On the Concrete Security of Goldreich's PRG

Yann Rotella
Joint work with Geoffroy Couteau, Aurélien Dupin,
Pierrick Méaux and Mélissa Rossi

January 31, 2019

Radboud University

PseudoRandom Generators

Seed: $(x_1, x_2, \ldots, x_n) \in \mathbb{F}_2^n$

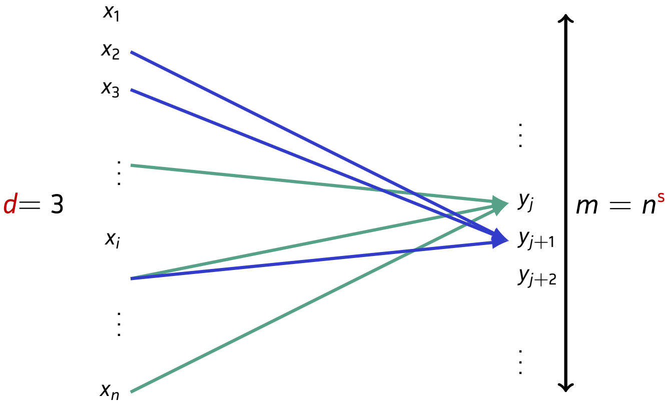
↓ PRG

Output: $(y_1, y_2, \ldots, y_n, y_{n+1}, \ldots, y_m) \in \mathbb{F}_2^n$

- $(y_i)_{i \leq m}$ should be indistinguishable from a random string;
- it is hard to recover $(x_i)_{i \leq n}$ using the knowledge of $(y_i)_{i \leq m}$.

Structure of this Talk

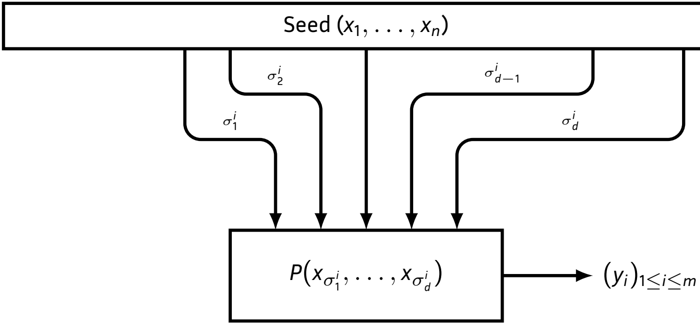
Stretch and locality



Theoretical applications

- Semi Secure computation with constant computational overhead [Ishai et al. STOC 2018, Applebaum et al. CRYPTO 2017]
- MPC-friendly primitives [Albrecht et al. EC 2015, Canteaut et al. FSE 2016, Méaux et al. EC 2016, Grassi et al. ACM-CCS 2016]
- Indistinguishability Obfuscation [Sahai and Waters STOC 2014, Lin and Tessaro CRYPTO 2017]
- Cryptographic Capsules [Boyle et al. ACN-CCS 2017]

Description of Goldreich's PRG



$m = n^s$, s is the stretch.

Parameters

- Stretch s > 1
- Subsets $(\sigma^i)_{i \leq 1}$
- Boolean function (predicate) P
- Locality d

Parameters

- Stretch s > 1 ?
- Subsets $(\sigma^i)_{i \leq 1}$?
- Boolean function (predicate) P ?
- Locality d ?

Subsets

The subsets should be sufficiently expanding: for some k, every k subsets should cover $k + \Omega(n)$ elements of $\{1, \ldots, n\}$.

Subsets

The subsets should be sufficiently expanding: for some k, every k subsets should cover $k + \Omega(n)$ elements of $\{1, \ldots, n\}$.

Ok if they are chosen uniformly random

Generic sub-exponential seed recovery

- Create a list of all possible values for $(2\varepsilon) * n$ variables.

Generic sub-exponential seed recovery

- Create a list of all possible values for $(2\varepsilon) * n$ variables.
- A value x' of the list can agree on $(1/2 + \varepsilon) * n$ output bits.

Generic sub-exponential seed recovery

- Create a list of all possible values for $(2\varepsilon) * n$ variables.
- A value x' of the list can agree on $(1/2 + \varepsilon) * n$ output bits.
- Final complexity:

$$2^{n^{1-(s-1/2d)}}$$

Generic sub-exponential seed recovery

- Create a list of all possible values for $(2\varepsilon) * n$ variables.
- A value x' of the list can agree on $(1/2 + \varepsilon) * n$ output bits.
- Final complexity:

$$2^{n^{1-(s-1/2d)}}$$

$$s = 1.45 \text{ and } d = 5 \Rightarrow 2^{n^{0.955}}$$

Predicate criteria

- degree [Goldreich 2000]
- rational degree (algebraic immunity) [Applebaum and Lovett STOC 2016]

$$\mathrm{AI}(P) > s$$

- resilience [O'Donnelland Witmer CCC 2014, Applebaum 2015]

$$\mathrm{res}(P) > 2s$$

locality

$$\left.\begin{array}{r} \textit{degree} \\ \textit{resilience} \\ \textit{Siegenthaler} \end{array}\right\} \Rightarrow d \geq 5$$

locality

$$\left.\begin{array}{r} \textit{degree} \\ \textit{resilience} \\ \textit{Siegenthaler} \end{array}\right\} \Rightarrow d \geq 5$$

$$P_5(x_1, x_2, x_3, x_4, x_5) = x_1 + x_2 + x_3 + x_4 x_5$$

Our results

- A new subexponential-time attack in $2^{O(n^{2-s})}$.

Our results

- A new subexponential-time attack in $2^{O(n^{2-s})}$.
- Linearization and Gröbner-based attacks.

Our results

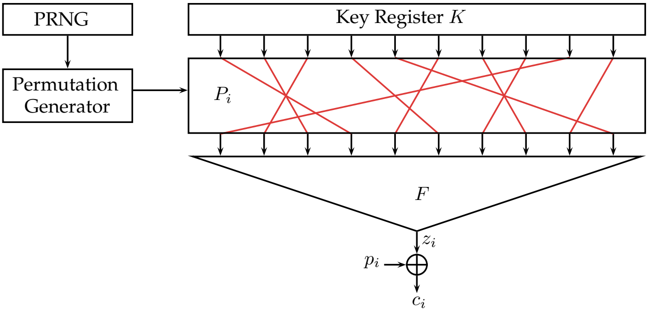
- A new subexponential-time attack in $2^{O(n^{2-s})}$.
- Linearization and Gröbner-based attacks.
- Generalization of the subexponential attack to all predicates.

Our results

- A new subexponential-time attack in $2^{O(n^{2-s})}$.
- Linearization and Gröbner-based attacks.
- Generalization of the subexponential attack to all predicates.
- locality and stretch are linked to the size of the seed.

Plan of this Section

Cryptanalysis of FLIP [Duval, Lallemand, Rotella CRYPTO 2016]



$$\begin{aligned} F(x) = & x_1 + x_2 + \cdots + x_{k_1} \\ & + x_{k_1+1}x_{k_1+2} + \cdots + x_{k_2-1}x_{k_2} \\ & + x_{k_3} + x_{k_3+1}x_{k_3+2} + \cdots + x_{n-14} \cdots x_{n-1}x_n \end{aligned}$$

FLIP vs Goldreich's PRG

- FLIP: overdetermined
- Goldreich's PRG: underdetermined

$$P_5(x_1, x_2, x_3, x_4, x_5) = x_1 + x_2 + x_3 + x_4 x_5$$

Collect linear equations

$$x_1 + x_4 + x_8 + x_9 x_{11} = 1$$
$$x_{14} + x_5 + x_7 + x_1 x_4 = 0$$
$$x_{13} + x_{10} + x_3 + x_{11} x_9 = 1$$

Collect linear equations

$$x_1 + x_4 + x_8 + x_9x_{11} = 1$$
$$x_{14} + x_5 + x_7 + x_1x_4 = 0$$
$$x_{13} + x_{10} + x_3 + x_{11}x_9 = 1$$

We get the following linear equation:

$$x_1 + x_4 + x_8 + x_{13} + x_{10} + x_3 = 0$$

Collect linear equations

$$x_1 + x_4 + x_8 + x_9 x_{11} = 1$$
$$x_{14} + x_5 + x_7 + x_1 x_4 = 0$$
$$x_{13} + x_{10} + x_3 + x_{11} x_9 = 1$$

We get the following linear equation:

$$x_1 + x_4 + x_8 + x_{13} + x_{10} + x_3 = 0$$

number of collisions $c \in O\left(n^{2(s-1)}\right)$

Guessing phase

- Choose the ℓ variables that appear the most in the quadratic terms, such that you get $n - c - \ell$ linear equations.

Guessing phase

- Choose the ℓ variables that appear the most in the quadratic terms, such that you get $n - c - \ell$ linear equations.
- For all possible values of the ℓ bits:

Guessing phase

- Choose the ℓ variables that appear the most in the quadratic terms, such that you get $n - c - \ell$ linear equations.
- For all possible values of the ℓ bits:
- Solve the correponding linear system of n linear equations.

Analysis and complexity

- **Complexity:** $\ell < n^{2-s} \rightarrow \mathcal{O}\left(n^3 2^{n^{2-s}}\right)$
  Conjectured secure up to $s < 1.5$.

Analysis and complexity

- **Complexity:** $\ell < n^{2-s} \rightarrow \mathcal{O}\left(n^3 2^{n^{2-s}}\right)$
  Conjectured secure up to $s < 1.5$.
- The equations might be linearly dependent (almost never the case).

This leads to a strong distinguisher and allows to determine if the Guess is right or wrong.

Analysis and complexity

- **Complexity:** $\ell < n^{2-s} \rightarrow \mathcal{O}\left(n^3 2^{n^{2-s}}\right)$
  Conjectured secure up to $s < 1.5$.
- The equations might be linearly dependent (almost never the case).

  This leads to a strong distinguisher and allows to determine if the Guess is right or wrong.
- If the equations aren't linearly dependent, then we solve a full rank linear system of size n.

Table: Average number of collisions

n	256	512	1024	2048	4096
$s = 1.45$	142	269	506	946	1771
$s = 1.4$	83	145	254	442	773
$s = 1.3$	28	42	64	97	147

Table: Theoretical number of guesses (worst case)

n	256	512	1024	2048
$s = 1.45$	4	7	11	18
$s = 1.4$	9	15	23	37
$s = 1.3$	20	34	56	94

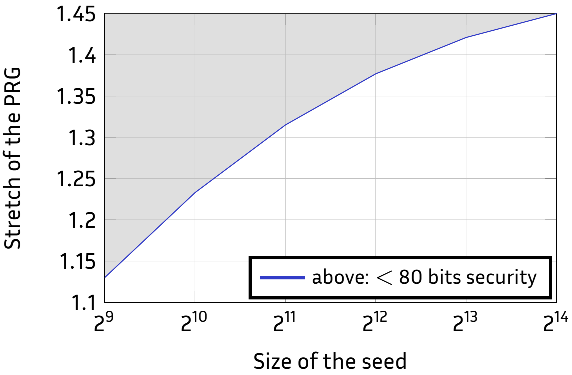
Table: Experimental number of guesses (average)

n	256	512	1024	2048	4096
$s = 1.45$	4	6	9	14	21
$s = 1.4$	6	11	17	27	44
$s = 1.3$	13	23	39	65	110

Table: Complexity of our attack.

	512	1024	2048	4096
$< 2^{80}$	1.120	1.215	1.296	1.361
$< 2^{128}$	1.048	1.135	1.222	1.295

Complexity


1.45
1.4
1.35
1.3
1.25
1.2
1.15
1.1
Stretch of the PRG
2^9
2^{10}
2^{11}
2^{12}
2^{13}
2^{14}
Size of the seed
above: $<$ 80 bits security

Plan of this Section

Collecting equations of degree 2

$$x_{i_1} + x_{i_2} + x_{i_3} + x_{i_4}x_{i_5} = y_i \quad (1)$$

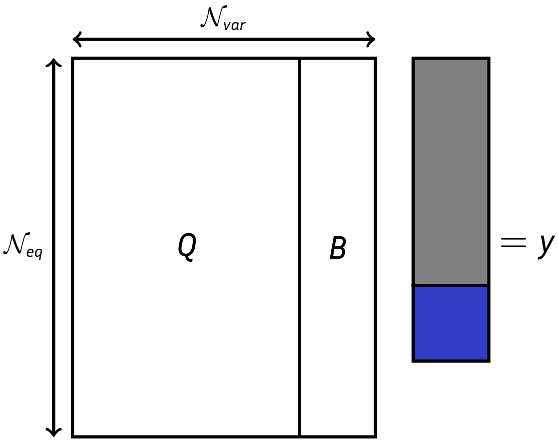
$$x_{j_1} + x_{j_2} + x_{j_3} + x_{j_4}x_{j_5} = y_j \quad (2)$$

using (1): $x_{i_4}x_{i_1} + x_{i_4}x_{i_2} + x_{i_4}x_{i_3} + x_{i_4}x_{i_5} = x_{i_4}y_i$

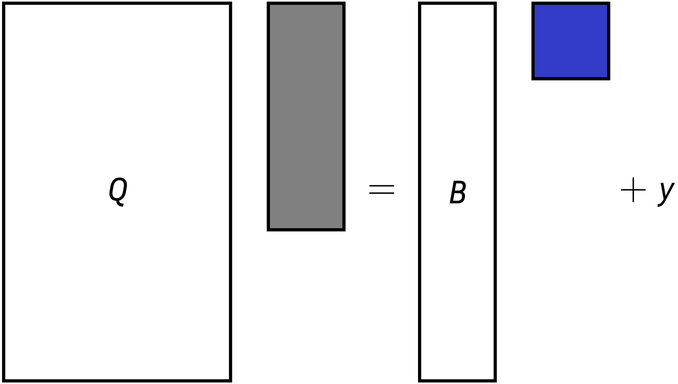
if $x_{i_4}x_{i_5} = x_{j_4}x_{j_5}$**:** $x_k y_i + x_k y_j = x_k x_{i_1} + x_k x_{i_2} + x_k x_{i_3} + x_k x_{j_1} + x_k x_{j_2} + x_k x_{j_3}$

if $x_{i_4} = x_{j_4}$**:** $x_{j_5} \times (1) + x_{i_5} \times (2)$

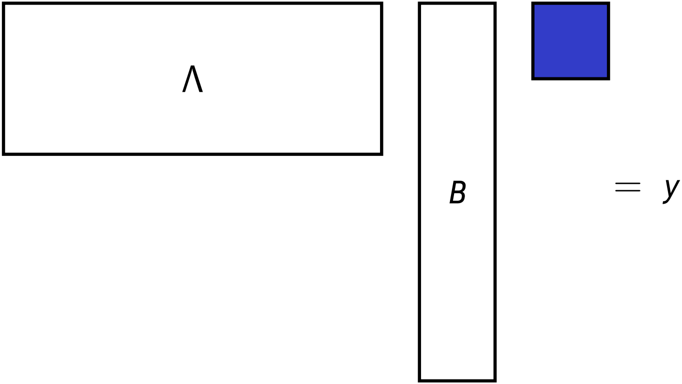
Solving the system



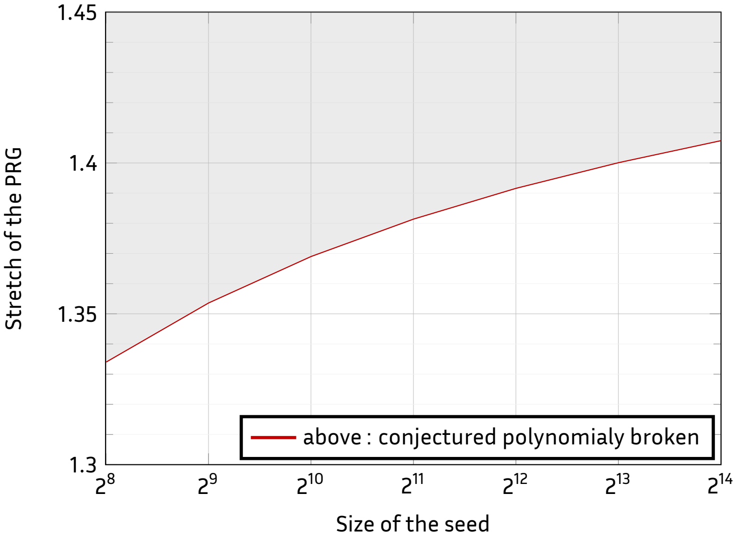
Solving the system



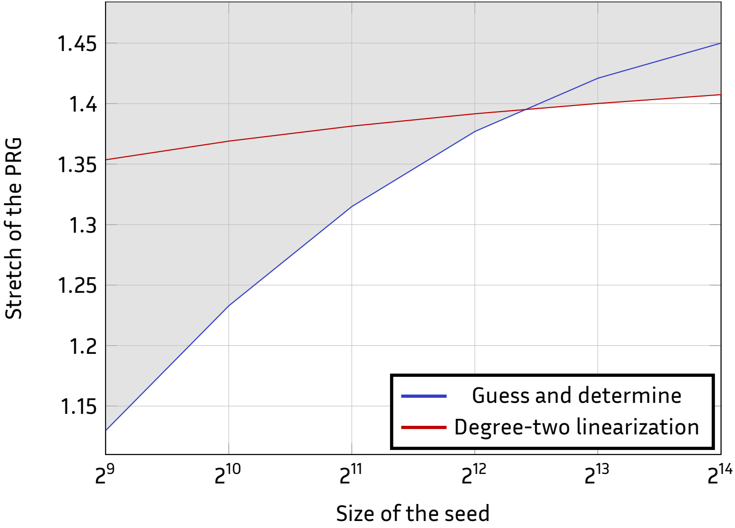
Solving the system



Experimental results



Results on P_5



Plan of this Section

General sub-exponential time attack

$$P = x_1 + x_2 + \cdots + x_\ell + f(x_{\ell+1}, \ldots, x_d)$$

$k = d - \ell \Rightarrow$

$$2^{n^{\frac{k-s}{k-1}}}$$

r-bit fixing Algebraic Immunity [MJSC, EC 2016]

$$\min_{(b,i)}\left(\mathsf{AI}\left(f_{(b,i)}\right)\right)$$

where bits at positions i are fixed.

For example, if $f(x_1, x_2, x_3, x_4, x_5) = x_1 + x_2x_3x_4 + x_5$, then

$$f_{(1,2),(0,1)} = x_3x_4 + x_5$$

Improvement

Fixing j bits on a predicate of the form

$$P = x_1 + x_2 + \cdots + x_\ell + f(x_{\ell+1}, \ldots, x_d)$$

gives equations of degree smaller than

$$\left\lceil \frac{k-j}{2} \right\rceil + 1$$

Improvement

Fixing j bits on a predicate of the form

$$P = x_1 + x_2 + \cdots + x_\ell + f(x_{\ell+1}, \ldots, x_d)$$

gives equations of degree smaller than

$$\left\lceil \frac{k-j}{2} \right\rceil + 1$$

If the stretch is "big enough", we can improve the previous generic attack using bounds on r-bit fixing algebraic immunity.

Application to XOR-MAJ predicates

- Fix enough bits to 0 (or 1).
- Recover linear equations.

$$O\left(2^{n^{1-\frac{s-1}{k/2+1}}}\right)$$

Polynomial Attack (AL theorem improvement)

Let N_e be the dimension of the vectorspace of annihilators of degree e, then if

$$s \geq e - \frac{\log(N_e)}{\log(n))}$$

then there exists a polynomial-time algorithm that breaks the PRG.

Plan of this Section

Conclusion

- First concrete parameters given.
- Symmetric Cryptanalysis can be applied to theoretical constructions.
- Several techniques that do not capture the same phenomenon.
- If s is close to 1.5, then the seed size has to be very big.
- New theorems and criteria on predicates.

Perspectives

- Link between expander graphs, first attack (Guess-and-Determine) and second attack (Gröbner).
- Capture the Gröbner success phenomenon.
- Find best predicate ?

Thank You !

Questions ?