. If it holds for fewer than half of the y's, $k_y(x_0, B(x_0)) < s/2$, which will be denoted as subcase 1 and is shown in figure 2. Subcase 2 is shown in figure 3. The minority of y-values with $C(x_0, y_j) = B(x_0) \oplus B(y_j)$ in subcase 1 correspond to a majority of y-values with $C(x_0, y_j) = B(x_0) \oplus B(y_j) \oplus 1$. Because of the clear behaviour of the y-part of the sample at x_0, it can be expected that for a sufficient majority of inputs x the same equations hold, i.e.

$$\begin{array}{ll} C(x_0, x) = B(x_0) \oplus B(x) \oplus 1 & \text{in subcase 1} \\ C(x_0, x) = B(x_0) \oplus B(x) & \text{in subcase 2} \end{array}$$

Therefore

$$\begin{array}{ll} D(x) := C(x_0, x) \oplus B(x_0) \oplus 1 & \text{in subcase 1} \\ D(x) := C(x_0, x) \oplus B(x_0) & \text{in subcase 2} \end{array}$$

seems worth trying.

If the second case holds, the number of hits in y-direction for $C(x_i, y_j) = B(x_i) \oplus B(y_j)$ are close to $s/2$ (figure 4). In this case, the input x is used as first argument in $C(.,.)$ and its behaviour in y-direction would be given by the value $k_y(x, B(x))$. If that value was known, x could be compared to all x_i in the x-direction of the sample with equal behaviour, i.e. $k_y(x_i, B(x_i)) = k_y(x, B(x))$. While $B(x)$ is not available, there are two possibilities only:

$$k_y(x, B(x)) = \begin{cases} k_y(x, 0) & \text{if} \quad B(x) = 0 \\ k_y(x, 1) = s - k_y(x, 0) & \text{if} \quad B(x) = 1 \end{cases}$$

Therefore D calculates $k_y(x, 0)$ and counts the number of x_i such that $k_y(x, 0) = k_y(x_i, B(x_i))$ and the number of x_i such that $k_y(x, 0) = s - k_y(x_i, B(x_i))$. The majority of these x_i are used as witnesses. D outputs 0 if the majority of the x_i give $k_y(x, 0) = k_y(x_i, B(x_i))$, or 1 if there is a majority for $s - k_y(x, 0) = k_y(x_i, B(x_i))$. If the number of witnesses for $B = 0$ equals that for $B = 1$, D makes a fair guess. In the example of figure 4, there are 4 witnesses for "$B = 0$" and 6 witnesses for "$B = 1$". D would return $B = 1$ in this example.

3.3 Probability for $\mathcal{D} = \mathcal{B}$

Let $m \in M$ still be fixed. It must be shown that the probability of of D's success is lower bounded by

$$\text{Prob}\,(x \in X : D(x) = B(x)) \geq \frac{1}{2} + \eta - \nu$$

This probability will be bounded for each of D's cases in the current section. In section 3.5 the probabilities will be combined using the total probability theorem.

Case 1 of $\mathcal{D}$

In the second subcase of D's case one $k_y(x_0, B(x_0)) - \frac{s}{2} \geq s\eta$ holds, and therefore

$$\frac{1}{2} + \eta \quad \leq \quad \text{Prob}\,(x \in \{y_1, \ldots, y_s\} : C(x_0, x) = B(x_0) \oplus B(x))$$

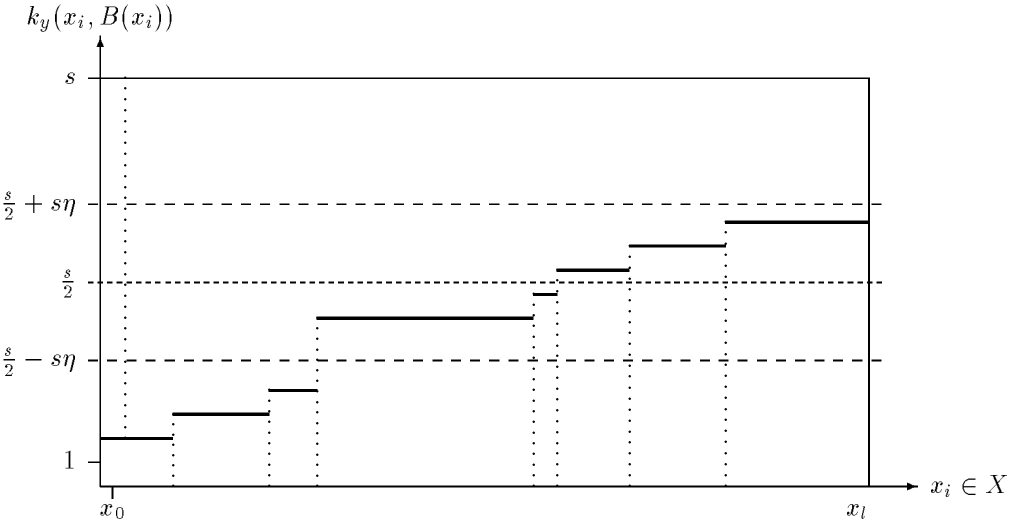

Figure 2: Subcase 1 of case 1

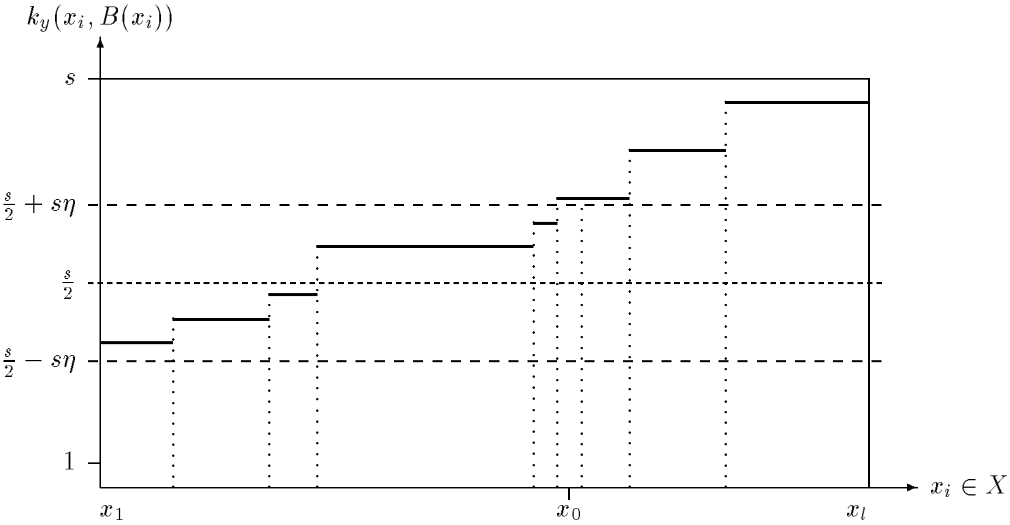

Figure 3: Subcase 2 of case 1

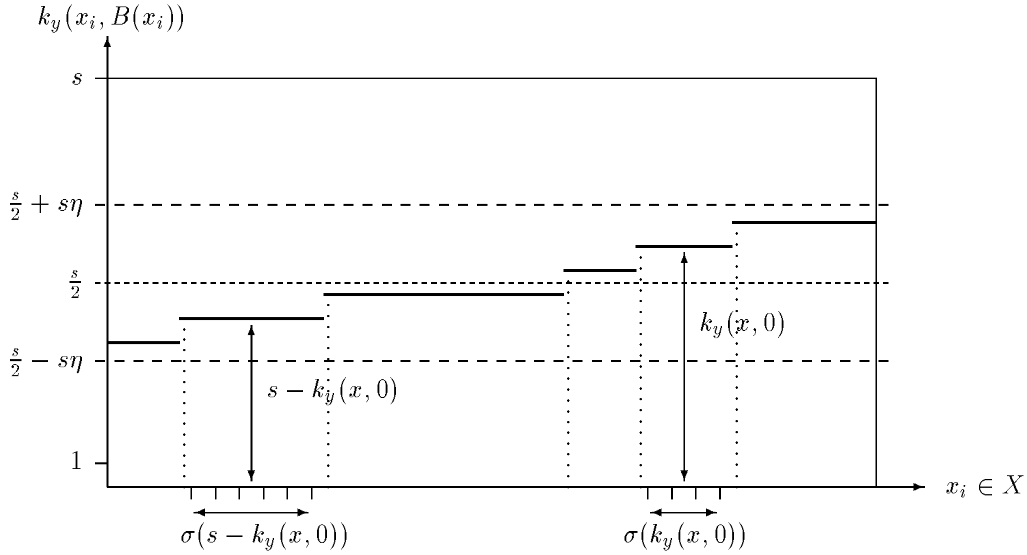

Figure 4: Case 2 (in this example the output would be $D(x) = 1$)

$$\leq \quad \text{Prob}\,(x \in X : C(x_0, x) = B(x_0) \oplus B(x)) + \frac{\nu}{2} \tag{1}$$

holds with probability at least $1 - \nu^2/2$. The latter probability is calculated with respect to the choice of the sample in y-direction and will be proved in the section 3.4. It follows that with at least the same probability

$$\begin{aligned} \frac{1}{2} + \eta - \frac{\nu}{2} &\leq \text{Prob}\,(x \in X : C(x_0, x) = B(x_0) \oplus B(x)) \\ &\overset{D}{=} \text{Prob}\,(x \in X : D(x) = B(x)) \end{aligned} \tag{2}$$

holds.

In the first subcase of case one $k_y(x_0, B(x_0)) \leq \frac{s}{2} - s\eta$ is valid, and therefore

$$\begin{aligned} \frac{1}{2} + \eta &\leq 1 - \frac{k_y(x_0, B(x_0))}{s} \\ &= \text{Prob}\,(x \in \{y_1, \ldots, y_s\} : C(x_0, x) = B(x_0) \oplus B(x) \oplus 1) \\ &\leq \text{Prob}\,(x \in X : C(x_0, x) = B(x_0) \oplus B(x) \oplus 1) + \frac{\nu}{2} \end{aligned} \tag{3}$$

holds with probability at least $1 - \nu^2/2$ (with respect to the choice of the y-sample, prove see below). This means at least the same probability for

$$\begin{aligned} \frac{1}{2} + \eta - \frac{\nu}{2} &\leq \text{Prob}\,(x \in X : C(x_0, x) = B(x_0) \oplus B(x) \oplus 1) \\ &\overset{D}{=} \text{Prob}\,(x \in X : D(x) = B(x)) \end{aligned} \tag{4}$$

Case 2 of $\mathcal{D}$

In the second case of D the probability of D's success is calculated in three steps.

In the first step the probability for $D = B$ is split into probabilities that can be treated seperately later on. These probabilities will be approximated by the success probabilities for $D = B$ applied to the members of the sample. Here the cases $B(x) = 0$ and $B(x) = 1$ will be distinguished. By the definition of D and $k_y(x,1) = s - k_y(x,0)$ can be obtained that

$$\begin{aligned}
&\mathrm{Prob}\,(x \in X : D(x) = B(x)) \\
&= \mathrm{Prob}\,(x \in X : D(x) = B(x) \wedge B(x) = 0) + \mathrm{Prob}\,(x \in X : D(x) = B(x) \wedge B(x) = 1) \\
&= \mathrm{Prob}\,(x \in X : D(x) = 0 \wedge B(x) = 0) + \mathrm{Prob}\,(x \in X : D(x) = 1 \wedge B(x) = 1) \\
&= \mathrm{Prob}\,(x \in X : \sigma(k_y(x,B(x))) > \sigma(s - k_y(x,B(x))) \wedge B(x) = 0) \\
&+ \mathrm{Prob}\,(x \in X : \sigma(k_y(x,B(x))) > \sigma(s - k_y(x,B(x))) \wedge B(x) = 1) \\
&+ \mathrm{Prob}\,(x \in X : \sigma(k_y(x,0)) = \sigma(s - k_y(x,0)) \wedge D(x) = B(x)) \\
&= \mathrm{Prob}\,(x \in X : \sigma(k_y(x,B(x))) > \sigma(s - k_y(x,B(x)))) \\
&+ \mathrm{Prob}\,(x \in X : \sigma(k_y(x,0)) = \sigma(s - k_y(x,0)) \wedge D(x) = B(x))
\end{aligned}$$

The second of these probabilities can be further manipulated. In this case, the events $D(x) = 0$ and $D(x) = 1$ each occur with probability $1/2$, independently of the precalculation and independently of the input x. Furthermore one has $k_y(x,1) = s - k_y(x,0)$ and therefore

$$\begin{aligned}
&\mathrm{Prob}\,(x \in X : \sigma(k_y(x,0)) = \sigma(s - k_y(x,0)) \wedge D(x) = B(x)) \\
&= \mathrm{Prob}\,(x \in X : \sigma(k_y(x,0)) = \sigma(s - k_y(x,0)) \wedge D(x) = 0 \wedge B(x) = 0) \\
&+ \mathrm{Prob}\,(x \in X : \sigma(k_y(x,0)) = \sigma(s - k_y(x,0)) \wedge D(x) = 1 \wedge B(x) = 1) \\
&= \frac{1}{2} \cdot \mathrm{Prob}\,(x \in X : \sigma(k_y(x,0)) = \sigma(s - k_y(x,0)) \wedge B(x) = 0) \\
&+ \frac{1}{2} \cdot \mathrm{Prob}\,(x \in X : \sigma(k_y(x,0)) = \sigma(s - k_y(x,0)) \wedge B(x) = 1) \\
&= \frac{1}{2} \cdot \mathrm{Prob}\,(x \in X : \sigma(k_y(x,0)) = \sigma(s - k_y(x,0))) \\
&= \frac{1}{2} \cdot \mathrm{Prob}\,(x \in X : \sigma(k_y(x,B(x))) = \sigma(s - k_y(x,B(x)))) \\
&= \frac{1}{2} \cdot \sum_{\substack{k=0 \\ \sigma(k)=\sigma(s-k)}}^{s} \mathrm{Prob}\,(x \in X : k_y(x,B(x)) = k)
\end{aligned}$$

This implies

$$\begin{aligned}
\mathrm{Prob}\,(x \in X : D(x) = B(x)) &= \sum_{\substack{k=0 \\ \sigma(k)>\sigma(s-k)}}^{s} \mathrm{Prob}\,(x \in X : k_y(x,B(x)) = k) \\
&+ \frac{1}{2} \cdot \sum_{\substack{k=0 \\ \sigma(k)=\sigma(s-k)}}^{s} \mathrm{Prob}\,(x \in X : k_y(x,B(x)) = k)
\end{aligned}$$

These probabilities are approximated by the sample. Because of the weak law of large numbers the probabilities in both sums fulfill

$$\mathrm{Prob}\,(x \in X : k_y(x,B(x)) = k) \geq \left(\frac{\sigma(k)}{l} - \frac{1}{4s^2}\right) \tag{5}$$

with probability at least $1-2\nu^4$, as will be proved in section 3.4.

The second step relates the success of D on the sample to the success of C on the sample. The success of D on the sample corresponds to the area below the graph in figure 4, which shall be called $\Omega_{l,s}$. Because s is odd and $\max\{\alpha,\beta\} = 1/2(\alpha+\beta+|\alpha-\beta|)$ and $\sum_{k=0}^{s}\sigma(k) = l$ hold, the probability is larger than $\geq 1-2\nu^4$ that

$$\begin{aligned}
&\mathrm{Prob}\,(x \in X : D(x) = B(x)) \\
\geq\ & \sum_{\substack{k=0\\ \sigma(k)>\sigma(s-k)}}^{s} \left(\frac{\sigma(k)}{l} - \frac{1}{4s^2}\right) + \frac{1}{2}\cdot \sum_{\substack{k=0\\ \sigma(k)=\sigma(s-k)}}^{s} \left(\frac{\sigma(k)}{l} - \frac{1}{4s^2}\right) \\
=\ & \sum_{\substack{k>s/2\\ \sigma(k)\neq\sigma(s-k)}} \left(\frac{\max\{\sigma(k),\sigma(s-k)\}}{l} - \frac{1}{4s^2}\right) + \frac{1}{2}\cdot \sum_{\substack{k=0\\ \sigma(k)=\sigma(s-k)}}^{s} \left(\frac{\sigma(k)}{l} - \frac{1}{4s^2}\right) \\
=\ & \sum_{\substack{k>s/2\\ \sigma(k)\neq\sigma(s-k)}} \left(\frac{\frac{1}{2}(\sigma(k)+\sigma(s-k)+|\sigma(k)-\sigma(s-k)|)}{l} - \frac{1}{4s^2}\right) \\
& +\frac{1}{2}\cdot \sum_{\substack{k=0\\ \sigma(k)=\sigma(s-k)}}^{s} \left(\frac{\sigma(k)}{l} - \frac{1}{4s^2}\right) \\
\geq\ & \frac{1}{2}\cdot\sum_{k=0}^{s}\frac{\sigma(k)}{l} + \frac{1}{2}\cdot \sum_{\substack{k=0\\ k>s/2}}^{s} \left|\frac{\sigma(k)-\sigma(s-k)}{l}\right| - \frac{s+1}{4s^2} \\
\geq\ & \frac{1}{2} + \frac{1}{2}\cdot \sum_{\substack{k=0\\ k>s/2}}^{s} \left|\frac{\sigma(k)-\sigma(s-k)}{l}\right| - \frac{1}{2s}
\end{aligned} \tag{6}$$

holds.

The value $\Omega_{l,s}$ can easily be calculated from figure 4 because the area below the graph does not change by the permutation of values in x-direction. Using the symmetry of the height of the steps with respect to the value $s/2$ and the second case ($|k-s/2| < s\eta$ and $k < s\eta + s/2$), one obtains

$$\begin{aligned}
\Omega_{l,s} &= \frac{sl}{2} + \sum_{\substack{k=0\\ s/2<k<s\eta+s/2}}^{s} \left(k-\frac{s}{2}\right)(\sigma(k)-\sigma(s-k)) \\
&< \frac{sl}{2} + \sum_{\substack{k=0\\ s/2<k<s\eta+s/2}}^{s} s\eta\,|\sigma(k)-\sigma(s-k)|
\end{aligned}$$

Further dividing by ls:

$$\begin{aligned}
\frac{1}{2}+\eta \sum_{\substack{k=0\\ s/2<k<s\eta+s/2}}^{s} \left|\frac{\sigma(k)-\sigma(s-k)}{l}\right| &> \frac{\Omega_{l,s}}{ls} \\
\Longrightarrow \sum_{\substack{k=0\\ s/2<k<s\eta+s/2}}^{s} \left|\frac{\sigma(k)-\sigma(s-k)}{l}\right| &> \frac{1}{\eta}\cdot\left(\frac{\Omega_{l,s}}{ls}-\frac{1}{2}\right)
\end{aligned}$$

Combined with (6) this yields

$$\text{Prob}\,(x \in X : D(x) = B(x)) > \frac{1}{2} + \frac{1}{2\eta} \cdot \left(\frac{\Omega_{l,s}}{ls} - \frac{1}{2}\right) - \frac{1}{2s} \tag{7}$$

Step three employs the fact that $\Omega_{l,s}/ls$ approximates the probability for $C(.,.) = B(.) \oplus B(.)$ with $l, s \longrightarrow t$. However this probability is lower bounded by the assumption of the theorem. Because of the weak law of large numbers the probability is at least $1 - \nu^8/16$ that

$$\frac{\Omega_{l,s}}{ls} \geq \text{Prob}\Big((x, y) \in X^2 : C(x, y) = B(x) \oplus B(y)\Big) - \frac{1}{s^2} \tag{8}$$

This approximation will be proved in section 3.4. The assumption of the theorem gives

$$\text{Prob}\Big((x, y) \in X^2 : C(x, y) = B(x) \oplus B(y)\Big) - \frac{1}{s^2} \geq \frac{1}{2} + \varepsilon - \frac{1}{s^2}$$

Using (8) and (7) one has

$$\text{Prob}\,(x \in X : D(x) = B(x)) \geq \frac{1}{2} + \frac{1}{2\eta} \cdot \left(\varepsilon - \frac{1}{s^2}\right) - \frac{1}{2s} \geq \frac{1}{2} + \eta - \frac{\nu}{2} \tag{9}$$

with probability at least

$$1 - \frac{33\nu^4}{16} \tag{10}$$

This is because the left inequality of (9) can only be hurt if (6) or (8) are wrong, which can occur with probability at most (consider $0 < \nu < 1/4$)

$$2\nu^4 + \frac{\nu^8}{16} = \frac{32\nu^4 + \nu^8}{16} \leq \frac{33\nu^4}{16}$$

The right inequality of (9) mainly is due to the definition of s and l. Look at $\eta + \nu$ more precisely:

$$\begin{aligned}
\sqrt{\frac{\varepsilon}{2}} &= \eta + \nu \\
\Longrightarrow \varepsilon &= 2\,(\eta + \nu)^2 = 2\eta^2 + 4\eta\nu + 2\nu^2 \\
&\geq 2\eta^2 + 4\eta\frac{\nu^4}{2} + \frac{\nu^8}{4} \overset{(11)}{\geq} 2\eta^2 + \frac{2\eta}{2s} + \frac{1}{s^2} \\
\Longrightarrow \frac{\varepsilon - 1/s^2}{2\eta} &\geq \eta + \frac{1}{2s} \\
\Longrightarrow \frac{\varepsilon - 1/s^2}{2\eta} - \frac{1}{2s} &\geq \eta \geq \eta - \frac{\nu}{2}
\end{aligned}$$

(2, 4, 9) complete the proof in section 3.5. Before, the approximations (1, 3, 5, 8) will be proved now.

3.4 Extent of the Approximation by the Random Sample

The calculation of the extent of the approximation of D's success probability on the set X by its success probability on the random sample uses Bernoulli's weak law of large numbers [2] several times. Given a random Bernoulli variable with expectancy p and empirical mean F_n after n trials, the probability for a distance of more than $\theta > 0$ between empirical mean and its limes is bounded above by

$$\mathrm{Prob}\left(|F_n - p| \geq \theta\right) \leq \frac{1}{4n\theta^2}$$

This instance of the weak law of large numbers is proved with Chebychev's inequality, considering the case of Bernoulli variables and hence the variance $p(1-p)$ being no larger than $1/4$.

Furthermore the following inequalities will be used that can be verified by the definition of s:

$$\frac{1}{s} = \frac{1}{2\left\lceil\frac{1}{\nu^2}\right\rceil^2 + 1} \leq \frac{1}{2\left(\frac{1}{\nu^4} + 1\right)} \leq \frac{\nu^4}{2} \qquad \text{und} \qquad \frac{1}{s^2} \leq \frac{\nu^8}{4} \tag{11}$$

Case 1 of $\mathcal{D}$

In case one inequalities (1) and (3) made the step from the probability on the random sample to the probability on X at the cost of $\frac{\nu}{2} = \frac{\delta}{4}\sqrt{\frac{\varepsilon}{2}}$.

(1) is obtained be the weak law of large numbers with parameters $n := s$ and $\theta := \nu/2$

$$\mathrm{Prob}\Big(\Big|\mathrm{Prob}(x \in \{y_1, \ldots, y_s\} : C(x_0, x) = B(x_0) \oplus B(x))$$
$$-\mathrm{Prob}(x \in X : C(x_0, x) = B(x_0) \oplus B(x))\Big| \geq \frac{\nu}{2}\Big) \leq \frac{1}{4s\theta^2} = \frac{1}{s\nu^2} \overset{(11)}{\leq} \frac{\nu^2}{2}$$

The complement of this yields the probability of $1 - \frac{\nu^2}{2}$ in (1). In the same way (3) in the other subcase of case one is proved.

Case 2 of $\mathcal{D}$

For the second case and (5) the y-sample $\{y_1 \ldots, y_s\}$ and a $k \in \{0 \ldots, s\}$ are fixed. At the l experiments with $x_1 \ldots, x_l$ one obtained the approximation $\sigma(k)/l$ for $\mathrm{Prob}\left(x \in X : k_y(x, B(x)) = k\right)$. Let now $\theta := 1/4s^2$ and $n := l$, and consequently

$$\mathrm{Prob}\Big(|\sigma(k)/l - \mathrm{Prob}\left(x \in X : k_y(x, B(x)) = k\right)| > 1/4s^2\Big) \leq \frac{1}{4l\theta^2} = \frac{16s^4}{4s^5} = \frac{4}{s} \overset{(11)}{\leq} 2\nu^4$$

therefore

$$\mathrm{Prob}\Big(\mathrm{Prob}\left(x \in X : k_y(x, B(x)) = k\right) \geq \sigma(k)/l - 1/4s^2\Big) \geq 1 - 2\nu^4$$

In (8) the expectancy of $\Omega_{l,s}/ls$ for $l, s \longrightarrow t$ is known:

$$\frac{\Omega_{l,s}}{ls} \longrightarrow \Omega := \text{Prob}\left((x,y) \in X^2 : C(x,y) = B(x) \oplus B(y)\right)$$

Using the weak law of large numbers with parameters $\theta := 1/s^2$ and ls points in the sample

$$\text{Prob}\left(|\Omega_{ls}/ls - \Omega| > 1/s^2\right) \leq \frac{1}{4ls\theta^2} = \frac{1}{4s^2} \overset{(11)}{\leq} \frac{\nu^8}{16}$$

which means that (8) holds with probability at least $1 - \frac{\nu^8}{16}$.

3.5 Final Combination of Results

The results of the two sections before combine to the assertion.

For finite sets A and B, the event C, a ρ with $0 \leq \rho \leq 1$, and

$$\lambda := \text{Prob}\,(a \in A : \text{Prob}\,(b \in B : C(a,b)) \geq \rho)$$

one has

$$\lambda + \rho\,(1-\lambda) \geq \text{Prob}\,((a,b) \in A \times B : C(a,b)) \geq \rho\lambda \tag{12}$$

The success probability of D can be simplified. Let $x \in X$ be an input to D and $z := (x_1, \ldots, x_l, y_1, \ldots, y_s) \in X^{l+s}$ a sample chosen by D. Consider the events

$P(x,z)$: D chooses sample z and $D(x) = B(x)$,
$Q_1(z)$: D chooses sample z and the first case happens,
$Q_2(z)$: D chooses sample z and the second case happens.

Using this notation and $\rho := 1/2 + \eta - \nu/2$ (2, 3, 9, 10) translate to:

$$\begin{aligned}
\text{Prob}\,(z : \text{Prob}\,(x : P(z,x)) \geq \rho \mid Q_1(z)) &\geq 1 - \frac{\nu^2}{2} \\
\text{Prob}\,(z : \text{Prob}\,(x : P(z,x)) \geq \rho \mid Q_2(z)) &\geq 1 - \frac{33\nu^4}{16}
\end{aligned}$$

The right inequality in (12) and the total probability theorem yield

$$\begin{aligned}
\text{Prob}\,(z,x : P(z,x)) &\geq \rho \cdot \text{Prob}\,(z : \text{Prob}\,(x : P(z,x)) \geq \rho) \\
&= \rho \cdot \Big(\text{Prob}\,(z : \text{Prob}\,(x : P(z,x)) \geq \rho \mid Q_1(z))\,\text{Prob}\,(z : Q_1(z)) \\
&\qquad + \text{Prob}\,(z : \text{Prob}\,(x : P(z,x)) \geq \rho \mid Q_2(z))\,\text{Prob}\,(z : Q_2(z))\Big) \\
&\geq \rho \cdot \min\left(1 - \frac{\nu^2}{2}, 1 - \frac{33\nu^4}{16}\right) \cdot \Big(\text{Prob}\,(z : Q_1(z)) + \text{Prob}\,(z : Q_2(z))\Big) \\
&= \left(\frac{1}{2} + \eta - \frac{\nu}{2}\right) \cdot \left(1 - \frac{\nu^2}{2}\right)
\end{aligned}$$

$$\begin{aligned} &\geq \left(\frac{1}{2} + \eta - \frac{\nu}{2}\right) - \frac{\nu^2}{2} \\ &\geq \frac{1}{2} + \eta - \nu \end{aligned}$$

Here $\nu < 1/4$ and consequently

$$\frac{\nu}{2} \geq \frac{\nu^2}{2} \geq \frac{33\nu^4}{16}$$

was employed ◇

4 Conclusion

The current paper contributes to earlier work of Yao and Kranakis [3, 2]. Although no doubt of the validity of Yao's statement was justified, it is good to see the ideas working in a consistent, full and simple proof.

Acknowledgement

Substantial thanks for continuous support and a great environment are due to Rainer Schrader, Georg Wambach and Andreas Nolte from Cologne University.

References

[1] F.Damm, *Konstruktion und Analyse beweisbar sicherer elektronischer Unterschriftenverfahren* (in German), (Dissertation, Universität zu Köln, 1994; published by Verlag Shaker, Aachen, Germany, 1995).

[2] E.Kranakis, *Primality and Cryptography*, (Wiley-Teubner Series in Computer Science, 1986).

[3] A.C.Yao, Theory and Applications of Trapdoor Functions, *Proceedings of 23rd Annual IEEE Symposium on Foundations of Computer Science*, (1982) 80–91.