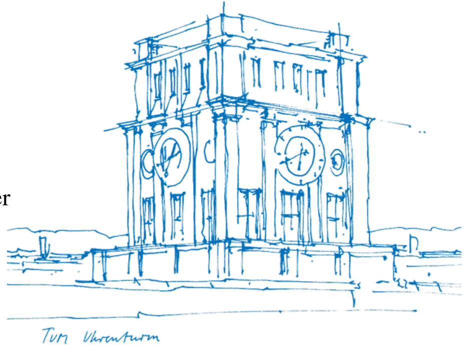
Interleaved Prange: A New Generic Decoder for Interleaved Codes

PQCrypto 2022

A. Porwal, L. Holzbaur, H. Liu, J. Renner, A. Wachter-Zeh, V. Weger

Technical University of Munich



Outline

Outline

Introduction

- McEliece system is a very promising candidate for post-quantum cryptography
- major drawback: large key size
- question: how can we do better?

Introduction

- potential solution: increase the error correction capability

  key (code) size $\uparrow$ $\longrightarrow$ error correction capability $\uparrow$ $\longrightarrow$ security level $\uparrow$

Introduction

- potential solution: increase the error correction capability

  key (code) size $\uparrow$ $\longrightarrow$ error correction capability $\uparrow$ $\longrightarrow$ security level $\uparrow$

- for example: use list decoding, **interleaving**, etc.

Interleaved Codes

- an ℓ-interleaved codeword is a concatenation of ℓ codewords from a constituent code C

$$\begin{bmatrix} \text{———} & c_1 & \text{———} \\ \text{———} & c_2 & \text{———} \\ \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots \\ \text{———} & c_\ell & \text{———} \end{bmatrix}$$

Interleaved Codes

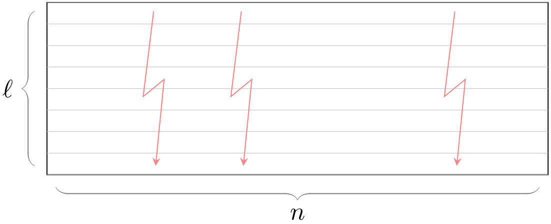
- an ℓ-interleaved codeword is a concatenation of ℓ codewords from a constituent code C

$$\begin{bmatrix} \text{———} & c_1 & \text{———} \\ \text{———} & c_2 & \text{———} \\ \vdots \; \vdots \; \vdots & \vdots & \vdots \; \vdots \; \vdots \\ \vdots \; \vdots \; \vdots & \vdots & \vdots \; \vdots \; \vdots \\ \text{———} & c_\ell & \text{———} \end{bmatrix}$$

- thus an ℓ-interleaved code is

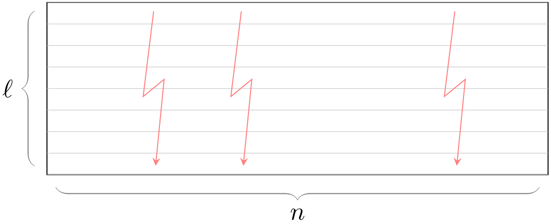
$$C_\ell = \left\{ \begin{bmatrix} c_1 \\ \vdots \\ c_\ell \end{bmatrix} : c_i \in C \right\}$$

Interleaved Codes



- interleaved decoders can correct up to t column errors
- here $\frac{d_{\min}-1}{2} < t < d_{\min}$ and typically t is close to $d_{\min}$

Interleaved Codes



- interleaved decoders can correct up to t column errors
- here $\frac{d_{\min}-1}{2} < t < d_{\min}$ and typically t is close to $d_{\min}$
- in particular: such decoders exist for interleaved Goppa codes

Interleaved Cryptosystems[1]

- Bob encodes his message matrix $\mathbf{M} \in \mathbb{F}_q^{\ell \times k}$ to get $\mathbf{M} \cdot \mathbf{G} = \mathbf{C} \in \mathbb{F}_q^{\ell \times n}$ (interleaved codeword)
- then the ciphertext is $\mathbf{R} = \mathbf{C} + \mathbf{E} \in \mathbb{F}_q^{\ell \times n}$ where $\mathbf{E}$ has column weight t
- Alice uses an interleaved Goppa decoder to decode $\mathbf{R}$

[1] Elleuch, Wachter-Zeh, and Zeh, "A Public-Key Cryptosystem from Interleaved Goppa Codes".

Interleaved Cryptosystems

SL [bits]	q	m	Method	r	n	k	t $(\ell, t_{\text{pub}}, d_{\text{E}})$	Rate	Key size [Bytes]
128	2	12	U. D.	70	2800	1960	70	0.70	205 800
	3	8	U. D.	100	2420	1620	75	0.67	256 763
			Int.		2130	1330	$(7, 131, 84)$	0.62	**210 800**
	4	6	U. D.	90	2150	1610	60	0.75	217 350
			Int.		1580	1040	$(7, 105, 76)$	0.66	**140 400**
	5	5	U. D.	100	1800	1380	62	0.74	200 266
			Int.		1290	790	$(7, 109, 84)$	0.61	**114 646**
256	2	13	U. D.	120	6740	5180	120	0.77	1 010 100
	3	8	U. D.	180	5100	3660	135	0.72	1 044 173
			Int.		4300	2860	$(7, 236, 156)$	0.67	**815 939**
	4	7	U. D.	240	4880	3200	160	0.66	1 344 000
			Int.		3760	2080	$(7, 280, 208)$	0.55	**873 600**
	5	6	U. D.	200	4690	3490	125	0.74	1 215 530
			Int.		3200	2000	$(7, 218, 171)$	0.63	**696 578**

Table 1[2]

[2] Holzbaur et al., "On Decoding and Applications of Interleaved Goppa Codes".

Interleaved Cryptosystems

SL [bits]	q	m	Method	r	n	k	t $(\ell, t_{\text{pub}}, d_E)$	Rate	Key size [Bytes]
128	2	12	U. D.	70	2800	1960	70	0.70	205 800
	3	8	U. D.	100	2420	1620	75	0.67	256 763
			Int.		2130	1330	$(7, 131, 84)$	0.62	**210 800**
	4	6	U. D.	90	2150	1610	60	0.75	217 350
			Int.		1580	1040	$(7, 105, 76)$	0.66	**140 400**
	5	5	U. D.	100	1800	1380	62	0.74	200 266
			Int.		1290	790	$(7, 109, 84)$	0.61	**114 646**
256	2	13	U. D.	120	6740	5180	120	0.77	1 010 100
	3	8	U. D.	180	5100	3660	135	0.72	1 044 173
			Int.		4300	2860	$(7, 236, 156)$	0.67	**815 939**
	4	7	U. D.	240	4880	3200	160	0.66	1 344 000
			Int.		3760	2080	$(7, 280, 208)$	0.55	**873 600**
	5	6	U. D.	200	4690	3490	125	0.74	1 215 530
			Int.		3200	2000	$(7, 218, 171)$	0.63	**696 578**

Table 1[2]

[2] Holzbaur et al., “On Decoding and Applications of Interleaved Goppa Codes”.

Hard Problem

Problem (Interleaved Decoding)

Given: $\mathbf{G} \in \mathbb{F}_q^{k \times n}$, $\mathbf{R} \in \mathbb{F}_q^{\ell \times n}$, *and* $t \in \mathbb{N}$

Find: is there an $\mathbf{E} \in \mathbb{F}_q^{\ell \times n}$ *of column weight at most* t*, such that each row of* $\mathbf{R} - \mathbf{E}$ *is in* $\langle \mathbf{G} \rangle$*?*

Goals

- understand generic decoding of interleaved codes

[3]Metzner and Kapturowski, “A General Decoding Technique Applicable to Replicated File Disagreement Location and Concatenated Code Decoding”.

Goals

- understand generic decoding of interleaved codes
  - important to assess security of interleaved cryptosystems

[3] Metzner and Kapturowski, “A General Decoding Technique Applicable to Replicated File Disagreement Location and Concatenated Code Decoding”.

Goals

- understand generic decoding of interleaved codes
  - important to assess security of interleaved cryptosystems
  - important also from a coding theoretic perspective:
    for $\ell \geq t$ (and full rank $\mathbf{E}$) there are efficient decoders for arbitrary linear constituent codes[3],
    but not true when $\ell \ll t$

[3] Metzner and Kapturowski, "A General Decoding Technique Applicable to Replicated File Disagreement Location and Concatenated Code Decoding".

Goals

- understand generic decoding of interleaved codes (when $\ell \ll t$)
  - important to assess security of interleaved cryptosystems
  - important also from a coding theoretic perspective:
    for $\ell \geq t$ (and full rank $\mathbf{E}$) there are efficient decoders for arbitrary linear constituent codes[3],
    but not true when $\ell \ll t$

[3] Metzner and Kapturowski, "A General Decoding Technique Applicable to Replicated File Disagreement Location and Concatenated Code Decoding".

Goals

- understand generic decoding of interleaved codes (when $\ell \ll t$)
  - important to assess security of interleaved cryptosystems
  - important also from a coding theoretic perspective:
    for $\ell \geq t$ (and full rank $\mathbf{E}$) there are efficient decoders for arbitrary linear constituent codes[3],
    but not true when $\ell \ll t$

- propose a new such generic decoder: interleaved Prange

[3] Metzner and Kapturowski, "A General Decoding Technique Applicable to Replicated File Disagreement Location and Concatenated Code Decoding".

Outline

Generic Decoding of Interleaved Codes

Three algorithms:

- SD-based: reduce the problem to the classical syndrome decoding (SD) problem
- CF-based: reduce the problem to a low-weight codeword finding (CF) problem
- a new algorithm: Interleaved Prange

Generic Decoding of Interleaved Codes

Reminder: our set-up is

$$\mathbf{C} = \begin{bmatrix} \text{———} & c_1 & \text{———} \\ \text{———} & c_2 & \text{———} \\ \vdots & \vdots & \vdots \\ \text{———} & c_\ell & \text{———} \end{bmatrix}$$ the (interleaved) codeword

$$\mathbf{E} = \begin{bmatrix} \text{———} & e_1 & \text{———} \\ \text{———} & e_2 & \text{———} \\ \vdots & \vdots & \vdots \\ \text{———} & e_\ell & \text{———} \end{bmatrix}$$ the error matrix which has only t non-zero columns

and the received word (the ciphertext) is $\mathbf{R} = \mathbf{C} + \mathbf{E}$

Generic Decoding of Interleaved Codes

Reminder: our set-up is

$$\mathbf{C} = \begin{bmatrix} \text{———} & c_1 & \text{———} \\ \text{———} & c_2 & \text{———} \\ \vdots & \vdots & \vdots \\ \text{———} & c_\ell & \text{———} \end{bmatrix}$$ the (interleaved) codeword

$$\mathbf{E} = \begin{bmatrix} \text{———} & e_1 & \text{———} \\ \text{———} & e_2 & \text{———} \\ \vdots & \vdots & \vdots \\ \text{———} & e_\ell & \text{———} \end{bmatrix}$$ the error matrix which has only t non-zero columns

and the received word (the ciphertext) is $\mathbf{R} = \mathbf{C} + \mathbf{E}$

We will be content with finding just a subset of the original t error positions

SD-based Algorithms

- pick non-zero vector from $\langle \mathbf{R} \rangle$ at random and solve the resulting SD problem
- most straightforward approach

SD-based Algorithms

- pick non-zero vector from $\langle \mathbf{R} \rangle$ at random and solve the resulting SD problem
- most straightforward approach
- information set decoding (ISD) attacks are one of the best known algorithms to solve the SD problem
- hence we call this approach Random $\langle$ISD$\rangle$ where $\langle$ISD$\rangle$ can be Prange, Stern, etc.

SD-based Algorithms: Random Prange

- for Random Prange, the success probability is

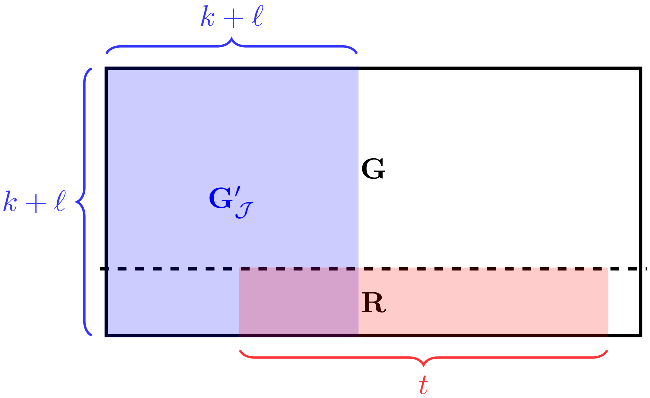
$$\sum_{v=0}^{t} \frac{\binom{t}{v}(q-1)^v}{q^t} \binom{n-k}{v} \binom{n}{v}^{-1}$$

- similarly, we can derive a expression for Random Stern

CF-based Algorithms

- note that the code generated by $\mathbf{G}' := \begin{bmatrix} \mathbf{G} \\ \mathbf{R} \end{bmatrix}$ is the same as the code generated by $\begin{bmatrix} \mathbf{G} \\ \mathbf{E} \end{bmatrix}$.
- thus the problem reduces to finding a low-weight codeword in the code $\langle \mathbf{G}' \rangle$ of dimension $k + \ell$.
- employ a CF-based algorithm (such as Stern's algorithm) to solve this problem

Interleaved Prange



is $\mathbf{G}'_{\mathcal{J}}$ rank-deficient?

Interleaved Prange

High-level description:

1. let $\mathbf{G}' := \begin{bmatrix} \mathbf{G} \\ \mathbf{R} \end{bmatrix}$
2. pick a set of $\mathcal{J}$ of column positions of size $k + \ell$
3. check if rank of $\mathbf{G}'_{\mathcal{J}}$ is less than $k + \ell$
4. if yes, search for an error-free vector in $\langle \mathbf{R} \rangle$ in the left null space of $\mathbf{G}'_{\mathcal{J}}$

Interleaved Prange

The work factor of interleaved Prange is $\frac{C}{P}$ where

$$P = \sum_{i=0}^{\min\{t,k+\ell\}} \frac{\binom{n-t}{k+\ell-i}\binom{t}{i}}{\binom{n}{k+\ell}} \cdot \left(1 - \prod_{j=0}^{\ell-1}(1 - q^{j-i})\right)$$

is the success probability

$$C \sim (k+\ell)^3 + 16 \prod_{j=0}^{k-1}(1 - q^{j-k}) \sum_{p=1}^{\ell} q^{-p^2+p}(k+\ell)(n-k-\ell)$$

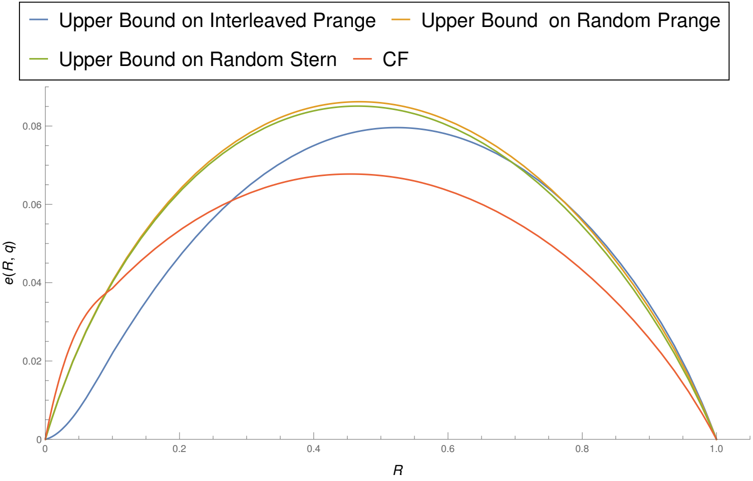
is the cost of one iteration

Outline

Comparison

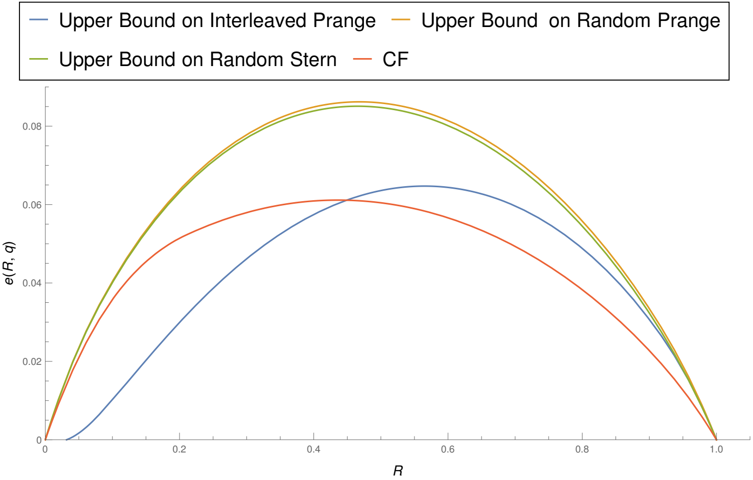
- we will do an asymptotic comparison
- some considerations for interleaved cryptosytems:
  - the greater the interleaving order ℓ, the closer t can be to d_{min}
  - but since the case $\ell \geq t$ can be efficiently decoded, we still want $\ell \ll t$

Comparison



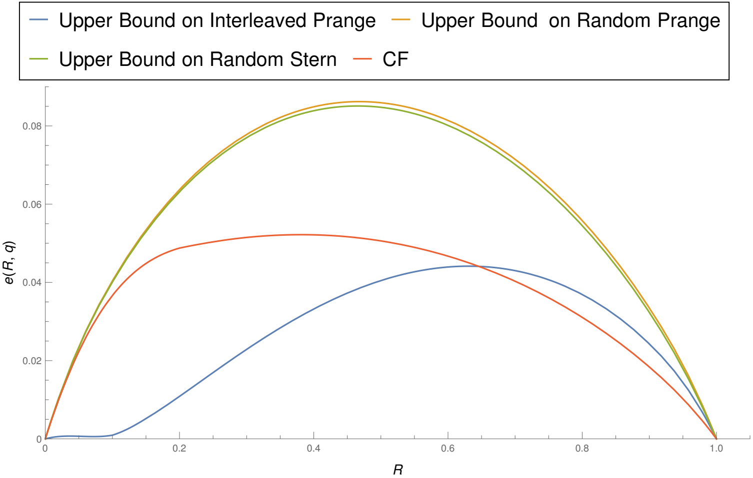
Asymptotic Cost for $q = 7$ and $\ell = t/20$

Comparison



Asymptotic Cost for $q = 7$ and $\ell = t/10$

Comparison



Asymptotic Cost for $q = 7$ and $\ell = t/5$

Comparison

ℓ	Algorithm	$e(R^*, q)$	R^*
$t/5$	Interleaved Prange (upper bound)	0.0441	0.631
	CF using Stern	0.0522	0.381
$t/10$	Interleaved Prange (upper bound)	0.06471	0.565
	CF using Stern	0.06114	0.436
$t/20$	Interleaved Prange (upper bound)	0.07961	0.524
	CF using Stern	0.06777	0.455

Maximum asymptotic cost $e(R^*, q = 7)$ with maximum at rate $R = R^*$

Outline

Conclusion

We looked at:

- how interleaved cryptosystems can be promising variant for code-based crypto
- three different algorithms for generic decoding of interleaved codes
- a new algorithm: Interleaved Prange

Interleaved Prange:

- asymptotically beats CF-based Stern in certain paramater ranges
- technique might also be applicable to decoders other than Prange