Pseudorandom Correlation Functions from Variable-Density LPN, Revisited

Geoffroy Couteau[1]

Clément Ducros[2]

[1]CNRS, IRIF, Université de Paris

[2]Université de Paris, IRIF, INRIA

May 10, 2023

A new primitive [BCG+20] [1]

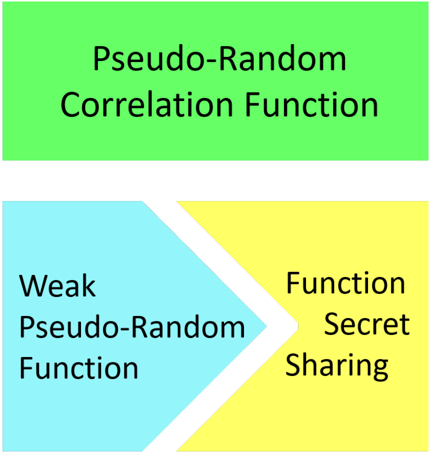
Pseudo-Random Correlation Function

Weak Pseudo-Random Function (WPRF)

A function $f, A \to B$ is a WPRF when the two distributions $\mathcal{D} = \{f(x), x \xleftarrow{\$} A\}$ and $\mathcal{D}' = \{y \xleftarrow{\$} B\}$ are indistinguishable.
i.e. the adversary can asks for random samples $(x, f(x))$ but can't evaluate the function on chosen inputs.

[1] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, and Peter Scholl. Correlated pseudorandom functions from variable-density LPN.
In *61st FOCS*, pages 1069–1080. IEEE Computer Society Press, November 2020

A new primitive [BCG$^+$20] [1]

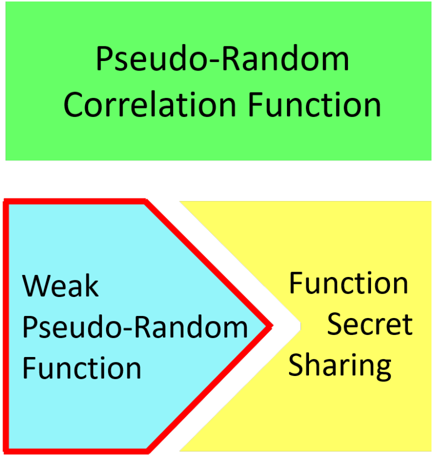


Weak Pseudo-Random Function (WPRF)

A function $f, A \to B$ is a WPRF when the two distributions
$\mathcal{D} = \{f(x), x \xleftarrow{\$} A\}$ and $\mathcal{D}' = \{y \xleftarrow{\$} B\}$
are indistinguishable.
i.e. the adversary can asks for random samples $(x, f(x))$ but can't evaluate the function on chosen inputs.

[1] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, and Peter Scholl. Correlated pseudorandom functions from variable-density LPN.
In *61st FOCS*, pages 1069–1080. IEEE Computer Society Press, November 2020

A new primitive [BCG+20] [1]



Weak Pseudo-Random Function (WPRF)

A function $f, A \to B$ is a WPRF when the two distributions $\mathcal{D} = \{f(x), x \xleftarrow{\$} A\}$ and $\mathcal{D}' = \{y \xleftarrow{\$} B\}$ are indistinguishable.
i.e. the adversary can asks for random samples $(x, f(x))$ but can't evaluate the function on chosen inputs.

---

[1] Elette Boyle, Geoffroy Couteau, Niv Gilboa, Yuval Ishai, Lisa Kohl, and Peter Scholl. Correlated pseudorandom functions from variable-density LPN.
In *61st FOCS*, pages 1069–1080. IEEE Computer Society Press, November 2020

Outline

A new WPRF

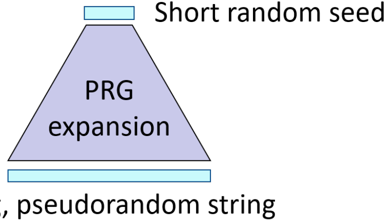
A framework of attacks

Our contribution

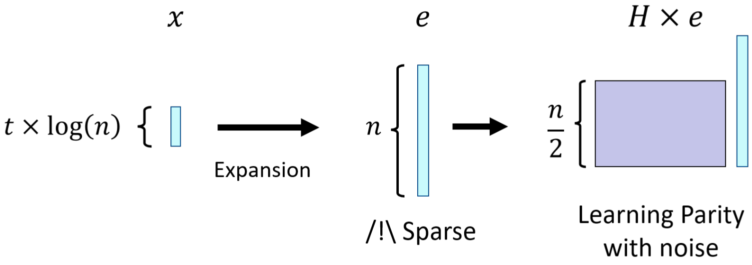
1 - A new WPRF

About Pseudo-Random Generators

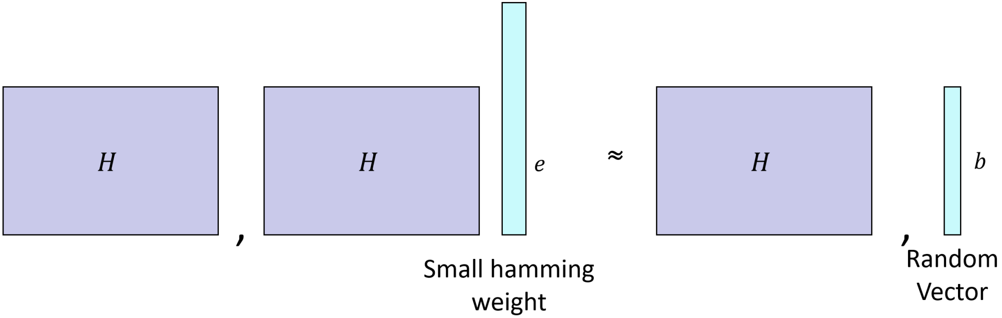
Pseudo Random Generators ?



How to construct PRG?



Learning Parity with Noise



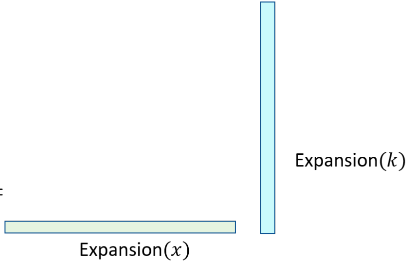
Syndrome Decoding Assumption

- Let H be a random matrix, e a random noise vector of small Hamming Weight. Then $H \cdot e^\top$ is indistinguishable from a random vector.
- What about more structured H?

Using this idea for WPRF

Same idea! Use LPN!

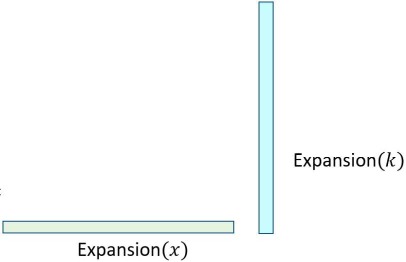
$f_k(x) =$



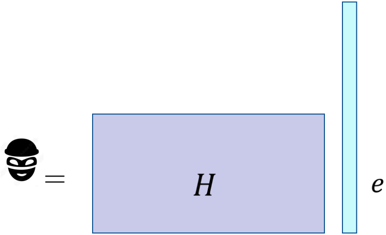
Using this idea for WPRF

Same idea! Use LPN!

$f_k(x) =$



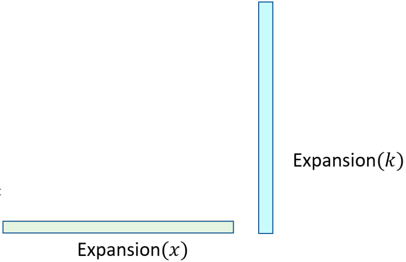
View of



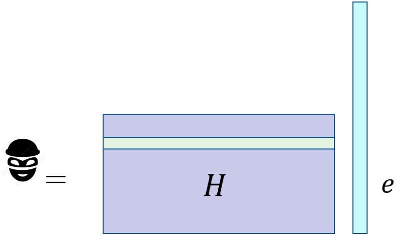
Using this idea for WPRF

Same idea! Use LPN!

$f_k(x) =$

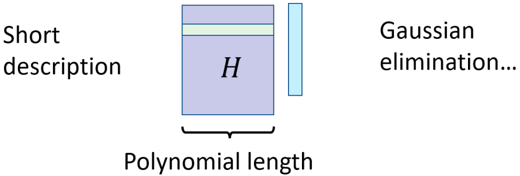


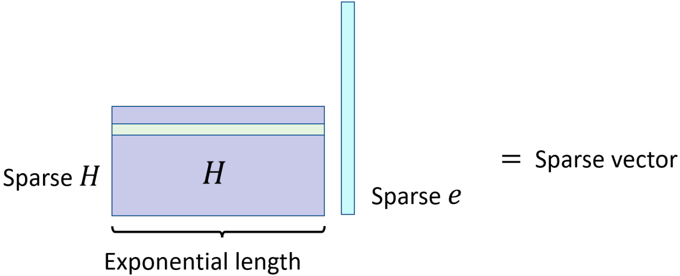
View of



Each row can be seen as an input. The adversary knows H, and the result of $H \cdot e^\top$.
Number N of samples $\rightarrow N$ rows in H. N should be exponentially big.

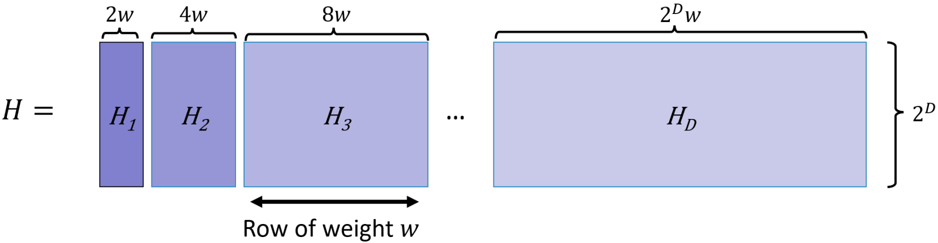
Two problems





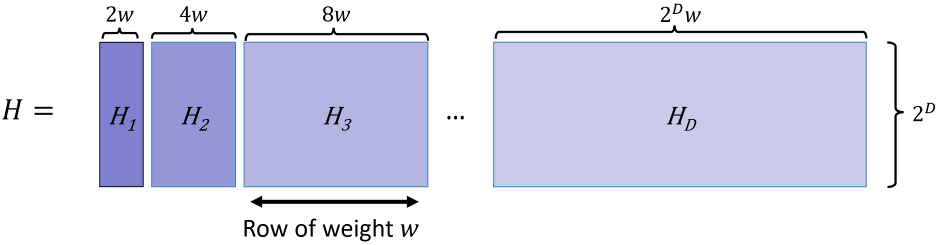
Variable Density Learning Parity with Noise [BCG+20]

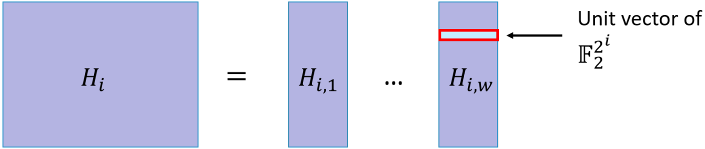
Solution : **Exponentially decreasing density**



Variable Density Learning Parity with Noise [BCG+20]

Solution : **Exponentially decreasing density**





The noise follows the same shape as one row of H.

2 - A framework of attacks

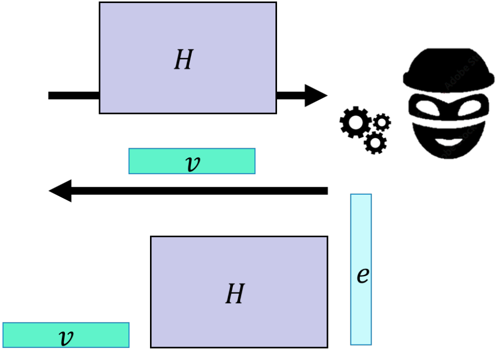
Linear attacks paradigm [BCG+20]

Bias of a distribution

Given a distribution $\mathcal{D}$ over $\mathbb{F}_2^n$, a vector $v \in \mathbb{F}_2^n$:

$$\text{bias}_v(\mathcal{D}) = \left| \frac{1}{2} - \Pr_{u \xleftarrow{\$} \mathcal{D}} \left[v^\top \cdot u = 1 \right] \right|$$

The bias of $\mathcal{D}$, denoted bias($\mathcal{D}$), is the maximum bias of $\mathcal{D}$ with respect to any nonzero vector v.



- Send H to the adversary
- The adversary returns a **test vector** v computed from H with unbounded time.
- Is $v^\top \cdot u = v^\top \cdot H \cdot e$ biased ?

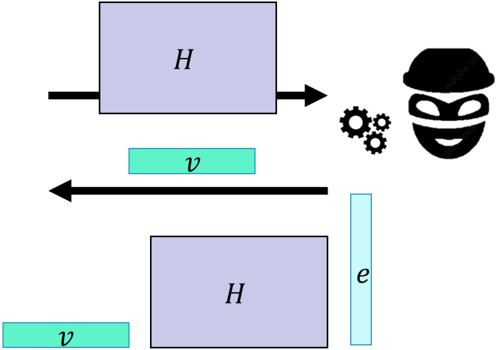
Linear attacks paradigm [BCG+20]

Bias of a distribution

Given a distribution $\mathcal{D}$ over $\mathbb{F}_2^n$, a vector $v \in \mathbb{F}_2^n$:

$$\text{bias}_v(\mathcal{D}) = \left| \frac{1}{2} - \Pr_{u \xleftarrow{\$} \mathcal{D}} \left[v^\top \cdot u = 1 \right] \right|$$

The bias of $\mathcal{D}$, denoted bias$(\mathcal{D})$, is the maximum bias of $\mathcal{D}$ with respect to any nonzero vector v.



- Send H to the adversary
- The adversary returns a **test vector** v computed from H with unbounded time.
- Is $v^\top \cdot u = v^\top \cdot H \cdot e$ biased ?

Resistance against linear attacks

Resistance against linear attacks

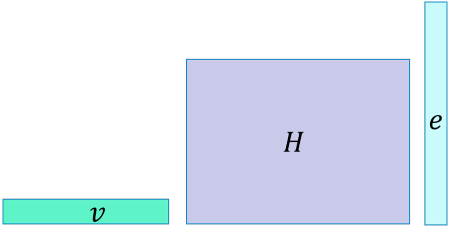
We obtain the resistance against linear attacks when

$$\Pr_{x^1,\cdots,x^{N(\lambda)} \xleftarrow{\$} \mathbb{F}_2^{n(\lambda)}}[\mathsf{bias}(\mathcal{D}(x) > \epsilon(\lambda)] < \delta(\lambda)$$

where ϵ and δ are small depending on the security parameter λ.

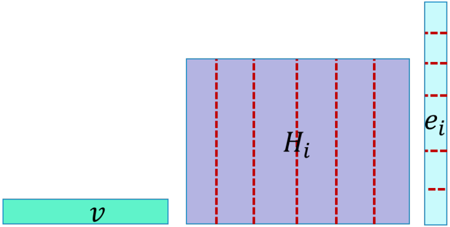
Attacks	Linear?
Gaussian elimination	✓
Statistical decoding	✓
Information set decoding	✓
BKW	✓
Algebraic attack	✗
Statistical Query Algorithm	✗

Analysis of security



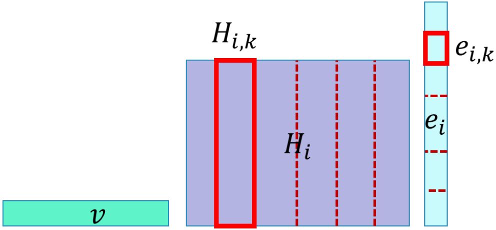
- Evaluation of the bias of $H \cdot e$

Analysis of security



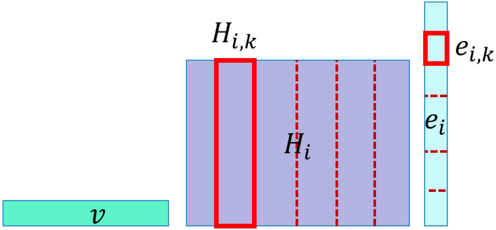
- The block H_i protects against vectors attacks v of Hamming Weight $l \in [2^{i-1}, 2^i]$

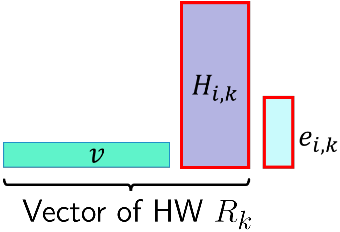
Analysis of security



- The block H_i protects against vectors attacks v of Hamming Weight $l \in [2^{i-1}, 2^i]$

Analysis of security





- The block H_i protects against vectors attacks v of Hamming Weight $l \in [2^{i-1}, 2^i]$

- We focus on the random value $Z_k = |2^{i-1} - R_k|$, e.g. the distance to the mean.

3 - Our contribution

Our contribution

- [BCG+20] proved VDLPN secure against linear attacks. Their construction was not intended to be efficient.

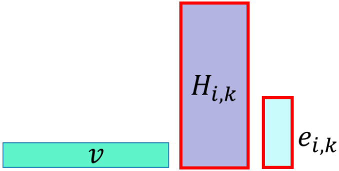
Our contribution

- [BCG+20] proved VDLPN secure against linear attacks. Their construction was not intended to be efficient.

Our contribution is divided in two parts:

- We provide a variant of VDLPN, with a new proof that offers results getting close to efficient.
- We found an error in the proof of security of [BCG+20] and fixed it.

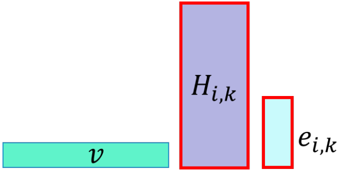
First axis, a better analysis



Bias for each sub-matrix :

$$\text{bias}_{\mathbf{v}}(O^{i,k}) = \frac{Z_k}{2^i}.$$

First axis, a better analysis



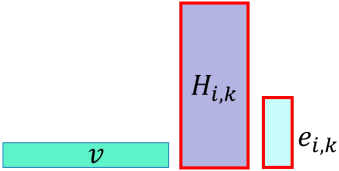
Bias for each sub-matrix :

$$\text{bias}_{\mathbf{v}}(O^{i,k}) = \frac{Z_k}{2^i}.$$

To obtain the bias of the entire bloc i, we use the Pilling-Up Lemma.

$$\text{bias}_{\mathbf{v}}(O^{i}) \leq \frac{1}{2} \cdot \prod_{k=1}^{w} \frac{Z_k}{2^{i-1}}.$$

First axis, a better analysis



Bias for each sub-matrix :

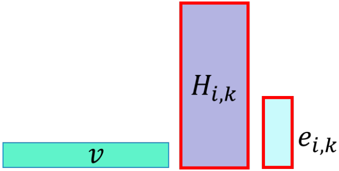
$$\text{bias}_{\mathbf{v}}(O^{i,k}) = \frac{Z_k}{2^i}.$$

To obtain the bias of the entire bloc i, we use the Pilling-Up Lemma.

$$\text{bias}_{\mathbf{v}}(O^{i}) \leq \frac{1}{2} \cdot \prod_{k=1}^{w} \frac{Z_k}{2^{i-1}}.$$

$$\Pr[\text{bias}_{\mathbf{v}}(O^{i}) > B]$$

First axis, a better analysis



Bias for each sub-matrix :

$$\text{bias}_{\mathbf{v}}(O^{i,k}) = \frac{Z_k}{2^i}.$$

To obtain the bias of the entire bloc i, we use the Pilling-Up Lemma.

$$\text{bias}_{\mathbf{v}}(O^{i}) \leq \frac{1}{2} \cdot \prod_{k=1}^{w} \frac{Z_k}{2^{i-1}}.$$

$$\Pr[\text{bias}_{\mathbf{v}}(O^{i}) > B] = \Pr\left[\prod_{k=1}^{w} Z_k > 2^{(i-1)w} \times (2B)\right] \leq \Pr\left[\sum_{k=1}^{w} Z_k > w \cdot 2^{(i-1)} \cdot c\right]$$

The previous proof taked into accounts only the top countributors.
Our key idea : transform the product of Z_k into a sum ; that we can afterwards bound with known concentration bounds.

Firt axis, a better analysis

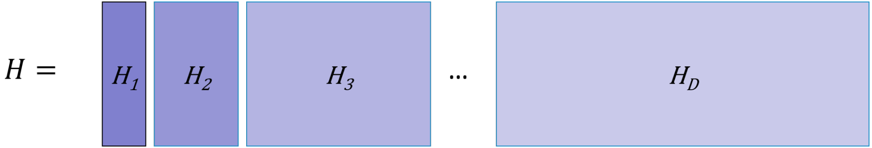
The expression we obtain is of the shape

$$\Pr\left[\mathsf{bias}_{\mathbf{v}}(O^i) \geq c^w\right] \leq \exp(-\frac{w}{a})$$

a is reduced by 3 order of magnitude.

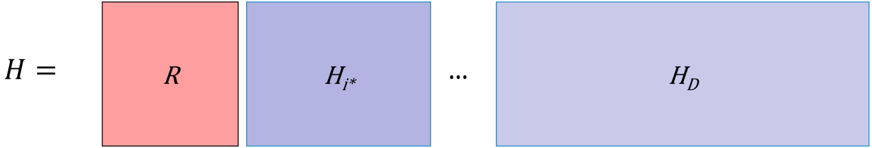
Second axis: a slightly different assumption

Loose bounds for small matrices.



Second axis: a slightly different assumption

Loose bounds for small matrices.



- The matrix R is random, and offer protection against all the attack vectors of Hamming Weight $l < 2^{i^*-1}$.
- We set the size of R according to our security parameter.

Third Axis : a simulation analysis

Natural question during the proof : estimate β such that $\mathbb{E}[Z_k] < \beta \cdot 2^i$.

- Loose upper bound on β
- Better estimation of β estimated via computer simulation. Security parameter divided by 4.

Results

- Our variant has bias at most 2^{-80} with probability at least $1 - 2^{-80}$; with $w = 380$ and maximum number of samples $N = 2^{30}$.
- Similar security parameters were proved in [BCG$^+$20] but for $w \geq 10^6$.

Results

- Our variant has bias at most 2^{-80} with probability at least $1 - 2^{-80}$; with $w = 380$ and maximum number of samples $N = 2^{30}$.
- Similar security parameters were proved in [BCG+20] but for $w \geq 10^6$.

Impact on the PCF construction scheme :

Variant	Seed size	PCF evaluations per second
This work	2.94MB	500
Aggressive variant	0.35MB	3890

Table: PCF seed size and speed using a 3.8GHz processor, on single core, estimation.

Results

- Our variant has bias at most 2^{-80} with probability at least $1 - 2^{-80}$; with $w = 380$ and maximum number of samples $N = 2^{30}$.
- Similar security parameters were proved in [BCG+20] but for $w \geq 10^6$.

Impact on the PCF construction scheme :

Variant	Seed size	PCF evaluations per second
This work	2.94MB	500
Aggressive variant	0.35MB	3890

Table: PCF seed size and speed using a 3.8GHz processor, on single core, estimation.

Thank you for your attention !