Codes with locality: constructions and applications to cryptographic protocols

Julien Lavauzelle
École Polytechnique & INRIA Saclay, Université Paris-Saclay
PhD defense
30/11/2018
1. Codes with locality
   - Locality in coding theory, examples
   - Lifted projective Reed-Solomon codes
   - A combinatorial point of view

2. Private information retrieval from transversal designs
   - Private information retrieval (PIR)
   - Transversal designs and codes
   - A new PIR construction
   - Instances

3. Proofs-of-retrievability

4. Conclusion
1. Codes with locality
   - Locality in coding theory, examples
   - Lifted projective Reed-Solomon codes
   - A combinatorial point of view

2. Private information retrieval from transversal designs
   - Private information retrieval (PIR)
   - Transversal designs and codes
   - A new PIR construction
   - Instances

3. Proofs-of-retrievability

4. Conclusion
1. Codes with locality
   Locality in coding theory, examples
   Lifted projective Reed-Solomon codes
   A combinatorial point of view

2. Private information retrieval from transversal designs
   Private information retrieval (PIR)
   Transversal designs and codes
   A new PIR construction
   Instances

3. Proofs-of-retrievability

4. Conclusion
Original goal: transmit information in the presence of noise.

<table>
<thead>
<tr>
<th>message $m \in \mathbb{F}_q^k$</th>
<th>codeword $c \in C \subseteq \mathbb{F}_q^n$</th>
<th>channel</th>
<th>noisy codeword $c' \in \mathbb{F}_q^n$</th>
<th>decoded message $m' (= m?)$</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\rightarrow$</td>
<td></td>
<td>$\uparrow$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>errors ($c_i \neq c_i' \in \mathbb{F}_q$) or erasures ($c_j' = \bot$)</td>
<td></td>
</tr>
</tbody>
</table>

Hamming distance $d(u, v) = |\{i \in [1, n], u_i \neq v_i\}|$. 

$\min d(C) = \min \{d(c, c'), c \neq c', (c, c') \in C^2\}$. 

$C$ linear over $\mathbb{F}_q$, with $k = \dim(C)$. 

$J. Lavauzelle$  

PhD defense – Codes with locality: constructions and applications to cryptographic protocols –  PhD defense
Original goal: transmit information in the presence of noise.

message $m \in \mathbb{F}_q^k \mapsto$ codeword $c \in \mathcal{C} \subseteq \mathbb{F}_q^n$ $\xrightarrow{\text{channel}}$ noisy codeword $c' \in \mathbb{F}_q^n$ $\mapsto$ decoded message $m'(=m?)$

errors ($c_i \neq c'_i \in \mathbb{F}_q$) or erasures ($c'_j = \bot$)

Hamming distance $d(u,v) := |\{i \in [1,n], u_i \neq v_i\}|.$
Original goal: transmit information in the presence of noise.

<table>
<thead>
<tr>
<th>Original code</th>
<th>Transmitted information</th>
</tr>
</thead>
<tbody>
<tr>
<td>Message $m \in \mathbb{F}_q^k$</td>
<td>Codeword $c \in C \subseteq \mathbb{F}_q^n$</td>
</tr>
<tr>
<td>Channel</td>
<td>Noisy codeword $c' \in \mathbb{F}_q^n$</td>
</tr>
<tr>
<td></td>
<td>Decoded message $m' (= m?)$</td>
</tr>
<tr>
<td>Errors ($c_i \neq c'_i \in \mathbb{F}_q$) or erasures ($c'_i = \bot$)</td>
<td></td>
</tr>
</tbody>
</table>

Hamming distance $d(u, v) := |\{i \in [1, n], u_i \neq v_i\}|$.

- $d = d_{\text{min}}(C) := \min\{d(c, c'), c \neq c', (c, c') \in C^2\}$,
- $C$ linear over $\mathbb{F}_q$, with $k = \dim(C)$. 
**Definition** (Reed-Solomon code). Let $x = (x_1, \ldots, x_n) \in \mathbb{F}_q^n$, pairwise distinct.

$$RS_q(r, x) := \{(f(x_1), \ldots, f(x_n)), f \in \mathbb{F}_q[X], \deg f \leq r\}$$
**Definition** (Reed-Solomon code). Let \( x = (x_1, \ldots, x_n) \in \mathbb{F}_q^n \), pairwise distinct.

\[
\text{RS}_q(r, x) := \{(f(x_1), \ldots, f(x_n)), f \in \mathbb{F}_q[X], \deg f \leq r\}
\]

- Dimension \( k = r + 1 \)
- Minimum distance \( d_{\text{min}} = n - r \)
- Can decode any \( b \) errors and \( e \) erasures
  - if \( e + 2b < d_{\text{min}} \)
  - in time \( \Theta(n \log^3 n) \).
Typical example: Reed-Solomon codes

**Definition (Reed-Solomon code).** Let \( x = (x_1, \ldots, x_n) \in \mathbb{F}_q^n \), pairwise distinct.

\[
\text{RS}_q(r, x) := \{ (f(x_1), \ldots, f(x_n)) | f \in \mathbb{F}_q[X], \deg f \leq r \}
\]

\[c_i = f(x_i)\]

- **Dimension** \( k = r + 1 \)
- **Minimum distance** \( d_{\min} = n - r \)
- **Can decode any** \( b \) **errors and** \( e \) **erasures**
  - \( \rightarrow \) if \( e + 2b < d_{\min} \)
  - \( \rightarrow \) in time \( \Theta(n \log^3 n) \).

In this talk,

\[
\text{RS}_q(r) := \text{RS}_q(r, \mathbb{F}_q)
\]
Goal: sublinear-time correction of some symbols of $c \in C$. 

**Definition [KT00]**. A code $C \subseteq \mathbb{F}_n^q$ is locally correctable with
- locality $\ell \leq n$,
- failure probability $\varepsilon \in (0, 1)$,
- admissible fraction of errors $\delta \in (0, 1)$,
if there exists a poly-time probabilistic algorithm $D$ such that, for every $y \in \mathbb{F}_n^q$ and $c \in C$ satisfying $d(y, c) \leq \delta n$ and for every $1 \leq i \leq n$:
- $\Pr(D(y)(i) = c_i) \geq 1 - \varepsilon$;
- $D(y)(i)$ makes at most $\ell$ queries to symbols of $y$.

$n = 16$, $\ell = 3$
Local correction [Katz, Trevisan ’00]

**Goal:** sublinear-time correction of some symbols of \( c \in C \).

**Definition [KT00].** A code \( C \subseteq \mathbb{F}_q^n \) is **locally correctable** with
- **locality** \( \ell \leq n \),
- failure probability \( \varepsilon \in (0, 1) \),
- admissible fraction of errors \( \delta \in (0, 1) \),

if there exists a poly-time **probabilistic algorithm** \( D \) such that, for every \( y \in \mathbb{F}_q^n \) and \( c \in C \) satisfying \( d(y, c) \leq \delta n \) and for every \( 1 \leq i \leq n \):
- \( \Pr(D(y)(i) = c_i) \geq 1 - \varepsilon \);
- \( D(y)(i) \) makes at most \( \ell \) queries to symbols of \( y \).

\( n = 16, \ell = 3 \)

\( \square \) = symbol to be corrected

\( \times \) = error

\( y \):

\[
\begin{array}{cccccccccc}
\text{error} & \text{symbol to be corrected} & \text{1} & \text{2} & \text{3} & \text{4} & \text{5} & \text{6} & \text{7} & \text{8} & \text{9} & \text{10} & \text{11} & \text{12} & \text{13} & \text{14} & \text{15} & \text{16}
\end{array}
\]
Goal: sublinear-time correction of some symbols of $c \in C$.

Definition [KT00]. A code $C \subseteq \mathbb{F}_q^n$ is **locally correctable** with

- **locality** $\ell \leq n$,
- failure probability $\varepsilon \in (0, 1)$,
- admissible fraction of errors $\delta \in (0, 1)$,

if there exists a poly-time **probabilistic algorithm** $D$ such that, for every $y \in \mathbb{F}_q^n$ and $c \in C$ satisfying $d(y, c) \leq \delta n$ and for every $1 \leq i \leq n$:

- $\Pr(D(y)(i) = c_i) \geq 1 - \varepsilon$;
- $D(y)(i)$ makes at most $\ell$ queries to symbols of $y$. 

$(n = 16, \ell = 3)$

\(\times\) = error

\(\square\) = symbol to be corrected

$y :$ \begin{array}{cccccccc}
\times & \square & \square & \square & \times & \square & \square & \square \\
\end{array}

\(i\)
Local correction [Katz, Trevisan ’00]

**Goal:** sublinear-time correction of some symbols of $c \in C$.

**Definition [KT00].** A code $C \subseteq \mathbb{F}_q^n$ is **locally correctable** with
- locality $\ell \leq n$,
- failure probability $\varepsilon \in (0, 1)$,
- admissible fraction of errors $\delta \in (0, 1)$,
if there exists a poly-time probabilistic algorithm $D$ such that, for every $y \in \mathbb{F}_q^n$ and $c \in C$ satisfying $d(y, c) \leq \delta n$ and for every $1 \leq i \leq n$:
  - $\Pr(D(y)(i) = c_i) \geq 1 - \varepsilon$;
  - $D(y)(i)$ makes at most $\ell$ queries to symbols of $y$.

(n = 16, $\ell = 3$)

$\Box$ = error  
□ = symbol to be corrected

$y : \quad \Box \quad \Box \quad \Box \quad \Box \quad \Box \quad \Box \quad \Box \quad i \quad \Box$
**Goal:** sublinear-time correction of some symbols of \( c \in C \).

**Definition [KT00].** A code \( C \subseteq \mathbb{F}_q^n \) is **locally correctable** with

- **locality** \( \ell \leq n \),
- failure probability \( \varepsilon \in (0, 1) \),
- admissible fraction of errors \( \delta \in (0, 1) \),

if there exists a poly-time **probabilistic algorithm** \( D \) such that, for every \( y \in \mathbb{F}_q^n \) and \( c \in C \) satisfying \( d(y, c) \leq \delta n \) and for every \( 1 \leq i \leq n \):

- \( \Pr(D(y)(i) = c_i) \geq 1 - \varepsilon \);
- \( D(y)(i) \) makes at most \( \ell \) queries to symbols of \( y \).

\((n = 16, \ell = 3)\)
\[\text{⊗} = \text{error}\]
\[\text{□} = \text{symbol to be corrected}\]
Goals:

- failure probability \( \varepsilon \leq f(\ell) \cdot \delta \), with \( f(\ell) \leq \text{cst} \).
- locality \( \ell \ll k \)
- large dimension \( k \)
LCCs: goals and previous works

Goals:
- failure probability $\varepsilon \leq f(\ell) \cdot \delta$, with $f(\ell) \leq \text{cst}.$
- locality $\ell \ll k$
- large dimension $k$

Some existing constructions:

- constant locality $\ell$:
  - Hadamard code (folklore) $\ell = 2$ and $k = \log(n)$
  - Matching vector codes [Yek08] $k$ subexponential in $\log(n)$

- constant rate $R = k/n$:
  - Reed-Muller codes (folklore) $\ell = n^{1/m}$ and $k \leq \frac{1}{m!} \cdot n$
  - Multiplicity codes [KSY14], \textit{lifted codes} [GKS13],
    expander codes [HOW14]
    \[ \ell \leq n^\varepsilon \quad \text{and} \quad k \geq \alpha \cdot n, \quad \forall \varepsilon, \alpha > 0, \quad n \to \infty \]
Example: Reed-Muller codes

\[ \text{RM}_q(m, r) := \{ (f(x) : x \in \mathbb{F}_q^m), f \in \mathbb{F}_q[X_1, \ldots, X_m], \deg f \leq r \} \]
Example: Reed-Muller codes

\[ \text{RM}_q(m, r) := \{ (f(x) : x \in \mathbb{F}_q^m), f \in \mathbb{F}_q[X_1, \ldots, X_m], \text{deg} f \leq r \} \]

Assume \( r \leq q - 2 \), and let:

- \( c = (f(x) : x \in \mathbb{F}_q^m) \in \text{RM}_q(m, r) \)
- \( \phi : \mathbb{F}_q \rightarrow \mathbb{F}_q^m \) affine and injective
  \[ \Rightarrow \text{affine line } L := \phi(\mathbb{F}_q) \subset \mathbb{F}_q^m \]

Then, the restriction of \( c \) to \( L \) (or to \( \phi \)):

\[ c|_L := ((f \circ \phi)(t) : t \in \mathbb{F}_q) \in \text{RS}_q(r) \]
Example: Reed-Muller codes

\[ \text{RM}_q(m, r) := \{ (f(x) : x \in \mathbb{F}_q^m), f \in \mathbb{F}_q[X_1, \ldots, X_m], \deg f \leq r \} \]

**Assume** \( r \leq q - 2 \), and let:

- \( c = (f(x) : x \in \mathbb{F}_q^m) \in \text{RM}_q(m, r) \)
- \( \phi : \mathbb{F}_q \to \mathbb{F}_q^m \) affine and injective
  \( \Rightarrow \) affine line \( L := \phi(\mathbb{F}_q) \subset \mathbb{F}_q^m \)

Then, the **restriction** of \( c \) to \( L \) (or to \( \phi \)):

\[ c|_L := ((f \circ \phi)(t) : t \in \mathbb{F}_q) \in \text{RS}_q(r) \]

**Local correction of** \( y \in \mathbb{F}_q^m \) **at coordinate** \( i \in \mathbb{F}_q^m \):

1. Pick at random a line \( L \subset \mathbb{F}_q^m \) such that \( i \in L \).
2. Correct \( y|_L \) as a noisy \( \text{RS}_q(r) \) codeword, and output \( \tilde{y}_i \).
Example: Reed-Muller codes

RM$_q(m, r) := \{ (f(x) : x \in \mathbb{F}_q^m), f \in \mathbb{F}_q[X_1, \ldots, X_m], \deg f \leq r \}$

Assume $r \leq q - 2$, and let:
- $c = (f(x) : x \in \mathbb{F}_q^m) \in$ RM$_q(m, r)$
- $\phi : \mathbb{F}_q \to \mathbb{F}_q^m$ affine and injective
  $\Rightarrow$ affine line $L := \phi(\mathbb{F}_q) \subset \mathbb{F}_q^m$

Then, the restriction of $c$ to $L$ (or to $\phi$):

$c|_L := ((f \circ \phi)(t) : t \in \mathbb{F}_q) \in$ RS$_q(r)$

Local correction of $y \in \mathbb{F}_q^m$ at coordinate $i \in \mathbb{F}_q^m$:
1. Pick at random a line $L \subset \mathbb{F}_q^m$ such that $i \in L$.
2. Correct $y|_L$ as a noisy RS$_q(r)$ codeword, and output $\tilde{y}_i$.

RM$_q(m, r)$ is locally correctable with $\ell = n^{1/m}$ and $\varepsilon = \frac{2}{1-r/q} \cdot \delta$
**Issue:** if $r \leq q - 2$, the rate of $\text{RM}_q(m, r)$ is $\simeq \frac{(r/q)^m}{m!}$. 

Idea: consider the set of all polynomials $f$ satisfying the “restriction property”: for every affine line $L$ given by $\phi$, 

$$(f \circ \phi)(t) : t \in F_q$$ 

is in $\text{RS}_q(r)$. Are there more polynomials than in $\text{RM}$ codes?

Example ($q = 4, m = 2, r = 2$).

$f(X, Y) = X^2Y^2 \in F_4[X, Y]$, hence $\deg(f) = 4 > 2$. Affine line $L$ given by $\phi(T) = (aT + b, cT + d)$.
High-rate construction: lifted codes (1)

**Issue:** if \( r \leq q - 2 \), the rate of \( \text{RM}_q(m, r) \) is \( \simeq \frac{(r/q)^m}{m!} \).

**Idea:** consider the set of all polynomials \( f \) satisfying the “restriction property”:
for every affine line \( L \) given by \( \phi \), \((f \circ \phi)(t) : t \in \mathbb{F}_q\) \( \in \text{RS}_q(r) \)

Are there more polynomials than in RM codes?
High-rate construction: lifted codes (1)

**Issue:** if $r \leq q - 2$, the rate of $\text{RM}_q(m, r)$ is $\approx \frac{(r/q)^m}{m!}$.

**Idea:** consider the set of all polynomials $f$ satisfying the “restriction property”: for every affine line $L$ given by $\phi$, $(f \circ \phi)(t) : t \in \mathbb{F}_q) \in \text{RS}_q(r)$

**Are there more polynomials than in RM codes?**

**Example** ($q = 4, m = 2, r = 2$).

$f(X, Y) = X^2Y^2 \in \mathbb{F}_4[X, Y]$, hence $\deg(f) = 4 > 2$

Affine line $L$ given by $\phi(T) = (aT + b, cT + d)$
High-rate construction: lifted codes (1)

**Issue:** if \( r \leq q - 2 \), the rate of \( \text{RM}_q(m, r) \) is \( \simeq \frac{(r/q)^m}{m!} \).

**Idea:** consider the set of all polynomials \( f \) satisfying the “restriction property”: for every affine line \( L \) given by \( \phi \), \( \{(f \circ \phi)(t) : t \in \mathbb{F}_q\} \in \text{RS}_q(r) \)

**Are there more polynomials than in RM codes?**

**Example** \((q = 4, m = 2, r = 2)\).

\( f(X, Y) = X^2Y^2 \in \mathbb{F}_4[X, Y] \), hence \( \deg(f) = 4 > 2 \)

Affine line \( L \) given by \( \phi(T) = (aT + b, cT + d) \)

\[(f \circ \phi)(T) = (aT + b)^2(cT + d)^2 \]

\[= (a^2T^2 + b^2)(c^2T^2 + d^2) \]

\[= (ac)^2T^4 + (ad + bc)^2T^2 + (bd)^2 \]
High-rate construction: lifted codes (1)

**Issue:** if \( r \leq q - 2 \), the rate of \( \text{RM}_q(m, r) \) is \( \simeq \frac{(r/q)^m}{m!} \).

**Idea:** consider the set of all polynomials \( f \) satisfying the “restriction property”: for every affine line \( L \) given by \( \phi \), \( ((f \circ \phi)(t) : t \in \mathbb{F}_q) \in \text{RS}_q(r) \)

**Are there more polynomials than in RM codes?**

**Example** \((q = 4, m = 2, r = 2)\).

\[
f(X, Y) = X^2 Y^2 \in \mathbb{F}_4[X, Y], \text{ hence } \deg(f) = 4 > 2
\]

Affine line \( L \) given by \( \phi(T) = (aT + b, cT + d) \)

\[
(f \circ \phi)(T) = (aT + b)^2(cT + d)^2 = (a^2T^2 + b^2)(c^2T^2 + d^2) = (ac)^2T^4 + (ad + bc)^2T^2 + (bd)^2 = (ad + bc)^2T^2 + (ac)^2T + (bd)^2 \mod (T^4 - T)
\]

\( \Rightarrow \) for every \( \phi \), the “restriction” \((f \circ \phi)(T)\) can be interpolated as a univariate polynomial of degree \( \leq 2 \).
High-rate construction: lifted codes (2)

- \( A^m := \mathbb{F}_q^m \quad \text{ev}_{A^m}(f) := (f(x) : x \in \mathbb{F}_q^m) \in \mathbb{F}_q^{A^m} \)

- \( \text{Emb}_{A}(m) := \{ \phi : \mathbb{F}_q \to \mathbb{F}_q^m, \text{injective and affine} \} \)

**Definition** (lifted Reed-Solomon code [GKS13] reformulated).

\[
\text{Lift}(\text{RS}_q(r), m) := \{ \text{ev}_{A^m}(f), f \in \mathbb{F}_q[X] \mid \forall \phi \in \text{Emb}_{A}(m), \text{ev}_{A^1}(f \circ \phi) \in \text{RS}_q(r) \}
\]
High-rate construction: lifted codes (2)

\[ A^m := \mathbb{F}_q^m \quad \text{ev}_{A^m}(f) := (f(x) : x \in \mathbb{F}_q^m) \in \mathbb{F}_q^{A^m} \]

\[ \text{Emb}_A(m) := \{ \phi : \mathbb{F}_q \to \mathbb{F}_q^m, \text{injective and affine} \} \]

**Definition** (lifted Reed-Solomon code [GKS13] reformulated).

\[
\text{Lift}(\text{RS}_q(r), m) := \{ \text{ev}_{A^m}(f), f \in \mathbb{F}_q[X] \mid \forall \phi \in \text{Emb}_A(m), \text{ev}_{A^1}(f \circ \phi) \in \text{RS}_q(r) \}
\]

\[
\text{Lift}(\text{RS}_q(r), m) \text{ is locally correctable with } \ell = n^{1/m} \text{ and } \varepsilon = \frac{2}{1-r/q} \cdot \delta.
\]
Definition (lifted Reed-Solomon code [GKS13] reformulated).

\[
\text{Lift}(\text{RS}_q(r), m) := \{ \text{ev}_{A^m}(f), f \in \mathbb{F}_q[X] \mid \forall \phi \in \text{Emb}_A(m), \text{ev}_A(f \circ \phi) \in \text{RS}_q(r) \}
\]

Lift(\text{RS}_q(r), m) is locally correctable with $\ell = n^{1/m}$ and $\varepsilon = \frac{2}{1-r/q} \cdot \delta$.

What about the dimension/rate?

Theorem (characteristic 2, simplified from [GKS13]).

For every $m \geq 2$ and $0 < R_0 < 1$, there exists $q > 0$ and $r \leq q - 2$ such that

\[
\text{Lift(\text{RS}_q(r), m)}
\]

is locally correctable with rate $R \geq R_0$. 
Bounds in [GKS13] are far from being tight.

- Ex: for $m = 2$ and $R_0 = 1/2$, GKS theorem requires $n = q^m \geq 2^{64}$.
Bounds in [GKS13] are **far from being tight**.

- **Ex:** for $m = 2$ and $R_0 = 1/2$, GKS theorem requires $n = q^m \geq 2^{64}$.

**Theorem** [characteristic 2, finite length $n = q^2 = 2^{2e}$].
For $m = 2$, $q = 2^e$ and $r = (1 - 2^{-c})q - 1$,

$$R = 1 - \frac{5}{4} \left( \frac{3}{4} \right)^c + \frac{1}{4} \left( \frac{1}{4} \right)^c + \frac{1}{2^e} \left( \frac{3^c - 1}{2^{c+2}} \right) .$$

- actually, $n = q^2 \geq 2^6 = 64$ is enough to achieve $R \geq 1/2$. 
Lifted codes are **monomial**, *i.e.* generated by evaluations of monomials

\[
ev_{\Lambda^m}(X_1^{d_1} \cdots X_m^{d_m}) = ev_{\Lambda^m}(X^d)
\]

**Degree set** of a monomial code [GKS13]:

\[
\text{Deg}(C) := \{ d \in [0, q - 1]^m, ev_{\Lambda^m}(X^d) \in C \}
\]
Lifted codes are **monomial**, *i.e.* generated by evaluations of monomials

\[ \text{ev}_{A^m}(X^{d_1} \ldots X^{d_m}) = \text{ev}_{A^m}(X^d) \]

**Degree set** of a monomial code [GKS13]:

\[ \text{Deg}(C) := \{ d \in [0, q - 1]^m, \text{ev}_{A^m}(X^d) \in C \} \]

A **representation** for \( m = 2 \):

- \( \text{RM}_4(2, 4) \)
- \( \text{RM}_4(2, 2) \)
- \( \text{Lift}(\text{RS}_4(2), 2) \)
"Fractal" representation of degree sets

$q = 16, r = 14$
$q = 8, r = 6$
$q = 4, r = 2$
1. Codes with locality
   - Locality in coding theory, examples
   - Lifted projective Reed-Solomon codes
   - A combinatorial point of view

2. Private information retrieval from transversal designs
   - Private information retrieval (PIR)
   - Transversal designs and codes
   - A new PIR construction
   - Instances

3. Proofs-of-retrievability

4. Conclusion
Projective space $\mathbb{P}^m := (\mathbb{A}^{m+1} \setminus \{0\}) / \sim$ where $a \sim b$ iff $\exists \lambda \in \mathbb{F}_q^\times, a = \lambda b$

Defining an evaluation map over $\mathbb{P}^m$ requires:

- \textbf{homogeneous} polynomials $f \in \mathbb{F}_q[X]^H_v$ of fixed degree $v$,
- to choose a \textbf{representative} for every $u \in \mathbb{P}^m$ (see [Lac86]):

\[
\begin{align*}
  u &= (0 : \cdots : 0 : 1 : * : \cdots : *) \in \mathbb{P}^m \\
  f(u) := f(0, \ldots, 0, 1, *, \ldots, *) \in \mathbb{F}_q \\
  \text{ev}_{\mathbb{P}^m}(f) := (f(u) : u \in \mathbb{P}^m) \in \mathbb{F}_q^{\mathbb{P}^m}
\end{align*}
\]
Example. Projective Reed-Solomon code:

\[ \text{PRS}_q(r) = \{ \text{ev}_{\mathbb{P}^1}(f) = (f(x) : x \in \mathbb{P}^1), f \in \mathbb{F}_q[X, Y]_r^H \} \]
Example. Projective Reed-Solomon code:

$$\text{PRS}_q(r) = \{ \text{ev}_{\mathbb{P}^1}(f) = (f(x) : x \in \mathbb{P}^1), f \in \mathbb{F}_q[X, Y]_r^H \}$$

Let $\text{Emb}_{\mathbb{P}}(m) := \{ \phi : \mathbb{F}_q^2 \to \mathbb{F}_q^{m+1} \text{ linear and injective} \}.

**Definition** (lifted projective RS codes). Let $v = r + (m - 1)(q - 1)$.

$$\text{Lift} (\text{PRS}_q(r), m) := \{ \text{ev}_{\mathbb{P}^m}(f), f \in \mathbb{F}_q[X]^H_v | \\
\forall \phi \in \text{Emb}_{\mathbb{P}}(m), \text{ev}_{\mathbb{P}^1}(f \circ \phi) \in \text{PRS}_q(r) \}$$
Main results on projective lifted codes

Projective lifted codes...

- are **locally correctable**, with parameters \((\ell = q + 1, \delta, \epsilon = \delta / \tau)\), where \(\tau\) is the relative correction capability of the small PRS code.
Main results on projective lifted codes

Projective lifted codes...

- are **locally correctable**, with parameters $(\ell = q + 1, \delta, \varepsilon = \delta / \tau)$, where $\tau$ is the relative correction capability of the small PRS code

- are **monomial**, with an **explicit bijection** between the degree sets of $\text{Lift}(\text{RS}_q(r - 1), m)$, $\text{Lift}(\text{PRS}_q(r), m)$ and $\text{Lift}(\text{PRS}_q(r), m - 1)$

Main results on projective lifted codes

Projective lifted codes...

- are **locally correctable**, with parameters $(\ell = q + 1, \delta, \varepsilon = \delta / \tau)$, where $\tau$ is the relative correction capability of the small PRS code
- are **monomial**, with an **explicit bijection** between the degree sets of $\text{Lift}(\text{RS}_q(r - 1), m)$, $\text{Lift}(\text{PRS}_q(r), m)$ and $\text{Lift}(\text{PRS}_q(r), m - 1)$
- satisfy the **puncturing/shortening** relation

\[ 0 \to \text{Lift}(\text{RS}_q(r - 1), m) \to \text{Lift}(\text{PRS}_q(r), m) \xrightarrow{\pi} \text{Lift}(\text{PRS}_q(r), m - 1) \to 0, \]

where $\pi$ is induced by $\mathbb{P}^m \to \mathbb{P}^{m-1}$.
Main results on projective lifted codes

Projective lifted codes...

- are **locally correctable**, with parameters \((\ell = q + 1, \delta, \epsilon = \delta / \tau)\), where \(\tau\) is the relative correction capability of the small PRS code
- are **monomial**, with an **explicit bijection** between the degree sets of 
  \(\text{Lift}(\text{RS}_q(r - 1), m)\), \(\text{Lift}(\text{PRS}_q(r), m)\) and \(\text{Lift}(\text{PRS}_q(r), m - 1)\)
- satisfy the **puncturing/shortening** relation

\[
0 \rightarrow \text{Lift}(\text{RS}_q(r - 1), m) \rightarrow \text{Lift}(\text{PRS}_q(r), m) \xrightarrow{\pi} \text{Lift}(\text{PRS}_q(r), m - 1) \rightarrow 0,
\]

where \(\pi\) is induced by \(\mathbb{P}^m \rightarrow \mathbb{P}^{m-1}\).

- are (up to equivalence)

  - **cyclic codes** if \(q - 1\) and \(n = \frac{q^{m+1}}{q-1}\) are coprime
  - **quasi-cyclic codes** if \(q - 1\) and \(\frac{n}{\gcd(n,q-1)}\) are coprime

Details in: 
Lifted Projective Reed-Solomon Codes, L., DCC, to appear
10.1007/s10623-018-0552-8
Main results on projective lifted codes

Projective lifted codes...

- are **locally correctable**, with parameters \((\ell = q + 1, \delta, \epsilon = \delta / \tau)\), where \(\tau\) is the relative correction capability of the small PRS code

- are **monomial**, with an explicit bijection between the degree sets of \(\text{Lift}(\text{RS}_{q}(r - 1), m)\), \(\text{Lift}(\text{PRS}_{q}(r), m)\) and \(\text{Lift}(\text{PRS}_{q}(r), m - 1)\)

- satisfy the **puncturing/shortening** relation

\[
0 \rightarrow \text{Lift}(\text{RS}_{q}(r - 1), m) \rightarrow \text{Lift}(\text{PRS}_{q}(r), m) \xrightarrow{\pi} \text{Lift}(\text{PRS}_{q}(r), m - 1) \rightarrow 0,
\]

where \(\pi\) is induced by \(\mathbb{P}^m \rightarrow \mathbb{P}^{m-1}\).

- are (up to equivalence)
  - **cyclic codes** if \(q - 1\) and \(n = \frac{q^{m+1}}{q-1}\) are coprime
  - **quasi-cyclic codes** if \(q - 1\) and \(\frac{n}{\gcd(n,q-1)}\) are coprime

- admit many explicit and easily computable **information sets**
Main results on projective lifted codes

Projective lifted codes...

- are **locally correctable**, with parameters $(\ell = q + 1, \delta, \epsilon = \delta / \tau)$, where $\tau$ is the relative correction capability of the small PRS code
- are **monomial**, with an **explicit bijection** between the degree sets of $\text{Lift}(\text{RS}_q(r-1), m)$, $\text{Lift}(\text{PRS}_q(r), m)$ and $\text{Lift}(\text{PRS}_q(r), m-1)$
- satisfy the **puncturing/shortening** relation

\[
0 \rightarrow \text{Lift}(\text{RS}_q(r-1), m) \rightarrow \text{Lift}(\text{PRS}_q(r), m) \xrightarrow{\pi} \text{Lift}(\text{PRS}_q(r), m-1) \rightarrow 0,
\]

where $\pi$ is induced by $\mathbb{P}^m \rightarrow \mathbb{P}^{m-1}$.

- are (up to equivalence)
  - **cyclic codes** if $q - 1$ and $n = \frac{q^{m+1}}{q-1}$ are coprime
  - quasi-cyclic codes if $q - 1$ and $\frac{n}{\gcd(n, q-1)}$ are coprime

- admit many explicit and easily computable **information sets**

Details in: 

* Lifted Projective Reed-Solomon Codes, L., DCC, to appear 10.1007/s10623-018-0552-8
1. Codes with locality
   Locality in coding theory, examples
   Lifted projective Reed-Solomon codes
   A combinatorial point of view

2. Private information retrieval from transversal designs
   Private information retrieval (PIR)
   Transversal designs and codes
   A new PIR construction
   Instances

3. Proofs-of-retrievability

4. Conclusion
Lifted codes when $r = q - 2$

**Remark.** Assume $r = q - 2$. Then, $RS_q(q - 2)$ is the parity-check code.

$$a \in RS_q(q - 2) \iff \sum_{i=1}^{q} a_i = 0$$

$$c \in \text{Lift}(RS_q(q - 2), m) \iff \forall L \subseteq \mathbb{F}_q^m, \sum_{x \in L} c_x = 0$$
Lifted codes when $r = q - 2$

**Remark.** Assume $r = q - 2$. Then, $\text{RS}_q(q - 2)$ is the parity-check code.

$$a \in \text{RS}_q(q - 2) \iff \sum_{i=1}^{q} a_i = 0$$

$$c \in \text{Lift}(\text{RS}_q(q - 2), m) \iff \forall L \subseteq \mathbb{F}_q^m, \sum_{x \in L} c_x = 0$$

A non-full-rank **parity-check matrix** for $\text{Lift}(\text{RS}_q(q - 2), m)$:

$$\begin{pmatrix}
\ast \\
0 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0
\end{pmatrix} \leftarrow \text{indicator vector of line } L$$
Point-line incidences in the affine space form the **affine geometry 2-design**.
Point-line incidences in the affine space form the **affine geometry 2-design**.

**Definition.** A *t-design* of parameters \((n, \ell, \lambda)\) consists in:
- a set \(X\) of points, \(|X| = n\),
- a set \(\mathcal{B}\) of blocks \(B \subseteq X, |B| = \ell\)
such that every \(t\)-set in \(X\) appears in exactly \(\lambda\) blocks.
Point-line incidences in the affine space form the **affine geometry 2-design**.

**Definition.** A *t*-design of parameters \((n, \ell, \lambda)\) consists in:

- a set \(X\) of points, \(|X| = n\),
- a set \(B\) of blocks \(B \subset X, |B| = \ell\)

such that every \(t\)-set in \(X\) appears in exactly \(\lambda\) blocks.

**Incidence matrix of a design:**

\[
\begin{pmatrix}
    \ast & 0 & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 \\
    0 & \ast & \cdots & 0 & 1 & \cdots & 1 & 0 & \cdots & 0 \\
\end{pmatrix}
\]

\(\leftarrow\) indicator vector of block \(B\)
The code based on the design $\mathcal{D} = (X, \mathcal{B})$ is the code $\mathcal{C} = \text{Code}(\mathcal{D}) \subseteq \mathbb{F}_q^X$ admitting the incidence matrix of $\mathcal{D}$ as a parity-check matrix.

$$\text{Code}(\mathcal{D}) = \{ c \in \mathbb{F}_q^X \mid \forall B \in \mathcal{B}, c|_B \in \text{Parity} \}$$

**Remark.** The dimension of Code$(\mathcal{D})$ is highly dependent on the field $\mathbb{F}_q$
The code based on the design $\mathcal{D} = (X, \mathcal{B})$ is the code $\mathcal{C} = \text{Code}(\mathcal{D}) \subseteq \mathbb{F}_q^X$ admitting the incidence matrix of $\mathcal{D}$ as a parity-check matrix.

$$\text{Code}(\mathcal{D}) = \{ c \in \mathbb{F}_q^X \mid \forall B \in \mathcal{B}, c|_B \in \text{Parity} \}$$

Remark. The dimension of $\text{Code}(\mathcal{D})$ is highly dependent on the field $\mathbb{F}_q$

Let $\mathcal{F} = (\mathcal{F}_B \subseteq \mathbb{F}_q^B : B \in \mathcal{B})$ be a family of codes indexed by blocks $B \in \mathcal{B}$. The generalised design-based code based on $(\mathcal{D}, \mathcal{F})$ is

$$\text{Code}(\mathcal{D}, \mathcal{F}) := \{ c \in \mathbb{F}_q^X \mid \forall B \in \mathcal{B}, c|_B \in \mathcal{F}_B \}.$$
**Generalised design-based code** $\mathcal{C} = \text{Code}(\mathcal{D}, \mathcal{F})$, where

- $\mathcal{D}$ be a $t$-$(n, \ell + 1, \lambda)$-design
- $\tau \in (0, \frac{1}{2})$ is fixed
- $\mathcal{F} = (\mathcal{F}_B : B \in \mathcal{B})$ s.t. every code in $\mathcal{F}$ corrects a fraction $\tau$ of errors

**Algorithm.** Local correction of $y \in \mathbb{F}_q^X$ at $i \in X$

- Pick uniformly at random a block $B \in \mathcal{B}$ such that $i \in B$.
- Correct $y|_B$ as a noisy codeword from $\mathcal{F}_B$, and output $\tilde{y}_i$.

**Proposition** $[t = 2]$. For every $\delta < \tau / 2$, $\text{Code}(\mathcal{D}, \mathcal{F})$ is a $(\ell, \delta, \varepsilon)$-LCC, where

$$\varepsilon = \frac{\delta}{\tau}.$$ 

**Proposition** $[t = 3]$. For every $\delta < \tau - 1 / \sqrt{2\ell}$, $\text{Code}(\mathcal{D}, \mathcal{F})$ is a $(\ell, \delta, \varepsilon)$-LCC where

$$\varepsilon = \frac{\delta(1 - \delta)}{(\tau - \delta)^2} \cdot \frac{1}{\ell} \leq \frac{1}{\tau^2 \ell} \cdot \delta.$$
1. Codes with locality
   Locality in coding theory, examples
   Lifted projective Reed-Solomon codes
   A combinatorial point of view

2. Private information retrieval from transversal designs
   Private information retrieval (PIR)
   Transversal designs and codes
   A new PIR construction
   Instances

3. Proofs-of-retrievability

4. Conclusion
1. Codes with locality
   Locality in coding theory, examples
   Lifted projective Reed-Solomon codes
   A combinatorial point of view

2. Private information retrieval from transversal designs
   Private information retrieval (PIR)
   Transversal designs and codes
   A new PIR construction
   Instances

3. Proofs-of-retrievability

4. Conclusion
Problem statement

Given a remote database $F \in \mathbb{F}_q^k$ and $1 \leq i \leq k$, can we \textbf{retrieve} the entry $F_i$, \textbf{without leaking} information on the index $i$?
Given a remote database $F \in \mathbb{F}_q^k$ and $1 \leq i \leq k$, can we retrieve the entry $F_i$, without leaking information on the index $i$?

**Trivial solution:** full download.
Problem statement

Given a remote database $F \in \mathbb{F}_q^k$ and $1 \leq i \leq k$, can we retrieve the entry $F_i$, without leaking information on the index $i$?

Trivial solution: full download.

Solutions with better communication complexity:

- With 1 server, only computational privacy is possible [CGKS95, CG97].
- With $\ell \geq 2$ servers, one can achieve information-theoretic privacy [CGKS95-98].
Given a file $F$ and $\ell$ servers $S_1, \ldots, S_\ell$, user $U$ wants to recover $F_i$ privately.

A Private Information Retrieval protocol is a set of algorithms $(Q, A, R)$:
Definition of PIR [CGKS95]

Given a file $F$ and $\ell$ servers $S_1, \ldots, S_\ell$, user $U$ wants to recover $F_i$ privately.

A Private Information Retrieval protocol is a set of algorithms $(Q, A, R)$:

1. $U$ generates a query vector $q = (q_1, \ldots, q_\ell) \leftarrow Q(i)$ and sends $q_j$ to server $S_j$
Definition of PIR [CGKS95]

Given a file $F$ and $\ell$ servers $S_1, \ldots, S_\ell$, user $U$ wants to recover $F_i$ privately.

A **Private Information Retrieval protocol** is a set of algorithms $(Q, A, R)$:

1. $U$ generates a query vector $q = (q_1, \ldots, q_\ell) \leftarrow Q(i)$ and sends $q_j$ to server $S_j$
2. Each server $S_j$ computes $a_j = A(q_j, F|_{S_j})$ and sends it back to $U$
Definition of PIR [CGKS95]

Given a file $F$ and $\ell$ servers $S_1, \ldots, S_\ell$, user $U$ wants to recover $F_i$ privately.

A Private Information Retrieval protocol is a set of algorithms $(Q, A, R)$:

1. $U$ generates a query vector $q = (q_1, \ldots, q_\ell) \leftarrow Q(i)$ and sends $q_j$ to server $S_j$
2. Each server $S_j$ computes $a_j = A(q_j, F|_{S_j})$ and sends it back to $U$
3. $U$ recovers $F_i = R(q, a, i)$
Definition of PIR [CGKS95]

Given a file $F$ and $\ell$ servers $S_1, \ldots, S_\ell$, user $U$ wants to recover $F_i$ privately.

A Private Information Retrieval protocol is a set of algorithms $(Q, A, R)$:

1. $U$ generates a query vector $q = (q_1, \ldots, q_\ell) \leftarrow Q(i)$ and sends $q_j$ to server $S_j$
2. Each server $S_j$ computes $a_j = A(q_j, F|_{S_j})$ and sends it back to $U$
3. $U$ recovers $F_i = R(q, a, i)$

Information-theoretic privacy: $I(i; q_j) = 0$, $\forall j = 1, \ldots, \ell$. 
Motivation

Usual goals for PIR:

- Low communication complexity
- Low storage overhead for the servers
- Low computation complexity for algorithms $A$ (server) and $R$ (user)
Usual goals for PIR:

- Low communication complexity
- Low storage overhead for the servers
- Low computation complexity for algorithms $A$ (server) and $R$ (user)

Most constructions focus on the download communication complexity

- up to the **PIR capacity** [SJ17]
- but require $\Omega(k)$ computation complexity for each server
Usual goals for PIR:

- Low communication complexity
- Low storage overhead for the servers
- Low computation complexity for algorithms $A$ (server) and $R$ (user)

Most constructions focus on the download communication complexity
- up to the PIR capacity [SJ17]
- but require $\Omega(k)$ computation complexity for each server

We here focus on the computation complexity, crucial for practicality [OG10].
1. Codes with locality
   Locality in coding theory, examples
   Lifted projective Reed-Solomon codes
   A combinatorial point of view

2. Private information retrieval from transversal designs
   Private information retrieval (PIR)
   Transversal designs and codes
   A new PIR construction
   Instances

3. Proofs-of-retrievability

4. Conclusion
A transversal design $\text{TD}(\ell, s) = (X, \mathcal{B}, \mathcal{G})$ is given by:

- $X$ a set of points, $|X| = n = s\ell$,
A transversal design $\text{TD}(\ell, s) = (X, \mathcal{B}, \mathcal{G})$ is given by:

- $X$ a set of points, $|X| = n = s\ell$,
- groups $\mathcal{G} = \{G_j\}_{1 \leq j \leq \ell}$ satisfying
  \[ X = \bigsqcup_{j=1}^{\ell} G_j \text{ and } |G_j| = s, \]

where $\mathcal{B}$ is a set of blocks $B \in \mathcal{B}$ satisfying:

- each block contains exactly $\ell$ points,
- for all $\{i, j\} \subset X$, $\{i, j\}$ lie in a single group $G \in \mathcal{G}$, or in a unique block $B \in \mathcal{B}$.
A transversal design \( \text{TD}(\ell, s) = (X, \mathcal{B}, \mathcal{G}) \) is given by:

- \( X \) a set of points, \( |X| = n = s\ell \),
- \( \textit{groups} \ \mathcal{G} = \{G_j\}_{1 \leq j \leq \ell} \) satisfying
  \[
  X = \bigsqcup_{j=1}^{\ell} G_j \quad \text{and} \quad |G_j| = s,
  \]
- \( \textit{blocks} \ \mathcal{B} \) satisfying
  - \( B \subseteq X \) and \( |B| = \ell \);
  - for all \( \{i, j\} \subseteq X \), \( \{i, j\} \) lie:
    - \textit{either} in a single group \( G \in \mathcal{G} \),
    - \textit{or} in a unique block \( B \in \mathcal{B} \).
A transversal design $TD(\ell, s) = (X, B, \mathcal{G})$ is given by:

- $X$ a set of points, $|X| = n = s\ell$,
- groups $\mathcal{G} = \{G_j\}_{1 \leq j \leq \ell}$ satisfying
  
  \[
  X = \bigsqcup_{j=1}^{\ell} G_j \quad \text{and} \quad |G_j| = s,
  \]

- blocks $B \in \mathcal{B}$ satisfying
  
  - $B \subset X$ and $|B| = \ell$;
  - for all $\{i, j\} \subset X$, $\{i, j\}$ lie:
    
    either in a single group $G \in \mathcal{G}$,
    
    or in a unique block $B \in \mathcal{B}$

Its incidence matrix (between points and blocks) defines a code.
The transversal design $\text{TD}(3, 3)$ represented by:

$$
\begin{align*}
G_1 & \quad G_2 & \quad G_3 \\
\bullet & \quad \bullet & \quad \bullet \\
\bullet & \quad \bullet & \quad \bullet \\
\bullet & \quad \bullet & \quad \bullet \\
\end{align*}
\text{with}
B = B_1 \cup B_2 \cup B_3
$$

gives an incidence matrix

$$
H = \begin{pmatrix}
1 & 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\
1 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \\
0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 0 \\
1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\
0 & 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 \\
0 & 0 & 1 & 1 & 0 & 0 & 0 & 1 & 0 \\
\end{pmatrix}
$$

Its rank over $\mathbb{F}_3$ is 6 $\implies$ the associated code $C$ is a $[9, 3]_3$ code.
1. Codes with locality
   Locality in coding theory, examples
   Lifted projective Reed-Solomon codes
   A combinatorial point of view

2. Private information retrieval from transversal designs
   Private information retrieval (PIR)
   Transversal designs and codes
   A new PIR construction
   Instances

3. Proofs-of-retrievability

4. Conclusion
Let $C \subseteq \mathbb{F}_q^n$ be a code based on a TD($\ell, s$).
Let $C \subseteq \mathbb{F}_q^N$ be a code based on a TD($\ell,s$).

- **Initialisation.** User $U$ encodes $F \mapsto c \in C$, and gives $c_{|G_j}$ to server $S_j$. 

1. User $U$ randomly picks a block $B \in B$ containing $i$.
   - Then $U$ defines: $q_j = Q(i)_j = \begin{cases} 
   \text{unique} & \text{if } i \not\in G_j \\
   \text{random point in } G_j & \text{otherwise.} 
   \end{cases}$
2. Each server $S_j$ sends back $c_{q_j}$.
3. $U$ recovers $c_i = - \sum_{j: i \not\in G_j} c_{q_j} = - \sum_{b \in B \setminus \{i\}} c_b$. 

---

The PIR scheme
Let $C \subseteq \mathbb{F}_q^n$ be a code based on a TD($\ell, s$).

- **Initialisation.** User $U$ encodes $F \mapsto c \in C$, and gives $c|_{G_j}$ to server $S_j$.

- **To recover** $F_i = c_i$, with $i \in X$:
  1. User $U$ randomly picks a block $B \in B$ containing $i$. Then $U$ defines:
     
     $$q_j = Q(i)_j := \begin{cases} 
     \text{unique} \in B \cap G_j & \text{if } i \notin G_j \\
     \text{a random point in } G_j & \text{otherwise.}
     \end{cases}$$

  2. Each server $S_j$ sends back $c_{q_j}$

  3. $U$ recovers

     $$c_i = - \sum_{j: i \notin G_j} c_{q_j} = - \sum_{b \in B \setminus \{i\}} c_b$$
Theorem. This PIR protocol is information-theoretically private.

Proof:
- the only server which holds $F_i$ received a random query;
- for each other server $S_j$, query $q_j$ gives no information on the block $B$ which has been picked $\Rightarrow$ no information leaks on $i$. 
Theorem. This PIR protocol is information-theoretically private.

Proof:
- the only server which holds $F_i$ received a random query;
- for each other server $S_j$, query $q_j$ gives no information on the block $B$ which has been picked ⇒ no information leaks on $i$.

Features.
- communication complexity: $\ell \log s$ uploaded bits, $\ell \log q$ downloaded bits
- computational complexity:
  - only 1 read for each server (somewhat optimal)
  - $\leq \ell$ additions over $\mathbb{F}_q$ for the user
- storage overhead: $(n - k) \log q$ bits
Theorem. This PIR protocol is information-theoretically private.

Proof:
- the only server which holds $F_i$ received a random query;
- for each other server $S_j$, query $q_j$ gives no information on the block $B$ which has been picked $\Rightarrow$ no information leaks on $i$.

Features.
- communication complexity: $\ell \log s$ uploaded bits, $\ell \log q$ downloaded bits
- computational complexity:
  - only 1 read for each server (somewhat optimal)
  - $\leq \ell$ additions over $F_q$ for the user
- storage overhead: $(n - k) \log q$ bits

Question: transversal designs with good $k$ depending on $(\ell, s)$?
1. Codes with locality
   Locality in coding theory, examples
   Lifted projective Reed-Solomon codes
   A combinatorial point of view

2. Private information retrieval from transversal designs
   Private information retrieval (PIR)
   Transversal designs and codes
   A new PIR construction
   Instances

3. Proofs-of-retrievability

4. Conclusion
$\mathcal{T}_A$, the classical affine transversal design:

- $X = \mathbb{F}_q^m, m \geq 2$,
- $\mathcal{G}$ a set of $q$ disjoint hyperplanes partitionning $X$,
- $\mathcal{B} = \{\text{affine lines } L \text{ secant to each group of } \mathcal{G}\}$.
\( \mathcal{T}_A \), the classical affine transversal design:

- \( X = \mathbb{F}_q^m, m \geq 2, \)
- \( \mathcal{G} \) a set of \( q \) disjoint hyperplanes partitionning \( X \),
- \( \mathcal{B} = \{ \text{affine lines } L \text{ secant to each group of } \mathcal{G} \}. \)

**Proposition.** The code based on \( \mathcal{T}_A \) is identical to the code based on the affine geometry design (i.e. the lifted code with \( r = q - 2 \)).
Instances with geometric designs

\( \mathcal{T}_A \), the classical affine transversal design:

- \( X = \mathbb{F}_q^m, m \geq 2, \)
- \( \mathcal{G} \) a set of \( q \) disjoint hyperplanes partitioning \( X \),
- \( \mathcal{B} = \{ \text{affine lines } L \text{ secant to each group of } \mathcal{G} \} \).

**Proposition.** The code based on \( \mathcal{T}_A \) is identical to the code based on the affine geometry design (i.e. the lifted code with \( r = q - 2 \)).

Instances:
- 3.2% storage overhead if \( \#\text{entries} \leq (\#\text{servers})^2 \)
- 27% storage overhead if \( \#\text{entries} \leq (\#\text{servers})^3 \)
\( \mathcal{T}_A \), the **classical affine transversal design**:  

- \( X = \mathbb{F}_q^m, m \geq 2 \),  
- \( \mathcal{G} \) a set of \( q \) disjoint hyperplanes partitioning \( X \),  
- \( \mathcal{B} = \{ \text{affine lines } L \text{ secant to each group of } \mathcal{G} \} \).

**Proposition.** The code based on \( \mathcal{T}_A \) is identical to the code based on the affine geometry design (i.e., the lifted code with \( r = q - 2 \)).

Instances:
- 3.2% storage overhead if \#entries \( \leq (\#\text{servers})^2 \)
- 27% storage overhead if \#entries \( \leq (\#\text{servers})^3 \)

**Question:** better instances?
An **orthogonal array** $\text{OA}(t, \ell, s)$ of strength $t$ is a list $A$ of words

- of length $\ell$,
- over a finite set $S$, $|S| = s$,
- such that, for every $I \subset [1, \ell]$ of size $t$, $A|_I = S^t$.

Equivalently, an $\text{OA}(t, \ell, s)$ is a code $A \subset S^\ell$ with dual distance $t + 1$.

$$S = \{a, b\}$$

$$\text{OA}(2, 3, 2) = \begin{bmatrix}
    a & b & b \\
    b & b & a \\
    b & a & b \\
    a & a & a \\
\end{bmatrix}$$
An **orthogonal array** $\text{OA}(t, \ell, s)$ of strength $t$ is a list $A$ of words
- of length $\ell$,
- over a finite set $S$, $|S| = s$,
- such that, for every $I \subset [1, \ell]$ of size $t$, $A|_I = S^t$.

Equivalently, an $\text{OA}(t, \ell, s)$ is a code $A \subset S^\ell$ with dual distance $t + 1$.

**Construction $\text{OA} \rightarrow \text{TD}$:**
- $X = S \times [1, \ell]$
- $\mathcal{G} = \{S \times \{i\}, 1 \leq i \leq \ell\}$
An **orthogonal array** $OA(t, \ell, s)$ of strength $t$ is a list $A$ of words
- of length $\ell$,
- over a finite set $S$, $|S| = s$,
- such that, for every $I \subset [1, \ell]$ of size $t$, $A|_I = S^t$.

Equivalently, an $OA(t, \ell, s)$ is a code $A \subset S^\ell$ with dual distance $t + 1$.

**Construction $OA \rightarrow TD$:**

- $X = S \times [1, \ell]$
- $\mathcal{G} = \{S \times \{i\}, 1 \leq i \leq \ell\}$
- $B = \{(c_i, i), 1 \leq i \leq \ell\}, c \in OA$

$$OA(2, 3, 2) = \begin{bmatrix}
a & b & b \\
b & b & a \\
b & a & b \\
a & a & a \\
\end{bmatrix}$$

$(a, 1)$ $(a, 2)$ $(a, 3)$
$(b, 1)$ $(b, 2)$ $(b, 3)$
**Instances with orthogonal arrays**

An **orthogonal array** $\text{OA}(t, \ell, s)$ of strength $t$ is a list $A$ of words

- of length $\ell$,
- over a finite set $S$, $|S| = s$,
- such that, for every $I \subset [1, \ell]$ of size $t$, $A|_I = S^t$.

Equivalently, an $\text{OA}(t, \ell, s)$ is a code $A \subset S^\ell$ with dual distance $t + 1$.

**Construction $\text{OA} \to \text{TD}$:**

- $X = S \times [1, \ell]$
- $\mathcal{G} = \{S \times \{i\}, 1 \leq i \leq \ell\}$
- $\mathcal{B} = \{(c_i, i), 1 \leq i \leq \ell\}, c \in \text{OA}$

$S = \{a, b\}$

$\text{OA}(2, 3, 2) = \begin{bmatrix} a & b & b \\ b & b & a \\ b & a & b \\ a & a & a \end{bmatrix}$

$\begin{align*}
(a, 1) & \quad (a, 2) & \quad (a, 3) \\
(b, 1) & \quad (b, 2) & \quad (b, 3)
\end{align*}$
An orthogonal array $OA(t, \ell, s)$ of strength $t$ is a list $A$ of words
- of length $\ell$,
- over a finite set $S$, $|S| = s$,
- such that, for every $I \subset [1, \ell]$ of size $t$, $A|_I = S^t$.

Equivalently, an $OA(t, \ell, s)$ is a code $A \subset S^\ell$ with dual distance $t + 1$.

Construction $OA \rightarrow TD$:
- $X = S \times [1, \ell]$
- $G = \{S \times \{i\}, 1 \leq i \leq \ell\}$
- $B = \{\{(c_i, i), 1 \leq i \leq \ell\}, c \in OA\}$

$S = \{a, b\}$

$OA(2, 3, 2) = \begin{bmatrix}
a & b & b \\
b & b & a \\
b & a & b \\
a & a & a \\
\end{bmatrix}$

(a, 1)  (a, 2)  (a, 3)  (b, 1)  (b, 2)  (b, 3)
Resisting collusions

**Proposition.** For $t = 2$, an OA$(t, \ell, s)$ gives a TD$(\ell, s)$. 

Experimentally, for $t = 2$ and small $\ell$ and $s$, codes based on classical affine TDs have the largest dimension. For $t \geq 3$, we get TDs such that: for every $t$-set $T$ of points lying in $t$ different groups, there exists a unique block $B \in B$ such that $T \subset B$. ⇒ The PIR protocol resists $t - 1$ colluding servers.

▶ OAs with $t > 2$ exist (e.g. from Reed-Solomon codes)
▶ But associated TDs lead to codes with poor rates except for $t \ll \ell$

Details in: Private Information Retrieval from Transversal Designs, L., IEEE TIT, to appear 10.1109/TIT.2018.2861747
**Proposition.** For $t = 2$, an OA($t, \ell, s$) gives a TD($\ell, s$).

Experimentally, for $t = 2$ and small $\ell$ and $s$, codes based on classical affine TDs have the largest dimension.
Proposition. For $t = 2$, an OA($t, \ell, s$) gives a TD($\ell, s$).

Experimentally, for $t = 2$ and small $\ell$ and $s$, codes based on classical affine TDs have the largest dimension.

For $t \geq 3$, we get TDs such that:

for every $t$-set $T$ of points lying in $t$ different groups,
there exists a unique block $B \in \mathcal{B}$ such that $T \subset B$.

$\Rightarrow$ The PIR protocol resists $t - 1$ colluding servers.
Proposition. For $t = 2$, an OA($t, \ell, s$) gives a TD($\ell, s$).

Experimentally, for $t = 2$ and small $\ell$ and $s$, codes based on classical affine TDs have the largest dimension.

For $t \geq 3$, we get TDs such that:

for every $t$-set $T$ of points lying in $t$ different groups, there exists a unique block $B \in \mathcal{B}$ such that $T \subset B$.

$\Rightarrow$ The PIR protocol resists $t - 1$ colluding servers.

- OAs with $t > 2$ exist (e.g. from Reed-Solomon codes)
- But associated TDs lead to codes with poor rates except for $t \ll \ell$
Resisting collusions

**Proposition.** For $t = 2$, an OA$(t, \ell, s)$ gives a TD$(\ell, s)$.

Experimentally, for $t = 2$ and small $\ell$ and $s$, codes based on classical affine TDs have the largest dimension.

For $t \geq 3$, we get TDs such that:

for every $t$-set $T$ of points lying in $t$ different groups,
there exists a unique block $B \in \mathcal{B}$ such that $T \subset B$.

⇒ The PIR protocol resists $t - 1$ colluding servers.

- OAs with $t > 2$ exist (e.g. from Reed-Solomon codes)
- But associated TDs lead to codes with poor rates except for $t \ll \ell$

Details in:

*Private Information Retrieval from Transversal Designs*, L., IEEE TIT, to appear

10.1109/TIT.2018.2861747
Outline

1. Codes with locality
   Locality in coding theory, examples
   Lifted projective Reed-Solomon codes
   A combinatorial point of view

2. Private information retrieval from transversal designs
   Private information retrieval (PIR)
   Transversal designs and codes
   A new PIR construction
   Instances

3. Proofs-of-retrievability

4. Conclusion
**Issue:** a client wants to verify if a file stored on a server is still retrievable, with a low communication challenge-response protocol.

"can I get my file?"

A few bits

---

Proofs-of-retrievability [Juels, Kaliski ’07]
**Issue:** a client wants to verify if a file stored on a server is still retrievable, with a low communication challenge-response protocol.

"can I get my file?"

**Additional constraints:** unbounded-use, low client storage, low computation.
PoR with lifted codes

\[ C = \text{Lift}(\text{RS}_q(r), m) \]

**Assumption:** one can compute independent pseudo-random permutations

\[ \sigma_i^{(\kappa)} \in \mathcal{G}(\mathbb{F}_q), \quad 1 \leq i \leq n, \ \kappa \in \mathcal{K} \]

**Initialisation:**

- User picks \( \kappa \in \mathcal{K} \) at random
- File \( F \) is encoded and permuted as follows:

\[
F \rightarrow c \in C \rightarrow w = \sigma(c) = (\sigma_1^{(\kappa)}(c_1), \ldots, \sigma_n^{(\kappa)}(c_n)) \in \mathbb{F}_q^n
\]

- User stores \( \kappa \), server stores \( w \)

**Verification:**

- User picks a line \( L \subset \mathbb{F}_q^m \) at random and sends it to the server
- Server reads \( w|_L \) and sends it back to the user
- User accepts iff \( \sigma^{-1}(w|_L) \in \text{RS}_q(r) \)
Informal result (for the lifted code with $m = 2$):

For every $\varepsilon \leq \varepsilon_0 \simeq 1$, we have:

if the server answers correctly to a fraction $\geq 1 - \varepsilon$ of the challenges,
then with probability $\geq 1 - O\left(\frac{1}{n(\varepsilon_0 - \varepsilon)^2}\right)$ the file is extractable from the server.


This idea can be generalised to other codes such as design-based codes.
Details in: Generic Constructions of PoRs from Codes and Instantiations, L. & Levy-dit-Vehel submitted, 2018
Informal result (for the lifted code with $m = 2$):

For every $\varepsilon \leq \varepsilon_0 \simeq 1$, we have:

\begin{itemize}
  \item [if] the server answers correctly to a fraction $\geq 1 - \varepsilon$ of the challenges,
  \item [then] with probability $\geq 1 - \mathcal{O}\left(\frac{1}{n(\varepsilon_0 - \varepsilon)^2}\right)$ the file is extractable from the server.
\end{itemize}

Details in:

*New Proofs of Retrievability using Locally Decodable Codes*, L. & Levy-dit-Vehel
IEEE International Symposium on Information Theory, 2016
**Informal result** (for the lifted code with $m = 2$):

For every $\varepsilon \leq \varepsilon_0 \simeq 1$, we have:

*if* the server answers correctly to a fraction $\geq 1 - \varepsilon$ of the challenges,

*then* with probability $\geq 1 - \mathcal{O}\left(\frac{1}{n(\varepsilon_0 - \varepsilon)^2}\right)$ the file is extractable from the server.

Details in: 
*New Proofs of Retrievability using Locally Decodable Codes*, L. & Levy-dit-Vehel
IEEE International Symposium on Information Theory, 2016

This idea can be **generalised** to other codes such as design-based codes.

Details in: 
*Generic Constructions of PoRs from Codes and Instantiations*, L. & Levy-dit-Vehel
submitted, 2018
1. Codes with locality
   - Locality in coding theory, examples
   - Lifted projective Reed-Solomon codes
   - A combinatorial point of view

2. Private information retrieval from transversal designs
   - Private information retrieval (PIR)
   - Transversal designs and codes
   - A new PIR construction
   - Instances

3. Proofs-of-retrievability

4. Conclusion
Conclusion

- **Analysis** and **generalisation** of a family of high-rate locally correctable codes, namely **lifted Reed-Solomon codes**

- **Combinatorial formalism** for the construction of locally correctable codes, thanks to **block designs**

- Application to **private information retrieval (PIR)**

- Application to **proofs of retrievability (PoR)**
Future works

- **PIR with low server computation complexity**
  - 1 server read $\rightarrow$ constant/sublinear number of server reads

- Extend the **lifting** process to other geometric varieties
  - e.g. the Hermitian variety

- Design-based codes allow us to **remove probabilistic decoders** from a definition of locally correctable codes
  - “usual” combinatorial coding-theoretic version of LCCs
  - new constructions? new bounds?