Coherent sampling based TRNG : a statistical and behavioral approach

Florent BERNARD, Viktor FISCHER

Cryptarchi 2009, june 24-27
Prague

Contents

Motivations

- Random numbers often employed in :
  - Key generation process,
  - Authentication protocols,
  - Padding,
  - Digital signature scheme,
  - Encryption algorithms (IV)
- Security depends greatly on the quality of the randomness source

Random Number Generation

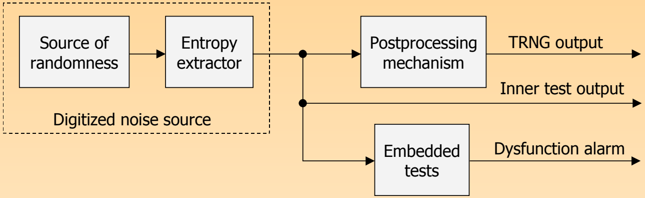


FIG.: General principle of random numbers generation

- Statistical tests needed at different levels,
- Analysis of statistical tests results must be done carefully,
- Derived conclusion from the tests about the RNG security must be done even more carefully...
- Question :
  « How can security be evaluated for random numbers generation ? »

Common ways of answering the question (1)

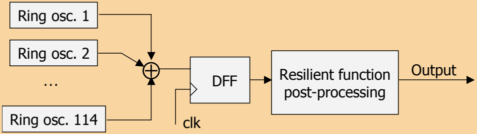
- **Usual (and quick) answer** : (T)RNG's ability to pass a battery of statistical tests : FIPS, NIST, DieHard
- Necessary but not sufficient condition

Common ways of answering the question (1)

- **Usual (and quick) answer** : (T)RNG's ability to pass a battery of statistical tests : FIPS, NIST, DieHard
- Necessary but not sufficient condition

Example

Sunar's principle with N Ring Oscillators, without any jitter :



$$Sunar_n = \left(N + \sum_{k=1}^{N} \left\lfloor \frac{\varphi_0 + n \times T_{clk} \mod T_k}{H_k} \right\rfloor \right) \mod 2$$

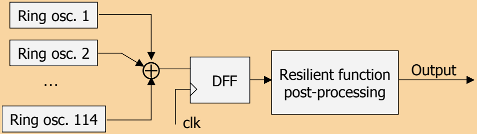
Sequence produced by this **deterministic equation** pass the FIPS 140-1 tests starting from $N \geq 17 << 114$...

Common ways of answering the question (1)

- **Usual (and quick) answer** : (T)RNG's ability to pass a battery of statistical tests : FIPS, NIST, DieHard
- Necessary but not sufficient condition

Example

Sunar's principle with N Ring Oscillators, without any jitter :



$$Sunar_n = \left(N + \sum_{k=1}^{N} \left\lfloor \frac{\varphi_0 + n \times T_{clk} \mod T_k}{H_k} \right\rfloor \right) \mod 2$$

Sequence produced by this **deterministic equation** pass the FIPS 140-1 tests starting from $N \geq 17 << 114$...

Can we conclude it is randomness ?

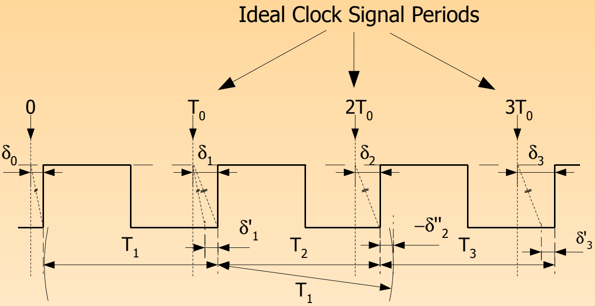
Common ways of answering the question (2)

- Problem in previous example : the produced sequence pass some statistical tests with a **null entropy**...
- The same problem appears if statistical tests are performed after post-processing (resilient function for example) - **Unchanged entropy**
- AIS 31 : « Entropy/random bit should be sufficiently large »
- Problem : Entropy **is not** a property of observed random numbers... but of random variables

Common ways of answering the question (2)

- Problem in previous example : the produced sequence pass some statistical tests with a **null entropy**...
- The same problem appears if statistical tests are performed after post-processing (resilient function for example) - **Unchanged entropy**
- AIS 31 : « Entropy/random bit should be sufficiently large »
- Problem : Entropy **is not** a property of observed random numbers... but of random variables
- **Preferable answer** : Mathematical model of the noise source is needed
- Difficulties : strong assumptions needed to have conclusion from mathematical equations...
  ... but not always easy to verify their validity in hardware

Source of randomness used in this work



- Phase jitter : $\delta_n = t_n - nT_0$
- Period jitter : $\delta'_n = (t_n - t_{n-1}) - T_0 = \delta_n - \delta_{n-1}$
- Cycle-to-cycle jitter : $\delta''_n = (t_n - t_{n-1}) - (t_{n-1} - t_{n-2}) = \delta'_n - \delta'_{n-1}$

Jitter components

- Deterministic jitter (DJ)
  - Power supply variation
  - Cross talks
  - Electro-magnetic interference
  - Simultaneous switching outputs
- Random jitter (RJ)
  - Sum of many independent contributor inherent to any electric circuits
    - Thermal vibrations : crystal structures, conductor atoms
    - **Many other** minor contributions

  Obeys the *central limit theorem* $\Rightarrow$ Gaussian probability distribution
- Difficulties to treat both jitter components in a model (not the same behaviour)
- Deterministic jitter remains always present in electronic devices but can sometimes be reduced

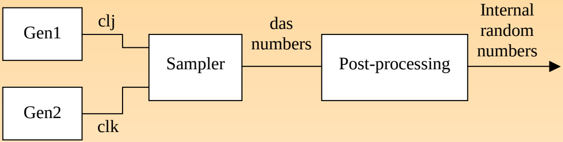
Jitter components

- Deterministic jitter (DJ)
  - Power supply variation
  - Cross talks
  - Electro-magnetic interference
  - Simultaneous switching outputs
- Random jitter (RJ)
  - Sum of many independent contributor inherent to any electric circuits
    - Thermal vibrations : crystal structures, conductor atoms
    - **Many other** minor contributions

  Obeys the *central limit theorem* $\Rightarrow$ Gaussian probability distribution
- Difficulties to treat both jitter components in a model (not the same behaviour)
- Deterministic jitter remains always present in electronic devices but can sometimes be reduced
- First approach in our model : study with the random part of the jitter only

Basic principle

- Class of RNGs based on sampling one clock signal with another

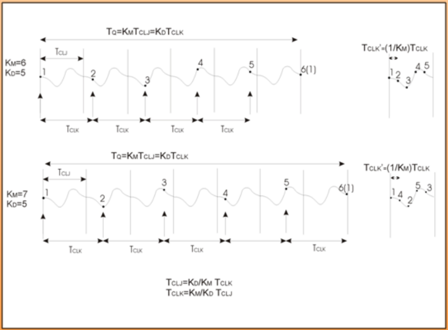


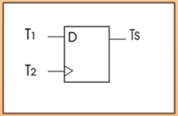
- Optional post-processing : increases statistical properties of produced sequences (not considered in this work)
- Two jittery clocks : one sampled by another → production of digitized analog signal (das) numbers

Coherent sampling

- Boyan :

Theoretical aspects (3/9)





$T_1=T_{clj}$, $T_2=T_{clk}$

$$\left(\frac{K_M}{K_D}T_{CLJ}\right)_{\mathrm{mod}(T_{CLJ})} \le \frac{T_{CLJ}}{K_D}$$

or

$$\Phi_2 \le \frac{1}{K_M}T_{CLK}$$

(1)

- Depending on K_m and K_d we can have either:
  - **consecutive equivalent sampling** - if condition (1) holds
  - **non consecutive equivalent sampling** - otherwise

8

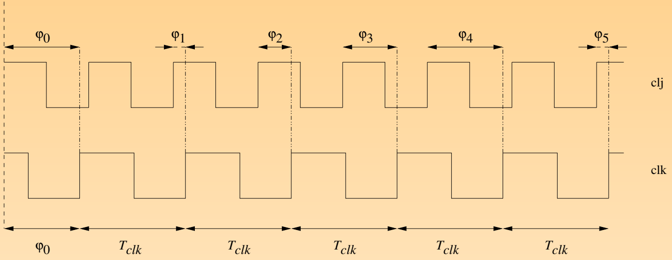
Assumptions - Strategy

- Focus on random (e.g. Gaussian) jitter only
  - The period T of one signal is considered as a random variable
  - T is supposed to follow a Gaussian distribution with mean μ and standard deviation σ :

$$T \sim \mathcal{N}(\mu, \sigma)$$

- Description of the ideal case ($\sigma = 0$) :
  - Easy !
  - Useful : corresponds to the mean behaviour
- Addition of the random jitter
- (Addition of the deterministic jitter $\longrightarrow$ future work)

Ideal case



- T_{clk} and T_{clj} are constant functions of time

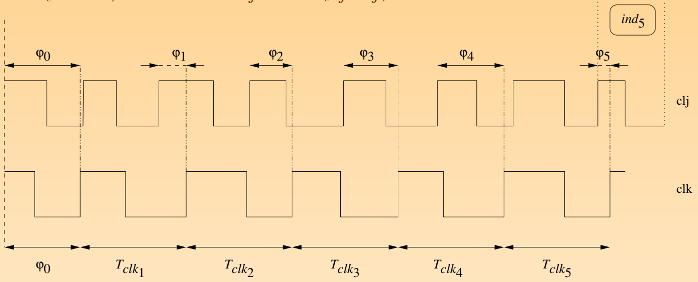
φ_i expression and corresponding sampled bit B_i logical value

$$\varphi_i = \varphi_0 + i \times T_{clk} \mod T_{clj}$$

$$B_i = 1 - \left\lfloor \frac{\varphi_i}{T_{clj}/2} \right\rfloor \text{ (assuming a 50/50 duty cycle)}$$

Adding a random jitter to each clock signal

$T_{clk} \sim \mathcal{N}(\mu_k, \sigma_k)$ and $T_{clj} \sim \mathcal{N}(\mu_j, \sigma_j)$

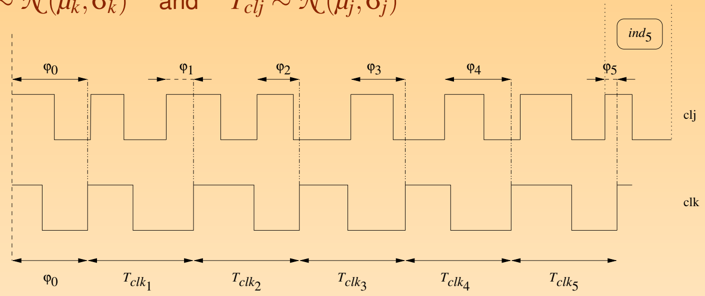


φ_i expression

$$\varphi_i = \varphi_0 + i \times T_{clk} \mod T_{clj}$$

Adding a random jitter to each clock signal

$T_{clk} \sim \mathcal{N}(\mu_k, \sigma_k)$ and $T_{clj} \sim \mathcal{N}(\mu_j, \sigma_j)$



φ_i expression

$$\varphi_i = \varphi_0 + (T_{clk_1} + \cdots + T_{clk_i}) - \left(T_{clj_1} + \cdots + T_{clj_{\mathbf{ind_i}-1}}\right)$$

$$ind_i = min\left\{m \mid \sum_{j=1}^{m} T_{clj_j} \geq \varphi_0 + \sum_{j=1}^{i} T_{clk_j}\right\}$$

Limitations

- Assuming $\{T_{clk_j}\}$ $\left(\text{resp. } \{T_{clj_j}\}\right)$ are independent realizations of the same random variable T_{clk} (resp. T_{clj})

$$\mathbf{T_{clk_{acc}}(i)} := \sum_{j=1}^{i} T_{clk_j} \sim \mathcal{N}\left(i \times \mu_k, \sqrt{i} \times \sigma_k\right)$$
$$T_{clj_{acc}}(\mathbf{ind_i} - 1) := \sum_{j=1}^{ind_i - 1} T_{clj_j} \sim \mathcal{N}\left((ind_i - 1) \times \mu_j, \sqrt{ind_i - 1} \times \sigma_j\right)$$

- Problem : $ind_i = min\left\{m \mid T_{clj_{acc}}(m) \geq \varphi_0 + \mathbf{T_{clk_{acc}}(i)}\right\}$
- Thus $\varphi_i = \varphi_0 + T_{clk_{acc}}(i) - T_{clj_{acc}}(ind_i - 1)$ cannot be expressed as a random variable following a Gaussian distribution.

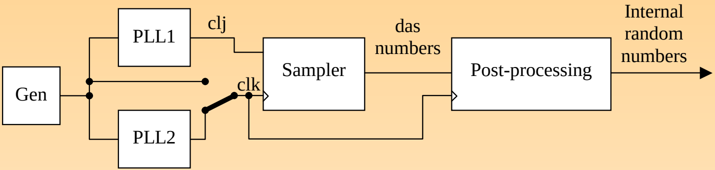
Absolute jitter : limitations

- Problem 1 : limitations of the mathematical model
- Problem 2 : absolute jitter is very difficult (if not impossible) to measure inside the chip
- Problem 3 : the generator extracts the relative jitter between two (or more) clocks and not the absolute jitters

From absolute to relative jitter (1)

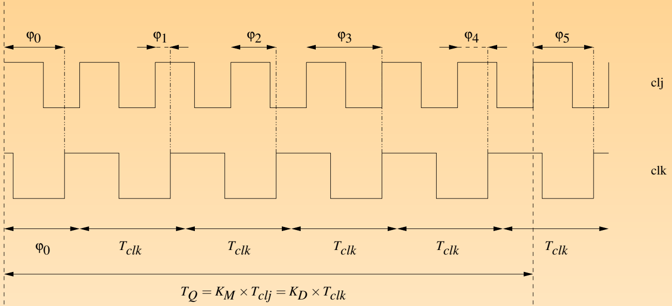
- Absolute jitter can describe a more general case (free running oscillator, jitter accumulation)
- Mathematical model limited without further (strong ?) assumptions
- Idea : the use of coherent sampling and the relationship between :
  - input frequency (f_i)
  - sampling frequency (f_s)
  - number of cycles (N_{cyc})
  - number of samples (M_{samp})
- Practical realization : PLLs

From absolute to relative jitter (2)



- Relative jitter between clj and clk
- Jitter accumulation : jitter accumulates !
- But : phase locking effect (PLL) $\longrightarrow$ bounded accumulation

Ideal case



$$\varphi_i = \varphi_0 + i \times T_{clk_{id}} - \left\lfloor \frac{\varphi_0}{T_{clj_{id}}} + \frac{i \times K_M}{K_D} \right\rfloor \times T_{clj_{id}}$$

$$B_i = 1 - \left\lfloor 2 \times \left(\frac{\varphi_0}{T_{clj_{id}}} + \frac{i \times K_M}{K_D} - \left\lfloor \frac{i \times K_M}{K_D} + \frac{\varphi_0}{T_{clj_{id}}} \right\rfloor \right) \right\rfloor$$

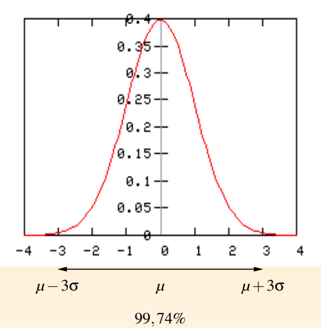
Adding the random jitter on *clj* (1)

- Relative jitter $\rightarrow T_{clk}$ supposed to be ideal
- $T_{clj_1} + \cdots + T_{clj_m} = T_{clj_{acc}}(m) \sim \mathcal{N}\left(m \times T_{clj_{id}}, \sigma_j\right)$

φ_i and ind_i

$$\varphi_i = i \times T_{clk_{id}} + \varphi_0 - T_{clj_{acc}}(ind_i - 1)$$

$$ind_i = max\{m \mid T_{clj_{acc}}(m-1) < i \times T_{clk_{id}} + \varphi_0\}$$



$$T_{clj_{acc}}(m-1) \leq_{99,74\%} (m-1) \times T_{clj_{id}} + 3\sigma_j$$

Then

$$ind_i = max\{m \mid (m-1) \times T_{clj_{id}} + 3\sigma_j < i \times T_{clk_{id}} + \varphi_0\}$$

$$ind_i = max\left\{m \mid m - 1 < \frac{i \times T_{clk_{id}} + \varphi_0 - 3\sigma_j}{T_{clj_{id}}}\right\}$$

$$ind_i = \left\lfloor \frac{i \times T_{clk_{id}} + \varphi_0 - 3\sigma_j}{T_{clj_{id}}} \right\rfloor + 1$$

Adding the random jitter on clj (2)

- Dependency between $T_{clk_{acc}}(i)$ and $T_{clj_{acc}}(ind_i - 1)$ has been removed
- φ_i can be seen as realizations of a random variable ϕ_i following a Gaussian distribution :

The random variable ϕ_i

$$\phi_i \sim \mathcal{N}\left(i \times T_{clk_{id}} + \varphi_0 - \underbrace{\left\lfloor \frac{i \times T_{clk_{id}} + \varphi_0 - 3\sigma_j}{T_{clj_{id}}} \right\rfloor}_{ind_i - 1} \times T_{clj_{id}}, \sigma_j \right)$$

- φ_i are defined by the difference between a sum of T_{clj} periods and a sum of T_{clk} periods
  Two equivalent interpretations :
  1. Set of random realizations of T_{clj} then compute exactly φ_i
  2. T_{clj} is supposed to be ideal and all φ_i are seen as realizations of the random variable ϕ_i above
- Second approach is chosen

LABORATOIRE Hubert CURIEN UMR • CNRS • 5516 • ST-ETIENNE

Period reconstruction and consequences

- φ_i realizations are not sorted
- Fischer and Drutarovsky proposed the following reconstruction (assuming K_M and K_D are relatively primes) :

$$i(j) = j \times K_M^{-1} \mod K_D$$

Then

$$0 < \varphi_{i(1)} - \varphi_0 \mod T_{clj_{id}} < \cdots < \varphi_{i(K_D-1)} - \varphi_0 \mod T_{clj_{id}}$$

The first ($i \geq 1$) sample after the reorganization is defined to be the closest one to the initial phase $\varphi_0 = \varphi_{j(0)}$

- Reconstruction also allows a simplified expression of the random variable $\phi_{i(j)}$:

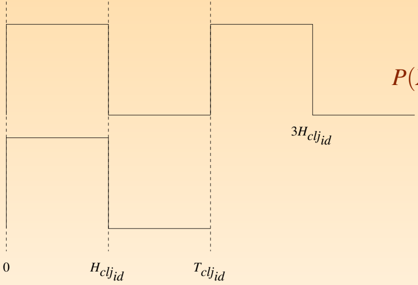
$$\phi_{i(j)} \sim \mathcal{N}\left(\varphi_0 + j \times \Delta \mod T_{clj_{id}}, \sigma_j\right)$$

where $\Delta = \dfrac{T_{clj_{id}}}{K_D}$ is the distance between $\varphi_{i(j)}$ and $\varphi_{i(j+1)}$
$\Rightarrow$ mean values of ϕ_i are uniformly distributed in the $T_{clj_{id}}$ period

Probability to sample a '1'

$$\phi_{i(j)} \sim \mathcal{N}\left(\varphi_0 + j \times \Delta \mod T_{clj_{id}}, \sigma_j\right)$$

- Means of ϕ_i are in the $[0, T_{clj_{id}}[$ interval
- Due to jitter, realizations $\varphi_{i(j)}$ can be outside this interval



$$P(X_{i(j)} = '1') = P(0 < \varphi_{i(j)} < H_{clj_{id}}) + P(T_{clj_{id}} < \varphi_{i(j)} < 3H_{clj_{id}})$$

$$\left(3\sigma_j << H_{clj_{id}} \Rightarrow P(\varphi_{i(j)} > 3H_{clj_{id}}) = 0\right)$$

Final expression

$$P\left(X_{i(j)} ='1'\right) = P\left(0 < \varphi_{i(j)} < H_{clj_{id}}\right) + P\left(T_{clj_{id}} < \varphi_{i(j)} < 3H_{clj_{id}}\right)$$

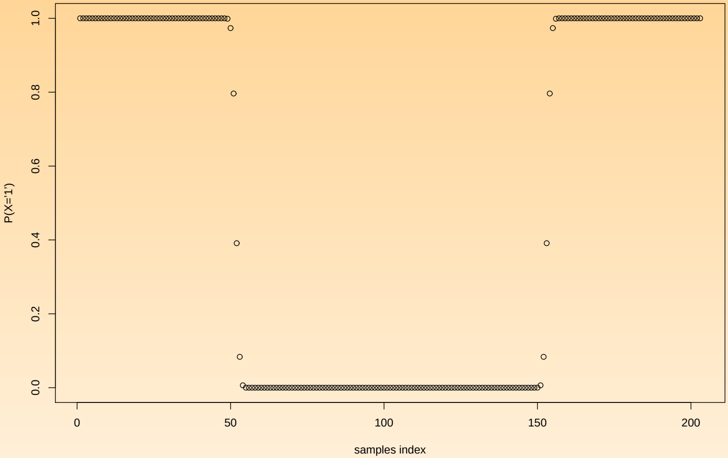
$$P(X_{i(j)} ='1') = P\left(\varphi_{i(j)} < H_{clj_{id}}\right) - P\left(\varphi_{i(j)} < 0\right) + 1 - P\left(\varphi_{i(j)} < T_{clj_{id}}\right)$$

We set $\mu_j = \varphi_0 + j \times \Delta \mod T_{clj_{id}}$, then

$$P(X_{i(j)} ='1') = \frac{1}{\sqrt{2\pi}\sigma_j}\left(\int_0^{H_{clj_{id}}} e^{-\frac{(x-\mu_j)^2}{2\sigma_j^2}}\,dx + 1 - \int_{-\infty}^{T_{clj_{id}}} e^{-\frac{(x-\mu_j)^2}{2\sigma_j^2}}\,dx\right) \quad (1)$$

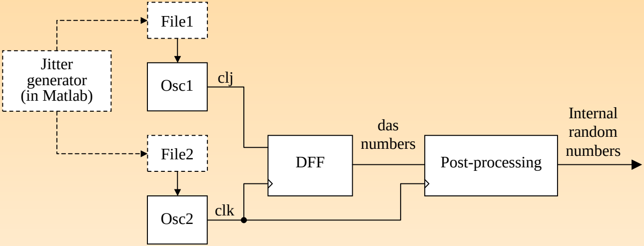
Reconstructed period ($\sigma_j = 60ps$, $\varphi_0 = T_{clj_{id}}/4$, $K_D = 203$, $K_M = 260$)

From equation 1, we plot all the $\left(i(j), P(X_{i(j)} =' 1')\right)$ for $j = 1$ to $j = K_D$



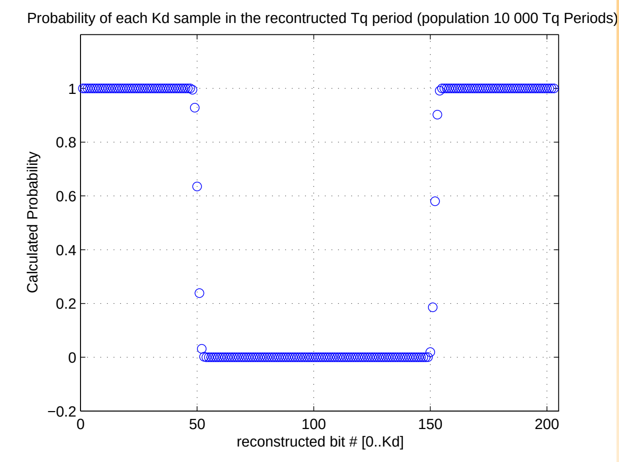
VHDL simulation

- Goal : validation of the mathematical model (random jitter only)
- Signals generated in half-periods with Matlab (to obtain a Gaussian population)
- Signals are injected in the behavioral VHDL simulation



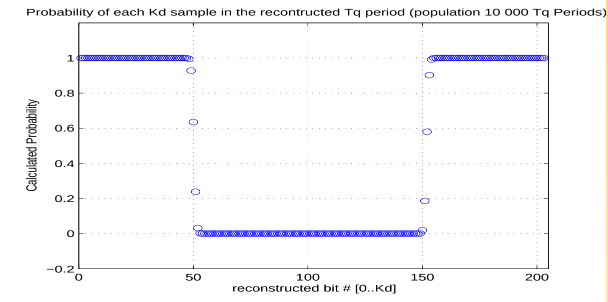
- Parameters : $f_{clj} = 74.286$ MHz, $f_{clk} = 58$ MHz, $\sigma_j = 60ps$
  $\left(\text{Note : } \frac{f_{clj}}{f_{clk}} = \frac{K_M}{K_D}\right)$

VHDL results



- Very close to the graph obtained with a mathematical equation... but still not the reality

Jitter measurement



- Relative jitter corresponds to the width of edges (rising and falling) in the $T_{clj_{id}}$ period
- Distance between two consecutives samples is $\Delta = \frac{T_{clj_{id}}}{K_D}$
- $99,74\%$ of the gaussian population is in an intervall of length $6\sigma_j$
- We count $5\Delta < x < 6\Delta$ on the rising (or falling) edge

  $x\Delta = 6\sigma_j \Rightarrow \sigma_j = \frac{x}{6}\Delta$

  $\sigma_j = \frac{x}{6} \times \frac{10^{12}}{74{,}286 \times 10^6 \times 203} = \frac{x}{6} \times 66,31ps \Rightarrow 55ps < \sigma_j < 66ps$

LABORATOIRE Hubert CURIEN UMR • CNRS • 5516 • ST-ETIENNE

Hardware experiment

- Actel AFS Evaluation board (Actel Fusion FPGA device AFS6000FG256ES) for RNG implementation
- External 40 MHz quartz oscillator
- Two embedded PLLs to generate two pairs of clock signals
  1. First configuration : $K_M = 260$ and $K_D = 203$

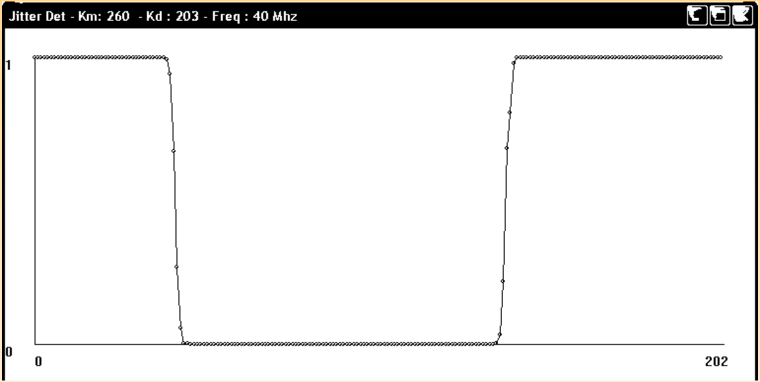
	division factor	multiplication factor	frequency (MHz)
PLL_1 (clj)	14	26	74,286
PLL_2 (clk)	10	29	58

  2. Second configuration : $K_M = 532$ and $K_D = 493$

	division factor	multiplication factor	frequency (MHz)
PLL_1 (clj)	17	28	65,88
PLL_2 (clk)	19	29	61,05

Hardware results

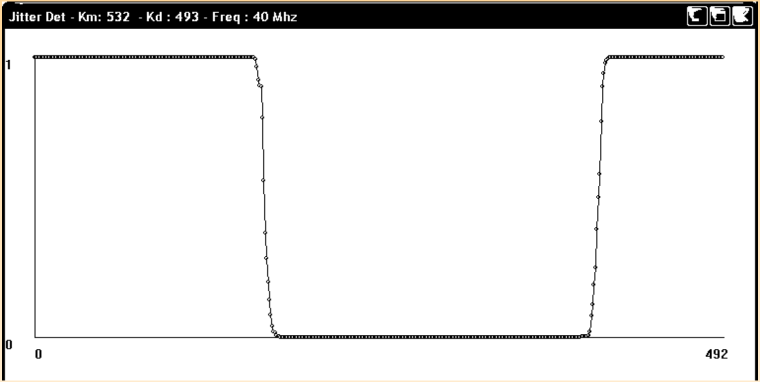
- First experiment



- We count between 5 and 6 $\Delta \Rightarrow$ real results very close to the mathematical and behavioral model
- In the case of random jitter only, we might conclude : $\sigma_j \approx 55ps$

Hardware results

- Second experiment



- We count $\approx 13\ \Delta$ on falling (or raising) edge
  $\Rightarrow \sigma_j = \frac{13}{6}\Delta = \frac{13}{6} \times \frac{10^6}{493 \times 61{,}053} \approx 72ps$ (far from $55ps$...)

Conclusion

- Security evaluation of TRNG cannot be reduced to an ability to pass a battery of statistical tests
- Entropy estimators must be computed on random variable as close as possible to the noise source
- Need of a mathematical model
  - Not an easy task
  - Based on assumptions that should be verified by hardware experiments
  - What we measure outside is not what is going on inside...
- Our model gives good results with behavioral VHDL simulation $\Rightarrow$ equations are correct
- But ! In reality, there are other aspects that influence the relative jitter (deterministic jitter component added by the PLL)
- Future work : include the deterministic jitter in the model
- Then compute entropy estimators

Thank you for your attention
Questions ?