# Post-processing functions for a biased physical random number generator

Patrick Lacharme

Université de Toulon,
Institut de Mathématique (Imath)

Fast Software Encryption 2008

# Overview

# Overview

# True random number generator

A true random number generator consists of two different parts :

- A physical non deterministic phenomenon produces a raw binary sequence.
- A deterministic function, called corrector, compress this sequence in order to reduce statistical weakness.

# Statistical model for the raw sequence

The bias $e$ is the deviation from 1/2 of the probability of occurence of a bit $x_i$ :

$$e = |P(x_i = 0) - 1/2| = |P(x_i = 1) - 1/2| \ .$$

**Hypothesis :** the bits $x_i$ of the raw sequence are independents and have a constant bias $e$.

A corrector is a function mapping a vector $x$ of $n$ bits to a vector $y$ of $m$ bits, with compression rate $m/n$.

# Output biais of a Boolean function

Let $e$ be the bias of the $n$ input $x_i$.

- For a **Boolean function** $f$ mapping $\mathbf{F}_2^n$ to $\mathbf{F}_2$, the output bias $\Delta_f$ of $f$ is

$$\Delta_f(e) = |P(f(x) = 1) - 1/2| = |P(f(x) = 0) - 1/2| \ . \quad (1)$$

- For a **vectorial function** $f$ mapping $n$ bits to $m$ bits, the output bias of a linear combination of $f_i$ $\sum_{i=1}^{m} u_i f_i(x) = u.f(x)$ is :

$$\Delta_{u.f}(e) = |P(u.f(x) = 1) - 1/2| = |P(u.f(x) = 0) - 1/2| \ .$$

# First formula of output bias of a Boolean function

From (1), the output bias $\Delta_f(e)$ is a polynomial in $e$ :

$$\Delta_f(e) = -\frac{1}{2}\sum_{x\in\mathbf{F}_2^n}(\frac{1}{2}-e)^{n-w_h(x)}(\frac{1}{2}+e)^{w_h(x)}(-1)^{f(x)} \qquad (2)$$

$$= a_0 + a_t e^t + a_{t+1}e^{t+1} + \ldots$$

Moreover, $\Delta_f(e)$ have no constant term if and only if $f$ is balanced.

Construction of functions which maximalise the valuation $t$ ?

# Three linear correctors

**M. Dichtl (FSE'07) :** *Bad and Good ways of post-processing biased physical random numbers*.

Three linear correctors mapping 16 bits to 8 bits :

1. $y_i = x_i + x_{i+1 \bmod 8} + x_{i+8} \bmod 2$.
2. $y_i = x_i + x_{i+1 \bmod 8} + x_{i+2 \bmod 8} + x_{i+8} \bmod 2$.
3. $y_i = x_i + x_{i+1 \bmod 8} + x_{i+2 \bmod 8} + x_{i+4 \bmod 8} + x_{i+8} \bmod 2$.

Same compression rate as $\mathrm{xor}$ corrector :

$$y_i = x_{2i} + x_{2i+1} \bmod 2 \ .$$

# Analysis of the bias

These correctors are designed to reduce the output bias.

Same hypothesis on input bias.

- **Approach :** determine probability of every inputs and sum up the probability of occurence leading the same output.
- **Results :** bias of any output bytes is a polynomial in $e$ and the lowest power in $e$ is respectly 3, 4 and 5.

Systematic construction of corrector with variable input sizes and compression rates ?

# Overview

# Matricial representation of a linear corrector

1. A linear corrector $f$ mapping $\mathbf{F}_2^n$ to $\mathbf{F}_2^m$ is defined by the matricial product :

$$f(x) = \begin{pmatrix} h_{1,1} \dots h_{1,n} \\ \vdots \\ h_{m,1} \dots h_{m,n} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_m \end{pmatrix},$$

   where $H = (h_{i,j})$ is a binary matrix with $m$ rows and $n$ collums.
2. A linear corrector is associated with a $[n, m]$ linear code.
3. It corresponds to a syndrom calculation.

# Minimal distance and bias

Let $f$ be the linear corrector represented by the matrix $H$ generating a $[n, m, d]$ linear code, and $e/2$ the input bias.

**Theorem**

*The output bias $\Delta_{u.f}(e)$ of $u.f(x)$ is less or equal than $e^d/2$.*

**Sketch of proof :**

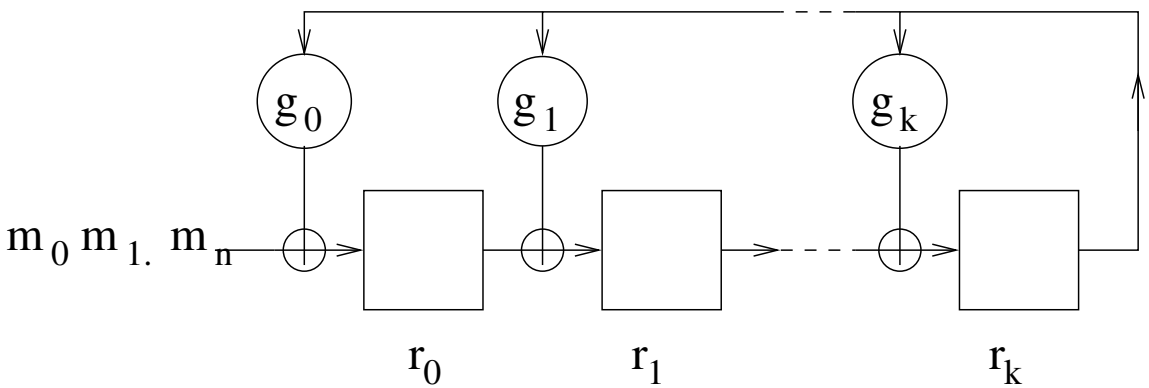
- The bias of $x_{i_1} + \ldots + x_{i_d} \bmod 2$ is $e^d/2$.
- Any linear combination of output bits is the sum of at least $d$ input bits, by definition of minimal distance of a linear code.

# Implementation with cyclic codes

The syndrom calculation is realized by a polynomial division

$$\sum_{i=0}^{n} m_i X^i \bmod \sum_{i=0}^{k} g_i X^i,$$

which is efficiently implemented with a shift register.



# Conclusion on linear corrector

Equivalence between minimal distance of a linear code and valuation of the bias of linear combination of output bits.

For example, the three correctors of M. Dichtl correspond respectly to generator matrix of [16,8,3], [16,8,4] and [16,8,5] linear codes.

Moreover, all this part can be generalized with non constant input bias using the $d$ greatest input bias.

Non linear corrector can be better than linear corrector ?

# Overview

# Fourier and Walsh Transform

## Definition

*The Fourier transform of a function $f$ with $n$ variables is :*

$$F_f(u) = \sum_{x \in \mathbf{F}_2^n} f(x)(-1)^{x.u}.$$

*The Walsh transform of a function $f$ with $n$ variables is :*

$$\widehat{f}(u) = \sum_{x \in \mathbf{F}_2^n} (-1)^{f(x) \oplus x.u}.$$

# Output bias with Walsh coefficients

## Theorem

*The output bias* $\Delta_{\phi_u}(e)$ *of* $\phi_u(x) = u.f(x)$ *is*

$$\Delta_{\phi_u}(e) = \frac{1}{2^{n+1}} \sum_{v \in \mathbf{F}_2^n} (2e)^{w_h(v)} (-1)^{w_h(v)+1} \widehat{\phi_u}(v). \qquad (3)$$

**Sketch of proof :** From formula (2) of the bias, we analyse the Fourier transform of the function

$$g(x) = (\frac{1}{2} - e)^{n-w_h(x)} (\frac{1}{2} + e)^{w_h(x)} .$$

# Example (1)

Let $f$ be the Boolean function defined by

$$f(x) = f(x_1, x_2, x_3) = x_2 + x_3 + x_1x_2 + x_2x_3 \bmod 2 \ ,$$

where the truth table and the Walsh coefficients are

| $x$ | $f(x)$ | $\widehat{f}(x)$ |
|---|---|---|
| 000 | 0 | 0 |
| 001 | 1 | 4 |
| 010 | 1 | 0 |
| 100 | 0 | -4 |
| 011 | 1 | 4 |
| 101 | 1 | 0 |
| 110 | 0 | 4 |
| 111 | 0 | 0 |

# Example (2)

The probability $P(f(x)=0)=\frac{1}{2}-e$ computed with formula (2) :

$$P(f(x)=0)=(\frac{1}{2}-e)^3+(\frac{1}{2}-e)^2(\frac{1}{2}+e)+(\frac{1}{2}-e)(\frac{1}{2}+e)^2+(\frac{1}{2}+e)^3$$

$$=\frac{1}{2}+2e^2\ .$$

The output bias computed with formula (3), with $u=1$, is

$$\Delta_f(e)=\frac{1}{16}(\hat{f}(000)+2e\hat{f}(001)+2e\hat{f}(010)+2e\hat{f}(100)$$

$$-4e^2\hat{f}(011)-4e^2\hat{f}(101)-4e^2\hat{f}(110)+8e^3\hat{f}(111))$$

$$=-2e^2\ .$$

# Valuation of bias

Coefficients and valuation of $\Delta_{\phi_u}(e)$ are determined with formula (3) :

For a Boolean function $\phi_u$, we denote

$$B_w = \sum_{\substack{v \in \mathbf{F}_2^n \\ w_h(v)=w}} \widehat{\phi_u}(v).$$

## Corollary

*If $\phi_u$ is balanced, then the output bias $\Delta_{\phi_u}(e)$ is a polynomial of valuation $W$, with $W = \min\{w \mid B_w \neq 0\}$.*

Systematic constructions of functions with high $W$ ?

# Overview

# Resilient functions

## Definition ($(n, m, t)$-resilient functions)

*A $(n, m, t)$-resilient function is a vectorial function from $\mathbf{F}_2^n$ to $\mathbf{F}_2^m$, such that for all $y \in \mathbf{F}_2^m$ and for any binary constant $c_i$ :*

$$P(f(x) = y \mid x_{i_1} = c_1, \ldots, x_{i_t} = c_t) = 2^{n-m} ,$$

*where all $x_i$, with $i \notin \{i_1, \ldots, i_t\}$ are viewed as independent binary random variables with probability 0.5.*

## Lemma (Xiao, Massey,1988)

*A function $f$ is $t$-resilient if and only if Walsh coefficients $\widehat{f}(u) = 0$, with $0 \leq w_h(u) \leq t$.*

# A resilient corrector

**Theorem**

*Let $f$ be a $(n, m, t)$-resilient function. Then for all $u \neq 0$, the output bias $\Delta_{u.f}(e)$ is a polynomial with valuation greater than $t + 1$.*

**Sketch of proof :** By previous Lemma and formula (3),

$$\Delta_{\phi_u}(e) = \frac{1}{2^{n+1}} \sum_{\substack{v \in \mathbf{F}_2^n \\ w_h(v) > t}} (2e)^{w_h(v)} (-1)^{w_h(v)+1} \widehat{\phi_u}(v).$$

# Resilience and bias

Non linear corrector are sometimes better than linear corrector : there exist a non linear (16,8,5) resilient function.

The bias of any linear combination of output bits is bounded in linear and non linear case using resilience degree.

We want an upper bound on the bias of any output $m$-tuple $y$ :

$$\left|P(f(x) = y) - 2^{-m}\right|$$

# Bias of any output $m$-tuple

**Theorem**

$$\forall y \in \{0,1\}^m \ \left|P(f(x)=y) - 2^{-m}\right| \leq 2 \max_{u \in \mathbf{F}_2^n} |\Delta_{\phi_u}(e)| \ .$$

**Sketch of proof :** Variant of a Theorem of Alon, Goldreich, Hastad, Peralta, 1992 on *biased sample space* and *almost $k$-wise independent random variables.*

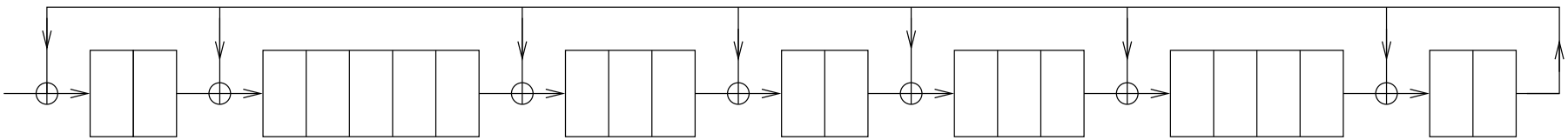
If $e^{t+1} << 2^{-m}$, then the minimal entropy of the output is very close to $m$.

# Example of linear corrector

- Let $C$ be a $[255, 21, 111]$ BCH code and $D$ the dual code of $C$ with parameters $[255, 234, 6]$, with generator polynomial

$$H(X) = X^{21} + X^{19} + X^{14} + X^{10} + X^7 + X^2 + 1.$$

- The linear corrector $f : \mathbf{F}_2^{255} \to \mathbf{F}_2^{21}$ is implemented with a shift register of length 21 with seven $\mathrm{xor}$ logic doors.



With an input bias of 0.25,

$$\forall y \in \mathbf{F}_2^{21} \quad \left| P(f(X) = y) - 2^{-21} \right| \leq 2^{-111}.$$

# Conclusion

Linear codes and resilient functions give construction of correctors reducing the bias with variable input sizes and compression rates.

Constant input bias assumption can be removed in the linear case.

Hardware implementation of post processing functions can be realized on a small component.