Evaluation and monitoring of free running oscillators serving as source of randomness

Elie Noumon Allini • Maciej Skórski • **Oto Petura** • Florent Bernard • Marek Laban • Viktor Fischer

CHES 2018, Amsterdam, September 2018

Introduction

Jittery clock – commonly used source of randomness in digital devices

- Clock jitter caused by several noise sources
  - White noise (thermal noise, ...)
    - ↪ Best source of randomness, non manipulable
  - Autocorrelated noise (low frequency noises, e.g. flicker noise)
    - ↪ Entropy rate (unpredictability measure) difficult to quantify
  - Data dependent noise
    - ↪ Dangerous (manipulable), must be avoided

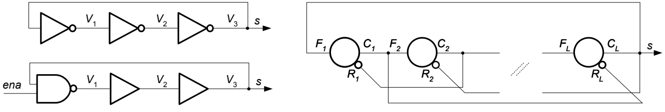
Jitter monitoring

- Continuous embedded monitoring is preferable
- Jitter – usually quantified using the variance

$$\mathrm{var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2 \quad (1)$$

Introduction

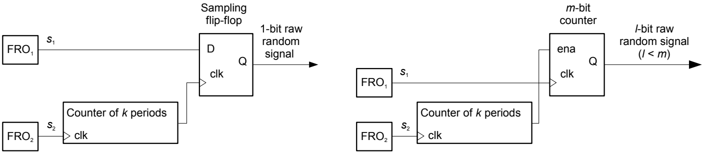
Free running oscillators – sources of the jittery clocks



Ring oscillators (RO)

Self-timed ring (STR) based on Müller gates

Randomness extraction methods from jittery clocks



Sampler based randomness extraction

Counter based randomness extraction

Introduction

Objectives

- Analyze the use of variance for entropy estimation
- Use high order Markov model to estimate entropy coming from auto-correlated noises
- Compare performance of ROs and STRs as sources of randomness

1. Variance and Allan variance
2. High order Markov model for entropy rate estimation from autocorrelated signals
3. Experimental results

Characterization of random fluctuations of the clock frequency

Power spectral density (PSD)

- Defined as:

$$S_y(f) = h_\alpha f^\alpha \tag{2}$$

  - y – dimensionless fractional frequency ($y = (\nu - \nu_0)/\nu_0$)
  - α – constant characterizing the noise process
  - h_α – intensity of this noise
- Characterizes random fluctuations of the clock frequency

α	**Type of the noise process**
-2	Random Walk Frequency (RWF)
-1	Flicker Noise Frequency (FF)
0	White Noise Frequency (WF) or Random Walk Phase (RWP)
1	Flicker Noise Phase (FP)
2	White Noise Phase (WP)

Variance of the frequency fluctuations

Main assumption

- y is an infinite zero-mean stationary process
  - characterized by its variance computed from a window of length τ

Variance can be computed using the power spectral density

- Corollary of the Wiener-Khinchin theorem
- Variance of y computed from the power spectral density $S_y(f)$:

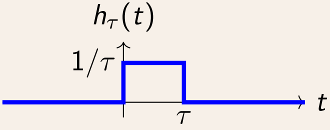
$$\sigma_y^2(\tau) = \int_0^{+\infty} S_y(f) \times |H_\tau(f)|^2 df, \tag{3}$$

whenever it exists.

  - $H_\tau(f)$ is the transfer function of the variance operator:
    - ↪ Fourier transform of the impulse response function h_τ
    - ↪ Depends on the type of variance computed

Computation of the statistical variance from the PSD

Time domain



Frequency domain

$$|H_\tau(f)|^2 = \left(\frac{\sin(\pi\tau f)}{\pi\tau f}\right)^2 \quad (4)$$

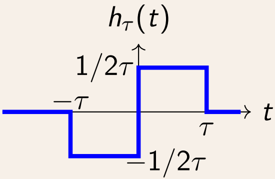
Variance of the jitter computed for $\alpha \in [-2; 2]$ from time window τ

$$\sigma_y^2(\tau) = \sum_{\alpha=-2}^{2} \frac{h_\alpha}{(\pi\tau)^2} \int_0^{f_h} f^{\alpha-2} \sin^2(\pi\tau f) df. \quad (5)$$

- Problem: if $\alpha \leqslant -1$, the integral does not converge as f tends to 0
  - The use of the statistical variance can cause entropy overestimation

Allan variance and its computation from the PSD

Time domain



Frequency domain

$$|H_\tau(f)|^2 = \left(\frac{\sin(\pi\tau f)}{\pi\tau f}\right)^2 \sin^2(\pi\tau f) \quad (6)$$

Allan Variance of the jitter computed for $\alpha \in [-2;2]$ from window τ

$$\sigma_y^2(\tau) = \sum_{\alpha=-2}^{2} \frac{2h_\alpha}{(\pi\tau)^2} \int_0^{f_h} f^{\alpha-2} \sin^4(\pi\tau f)df \quad (7)$$

- Convergence ensured for $\alpha > -3$ as f tends to 0:
  - Allan variance is accurate, even in presence of low frequency noises

Allan variance estimation from a limited data set

An average fractional frequency can be used

- Average frequency deviation $\overline{y}_k$ over a time interval of length τ
  - Corresponds to the fluctuations while counting the number of periods of the jittery signal over τ

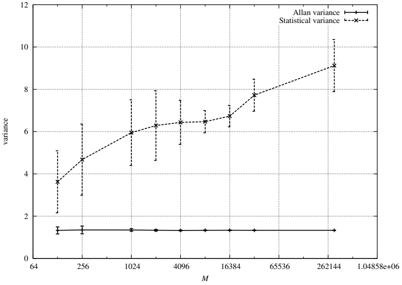
- Estimate of the Allan variance:

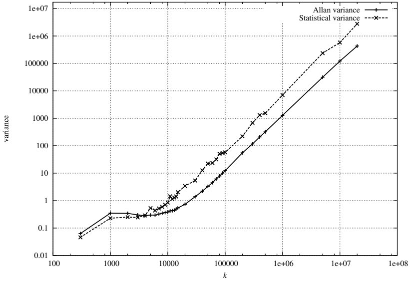
$$\sigma_y^2(\tau) = \frac{1}{2(M-1)} \sum_{i=1}^{M-1} \left(\overline{y}_{i+1} - \overline{y}_i\right)^2 . \quad (8)$$

$\hookrightarrow$ M : total number of $\overline{y}_k$'s.

- For $\alpha = 0$, $\sigma_y^2(\tau)$ is an unbiased estimator of the variance even for a finite M

Experimental results

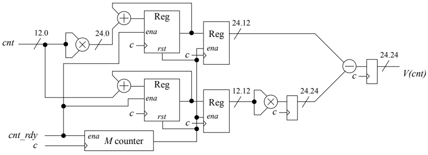




- Variance dependence on the number of samples M
  - Allan variance stable
  - Statistical variance increases with M
- Variance dependence on the jitter accumulcation period k
  - Allan variance always below statistical variance
  - Statistical variance causes entropy rate overestimation
- Similar results for both types of free running oscillators studied

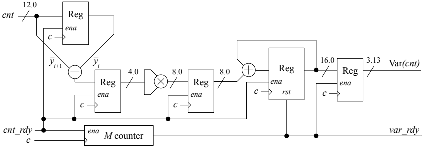
Hardware implementations

- Statistical variance



3 adders/subtractors, 2 multipliers

- Allan variance



1 adder/subtractor, 1 multiplier

Comparison with the state-of-the-art methods

Method	Area		f_{max}	Power
	ALM/Regs	DSPs	[MHz]	[mW]
Haddad *et al.*	119/160	2	178.3	6-7
Fischer and Lubicz	169/200	4	187.7	7-8
Proposed method, Eq. (8)	49/117	1	238.5	4-5

1 Variance and Allan variance

2 High order Markov model for entropy rate estimation from autocorrelated signals

3 Experimental results

The use of high order Markov chain models for entropy estimation

Min-entropy

- Min-entropy is the most conservative entropy measure
  - Avoids entropy rate overestimation
  - Hard to estimate in general
- Recent approach offers efficient way to estimate min-entropy[a]:
  - Information sources modeled as high order Markov chains

[a] S. Kamath and S. Verdu, Estimation of entropy rate and Renyi entropy rate for Markov chains, IEEE International Symposium on Information Theory 2016

Markov chain

- Convenient to model temporal short-term dependencies
  - Higher order models give more accuracy but are much more complex
- Depending on jitter properties and the randomness extraction process, we use an 8-th order Markov model to study dependencies
  - Model parameters: $\{0,1\}^8$ states, transition matrix $2^8 \times 2^8$

Entropy estimates from the 8-th order Markov chain model

Randomness extraction method: sampling the jittery clock

Jitter accumulation time	Markov chain	AIS 31 Procedure B	AIS 31 T8	NIST 800-90B	NIST 800-90B
Periods of s_2	min-entropy		Shannon entropy	IID	min-entropy
10 000	0.8102	failed	0.9844	non-IID	0.648
20 000	0.8105	failed	0.9851	non-IID	0.647
30 000	0.8102	failed	0.9847	non-IID	0.648
50 000	0.9369	failed	0.9992	non-IID	0.673
100 000	0.9012	failed	0.9935	non-IID	0.670

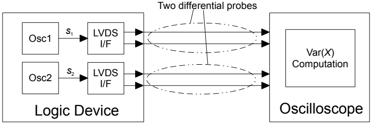
Randomness extraction method: counting the jittery clock periods

Jitter accumulation time	Markov chain	AIS 31 Procedure B	AIS 31 T8	NIST 800-90B	NIST 800-90B
Periods of s_2	min-entropy		Shannon entropy	IID	min-entropy
10 000	0.8089	failed	0.9966	non-IID	0.844
15 000	0.9769	passed	0.9998	non-IID	0.931
20 000	0.9865	passed	0.9999	IID	0.999
25 000	0.9907	passed	0.9999	IID	0.998
100 000	0.9910	passed	0.9999	IID	0.998

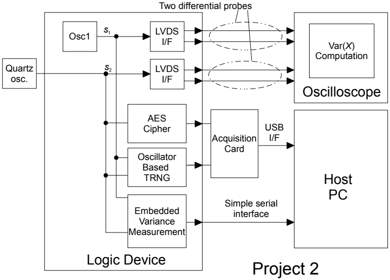
1. Variance and Allan variance
2. High order Markov model for entropy rate estimation from autocorrelated signals
3. Experimental results

Impact of the surrounding logic on the jitter and entropy rate

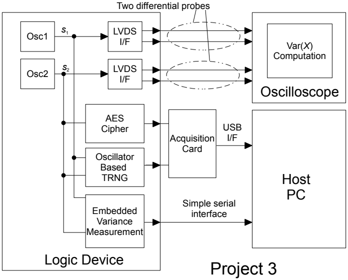
- Three projects implemented
- Blocks placed exactly on the same place in the same FPGA



Project 1



Project 2



Project 3

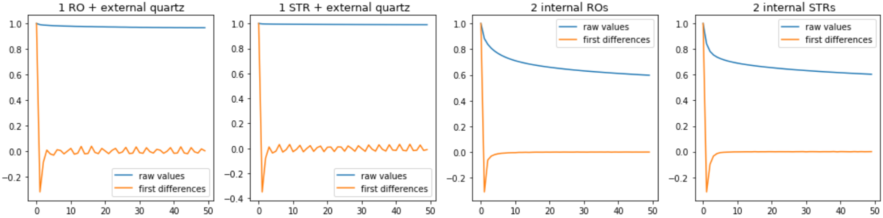
Impact of the surrounding logic on the jitter and entropy rate

Project	σ_1 [ps]	σ_2 [ps]	$Var(cnt)$	$Avar(N)$
Project 1 (just two rings)	3.9	3.3	14.01	2.79
Project 2 (ring + ext.osc. + other logic)	9.7	7.3	26.94	4.33
Project 3 (two rings + other logic)	10.6	10.0	14.72	2.76

- Oscillator jitter increases when a full cryptosystem is implemented
  - Surrounding logic has inevitable impact on clock jitters
- However, variances of counter values do not change when both oscillators are implemented inside the device!
- External clocks
  - Cause entropy rate overestimation
  - Introduce manipulable global noise sources into the generator

Comparison of RO and STR as sources of randomness

- Autocorrelation of raw counter values and their first differences
  - Two identical rings (RO or STR)
  - One ring (RO or STR) and an external quartz oscillator



- RO and STR exhibit the same behavior in terms of jitter produced
- The use of identical oscillators reduces autocorrelations
- First order difference removes large portion of autocorrelation

Conclusions

- Counting jittery clock periods gives higher quality random numbers
  - Higher bit rate with higher entropy rate
  - Counter values can be used for online jitter monitoring
- Allan variance should be used to estimate entropy rate rather than the statistical variance
  - Not sensitive to window size – impact of low frequency noises can be reduced using small windows without loosing precision
  - Smaller circuitry required for implementation
- Differential principle of the TRNG design is a stringent requirement, not a recommendation
  - Global, manipulable noises are strong and always present
- High order Markov chain models provide good min-entropy estimates and are efficient to detect dependencies in generated numbers

Acknowledgments

This work was performed in the framework of the project

HECTOR

Hardware Enabled Crypto and Randomness

The HECTOR project has received funding from the
European Union's Horizon 2020 research and innovation programme
under grant agreement number 644052 starting from March 2015

www.hector-project.eu