BIKE

Bit-flipping Key Encapsulation

NIST PQC Seminars
September 27th, 2024

Nicolas Aragon, University of Limoges, France
Paulo L. Barreto, University of Washington Tacoma, USA
Slim Bettaieb, TII, UAE
Loïc Bidoux, TII, UAE
Olivier Blazy, Ecole Polytechnique, France
Jean-Christophe Deneuville, ENAC, Federal University of Toulouse, France
Philippe Gaborit, University of Limoges, France
Santosh Ghosh, Nvidia, USA
Shay Gueron, University of Haifa, and Meta, Israel & USA
Tim Güneysu, Ruhr-Universität Bochum & DFKI, Germany
Carlos Aguilar Melchor, University of Toulouse, France
Rafael Misoczki, Meta, USA
Edoardo Persichetti, Florida Atlantic University, USA
Jan Richter-Brockmann, Ruhr-Universität Bochum, Germany
Nicolas Sendrier, INRIA, France
Jean-Pierre Tillich, INRIA, France
Valentin Vasseur, Thales, France
Gilles Zémor, IMB, University of Bordeaux, France

https://bikesuite.org

Agenda

- BIKE Overview
  - Key Features
  - Performance
  - Specification
  - Security
  - Modes of usage
- BIKE Decoding
  - Improved Results

Key Features

- Competitive performance
  - "*BIKE has **the most competitive performance** among the non-lattice based KEMs*" [NIST-IR8413]
- Robust security
  - Based on **well-known** coding-theory problems
  - Foundational cryptosystem was proposed **more than a decade ago** [MTSB'13]
- Simple implementation
  - Polynomial ring arithmetic
  - Bit flipping decoding

BIKE Performance

Security Level 1

	Communication Bandwidth (bytes)		Speed (kcycles)		
	Public Key	Ciphertext	Key Generation	Encapsulation	Decapsulation
BIKE	1,540	1,572	589	97	1,135
HQC	2,249	4,497	187	419	833
Classic McEliece	261,120	128	140,870	46	137
Kyber	800	768	123	155	289

BIKE performance numbers from Drucker, Gueron, Kostic, "Additional implementation of BIKE (Bit Flipping Key Encapsulation)". https://github.com/awslabs/bike-kem.

BIKE Performance

Security Level 3

	Communication Bandwidth (bytes)		Speed (kcycles)		
	Public Key	Ciphertext	Key Generation	Encapsulation	Decapsulation
BIKE	3,082	3,114	1,823	223	3,887
HQC	4,522	9,042	422	946	1,662
Classic McEliece	524,160	188	441,517	83	273
Kyber	1,184	1,088	213	249	275

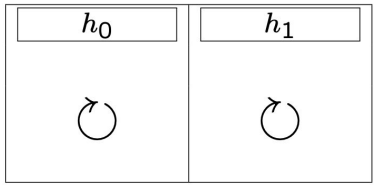
BIKE performance numbers from Drucker, Gueron, Kostic, "Additional implementation of BIKE (Bit Flipping Key Encapsulation)". https://github.com/awslabs/bike-kem.

BIKE Overview

- Niederreiter-based KEM instantiated with QC-MDPC codes.
- Leverage Fujisaki-Okamoto Transform [DGKP'21].
- [**New**] State-of-the-art QC-MDPC Decoding Failure Rate analysis.
- [**New**] BIKE-Flip Decoder with strengthened resilience to weak keys.

BIKE Overview

- Block circulant matrices are isomorphic to $\mathcal{R} = \mathbb{F}_2[x]/(x^r - 1)$.
- Index 2 Quasi-Cyclic codes admit a parity check matrix $(n = 2r)$.

$H =$ 

with $(h_0, h_1) \in \mathcal{R}^2$

- QC-MDPC if $w = |h_0| + |h_1| \approx \sqrt{n}$.
- Decoding: find $(e_0, e_1) \in \mathcal{R}^2$ s.t. $s = e_0 h_0 + e_1 h_1$ for small $t = |e_0| + |e_1|$.
- Bit-flipping decoders succeed (i.e. $(e_0, e_1) = decode(e_0 h_0 + e_1 h_1, h_0, h_1)$) with high probability if $t \approx w \approx \sqrt{n}$.

BIKE Specification

KeyGen : $() \mapsto (h_0, h_1, \sigma), h$

Output: $(h_0, h_1, \sigma) \in \mathcal{H}_w \times \mathcal{M}$, $h \in \mathcal{R}$

1: $(h_0, h_1) \xleftarrow{\mathcal{D}} \mathcal{H}_w$ ▷ (1)
2: $h \leftarrow h_1 h_0^{-1}$
3: $\sigma \xleftarrow{\$} \mathcal{M}$

Encaps : $h \mapsto K, c$

Input: $h \in \mathcal{R}$

Output: $K \in \mathcal{K}$, $c \in \mathcal{R} \times \mathcal{M}$

1: $m \xleftarrow{\$} \mathcal{M}$
2: $(e_0, e_1) \leftarrow \mathbf{H}(m)$
3: $c \leftarrow (e_0 + e_1 h, m \oplus \mathbf{L}(e_0, e_1))$
4: $K \leftarrow \mathbf{K}(m, c)$

Decaps : $(h_0, h_1, \sigma), c \mapsto K$

Input: $((h_0, h_1), \sigma) \in \mathcal{H}_w \times \mathcal{M}$, $c = (c_0, c_1) \in \mathcal{R} \times \mathcal{M}$

Output: $K \in \mathcal{K}$

1: $e' \leftarrow \texttt{decoder}(c_0 h_0, h_0, h_1)$ ▷ $e' \in \mathcal{R}^2 \cup \{\perp\}$
2: $m' \leftarrow c_1 \oplus \mathbf{L}(e')$ ▷ with the convention $\perp = (0, 0)$
3: **if** $e' = \mathbf{H}(m')$ **then** $K \leftarrow \mathbf{K}(m', c)$ **else** $K \leftarrow \mathbf{K}(\sigma, c)$

Parameters

r : block length
w : row weight
t : error weight
ℓ : shared secret size
$\mathcal{M}$: message space in $\{0,1\}^\ell$
$\mathcal{K}$: key space in $\{0,1\}^\ell$

NOTATION	
$\mathbb{F}_2$:	Binary finite field.
$\mathcal{R}$:	Cyclic polynomial ring $\mathbb{F}_2[X]/(X^r - 1)$.
$\mathcal{H}_w$:	Private key space $\{(h_0, h_1) \in \mathcal{R}^2 \mid \lvert h_0\rvert = \lvert h_1\rvert = w/2\}$
$\mathcal{E}_t$:	Error space $\{(e_0, e_1) \in \mathcal{R}^2 \mid \lvert e_0\rvert + \lvert e_1\rvert = t\}$
$\lvert g\rvert$:	Hamming weight of a binary polynomial $g \in \mathcal{R}$.
$u \xleftarrow{\$} U$:	Variable u is sampled uniformly at random from the set U.
$\oplus$:	exclusive or of two bits, componentwise with vectors

Functions

- $\mathbf{H} : \mathcal{M} \to \mathcal{E}_t$.
- $\mathbf{K} : \mathcal{M} \times \mathcal{R} \times \mathcal{M} \to \mathcal{K}$.
- $\mathbf{L} : \mathcal{R}^2 \to \mathcal{M}$

BIKE Security

- **Assumption 1**. Hardness of $QCSD_{r,t}$
- **Assumption 2**. Hardness of $QCCF_{r,w}$
- **Assumption 3**. Correctness of the decoder.

BIKE is proven IND-CPA secure under assumptions 1 and 2.

BIKE is proven IND-CCA secure under assumptions 1, 2, and 3.

BIKE Security

$QCSD_{(r,t)}$: Quasi-Cyclic Syndrome Decoding.
Instance: $(h, s) \in \mathcal{R}_{odd} \times \mathcal{R}_{p(t)}$ an integer $t > 0$.
Property: There exists $(e_0, e_1) \in \mathcal{E}_t$ such that $e_0 + e_1 h = s$.

$QCCF_{(r,w)}$: Quasi-Cyclic Codeword Finding
Instance: $h \in \mathcal{R}_{odd}$, an even integer $w > 0$, with $w/2$ odd.
Property: There exists $(h_0, h_1) \in \mathcal{H}_w$ such that $h_1 + h_0 h = 0$.

Decoding correctness: refers to [HHK'17] where a KEM is δ-correct if the decapsulation fails with probability at most δ on average over all keys and messages. Similarly, a decoder will be δ-correct if its failure rate is at most δ on average when the input is drawn uniformly.

BIKE Practical Security

- The best techniques to solve codeword finding and syndrome decoding:
  - Variants of Prange's Information Set Decoding (ISD) [Pran**62**]
- Work factor of any ISD variant A to decode t errors in a (n, k)-binary code:

$$WF_A(n, k, t) = 2^{ct(1+o(1))}$$

- The quasi-cyclic case
  - Codeword finding and decoding are a bit easier for quasi-cyclic codes.
  - Adversary gains a factor r for codeword finding and factor $\sqrt{r}$ for decoding
- The best quantum attack against BIKE
  - Grover's algorithm [Gro96] applied to ISD

BIKE - Modes of Usage

- **BIKE can use different decoders without affecting interoperability**.
  - The decoder must be implementable in constant-time to avoid side-channel attacks, and its DFR must be low enough to match the security requirement.

- **BIKE ephemeral keys usage.**
  - The party that initiates a session needs to: a) Generate a fresh private/public key pair for every session; b) Refuse to decapsulate more than one incoming ciphertext with that key.
  - The IND-CPA security property suffices for this type of usage.

- **BIKE long-term static keys usage**.
  - This usage requires IND-CCA security and therefore a low enough DFR for the specified decoder (see next discussion). However, this usage model implies the loss of forward secrecy.

Nicolas' Slides on Decoding and DFR recent results.

References

- [DGK]: Drucker, Gueron, Kostic, "Additional implementation of BIKE (Bit Flipping Key Encapsulation)". https://github.com/awslabs/bike-kem.
- [DGK'20]: Drucker, N., Gueron, S., Kostic, D.: Fast Polynomial Inversion for Post Quantum QC-MDPC Cryptography. Cyber Security Cryptography and Machine Learning. pp. 110–127. Springer International Publishing, Cham (2020). https://doi.org/10.1007/978-3-030-49785-9_8
- [DGK'23]: Drucker, N., Gueron, S., Kostic, D.: To Reject or Not Reject: That Is the Question. The Case of BIKE Post Quantum KEM. In: Latifi, S. (ed.) ITNG 2023 20th International Conference on Information Technology-New Generations. pp. 125–131. Springer International Publishing, Cham (2023).
- [DGKP'21] Drucker, N., Gueron, S., Kostic, D., Persichetti, E.: On the applicability of the Fujisaki-Okamoto transformation to the BIKE KEM. Int. J. Comput. Math. Comput. Syst. Theory 6(4), 364–374 (2021). https://doi.org/10.1080/23799927.2021.1930176
- [GHJ'22]: Qian Guo, Clemens Hlauschek, Thomas Johansson, Norman Lahr, Alexander Nilsson, and Robin Leander Schröder. Don't reject this: Key-recovery timing attacks due to rejection-sampling in HQC and BIKE. IACR Trans. Cryptogr. Hardw. Embed. Syst., 2022(3):223–263, 2022.
- [HHK'17]: Dennis Hofheinz, Kathrin Hövelmanns, and Eike Kiltz. A modular analysis of the Fujisaki-Okamoto transformation. In Theory of Cryptography Conference, pages 341–371. Springer, 2017.
- [MTSB'13]: Misoczki, Rafael, Jean-Pierre Tillich, Nicolas Sendrier, and Paulo SLM Barreto. "MDPC-McEliece: New McEliece variants from moderate density parity-check codes." In 2013 IEEE international symposium on information theory, pp. 2069-2073. IEEE, 2013.
- [Sen'23]: Nicolas Sendrier. Secure sampling of constant-weight words – Application to BIKE. Cryptology ePrint Archive, Report 2021/1631, August 2023.
- [WWW'23]: Wang, Tianrui, Anyu Wang, and Xiaoyun Wang. "Exploring Decryption Failures of BIKE: New Class of Weak Keys and Key Recovery Attacks." Annual International Cryptology Conference. Cham: Springer Nature Switzerland, 2023.

Questions?

https://bikesuite.org

BIKE Decoding Failure
Weak Keys & Error Floors

Nicolas Sendrier, Inria, France

NIST PQC Seminars, September 27, 2024

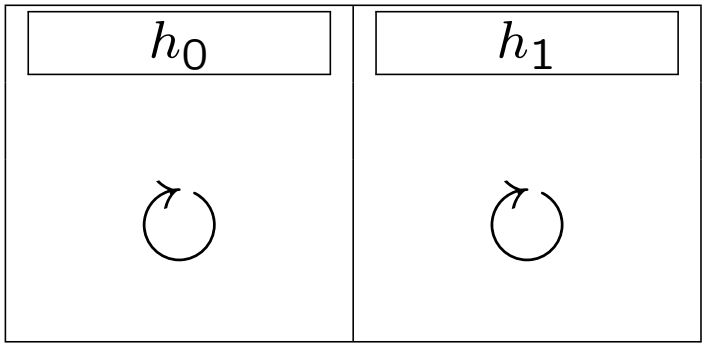
Introduction:
BIKE Security & DFR

QC-MDPC (Quasi-Cyclic Moderate Parity Check) Codes

Block circulant $r \times r$ binary matrices are isomorphic to $\mathcal{R} = \mathbf{F}_2[x]/(x^r - 1)$

—

Index 2 QC codes admit a parity check matrix ($n = 2r$)

$H =$ 

with $(h_0, h_1) \in \mathcal{R}^2$

QC-MDPC if $w = |h_0| + |h_1| \approx \sqrt{n}$

—

Decoding: find $(e_0, e_1) \in \mathcal{R}^2$ s.t. $s = e_0 h_0 + e_1 h_1$ for small $t = |e_0| + |e_1|$

Bit-flipping decoding succeeds* with high probability if $t \approx w \approx \sqrt{n}$

*successful decoding means $(e_0, e_1) = \mathsf{decode}(e_0 h_0 + e_1 h_1, h_0, h_1)$

BIKE

McEliece (Niederreiter) scheme with index 2 binary QC-MDPC codes

$\mathcal{R} = \mathbf{F}_2[x]/(x^r - 1)$

$\mathcal{H}_w = \{(h_0, h_1) \in \mathcal{R}^2 \mid |h_0| = |h_1| = d = w/2\}$

$\mathcal{E}_t = \{(e_0, e_1) \in \mathcal{R}^2 \mid |e_0| + |e_1| = t\}$

Secret: $(h_0, h_1) \in \mathcal{H}_w$

Public: $h = h_0^{-1} h_1 \in \mathcal{R}$

Encrypt: $(e_0, e_1) \in \mathcal{E}_t \mapsto c = e_0 + h e_1$

Decrypt: $c \mapsto (e_0, e_1) = \mathsf{decode}(h_0 c, h_0, h_1)$

Security	λ	r	d	t
Level 1	128	12 323	71	134
Level 3	192	24 659	103	199
Level 5	256	40 973	137	264

Decoding Failure Rate

$\mathcal{R} = \mathbf{F}_2[x]/(x^r - 1)$

$\mathcal{H}_w = \{(h_0, h_1) \in \mathcal{R}^2 \mid |h_0| = |h_1| = d = w/2\}$

$\mathcal{E}_t = \{(e_0, e_1) \in \mathcal{R}^2 \mid |e_0| + |e_1| = t\}$

Decoding Failure Rate:

$$\mathsf{DFR} = \Pr\left((e_0, e_1) \neq \mathsf{decode}(e_0 h_0 + e_1 h_1, h_0, h_1)\right)$$

with (e_0, e_1) and (h_0, h_1) uniformly distributed in $\mathcal{E}_t$ and $\mathcal{H}_w$

DFR & CCA Security

IND-CCA reduction [HHK17]: breaking BIKE requires a computational effort

$$\geq \min\left(\frac{1}{\mathrm{DFR}}, 2^{\lambda_{\mathrm{CPA}}}\right)$$

$2^{\lambda_{\mathrm{CPA}}}$: "hardness of decoding"

Proving $\mathrm{DFR} \leq 2^{-\lambda}$ (*e.g.* $\lambda = 128$) by simulation is out of reach

—

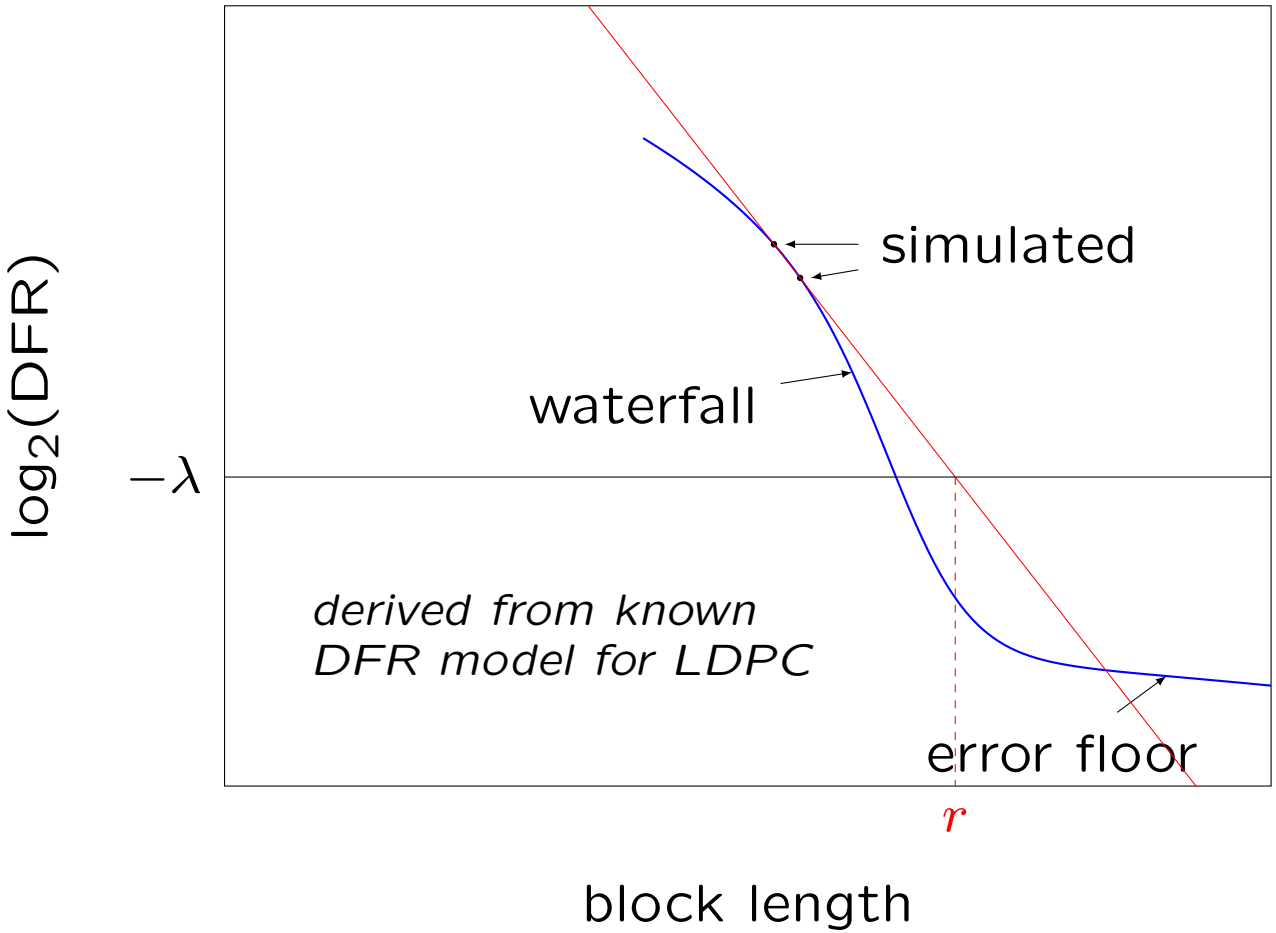
High DFR allows failure attacks:

[Guo, Johansson, Stankovski, Asiacrypt 2016]

[Wang, Wang, Wang, Crypto 2023]

[HHK17] Hofheinz, Hövelmanns, Kiltz, *A modular analysis of the Fujisaki-Okamoto transformation*, TCC 2017

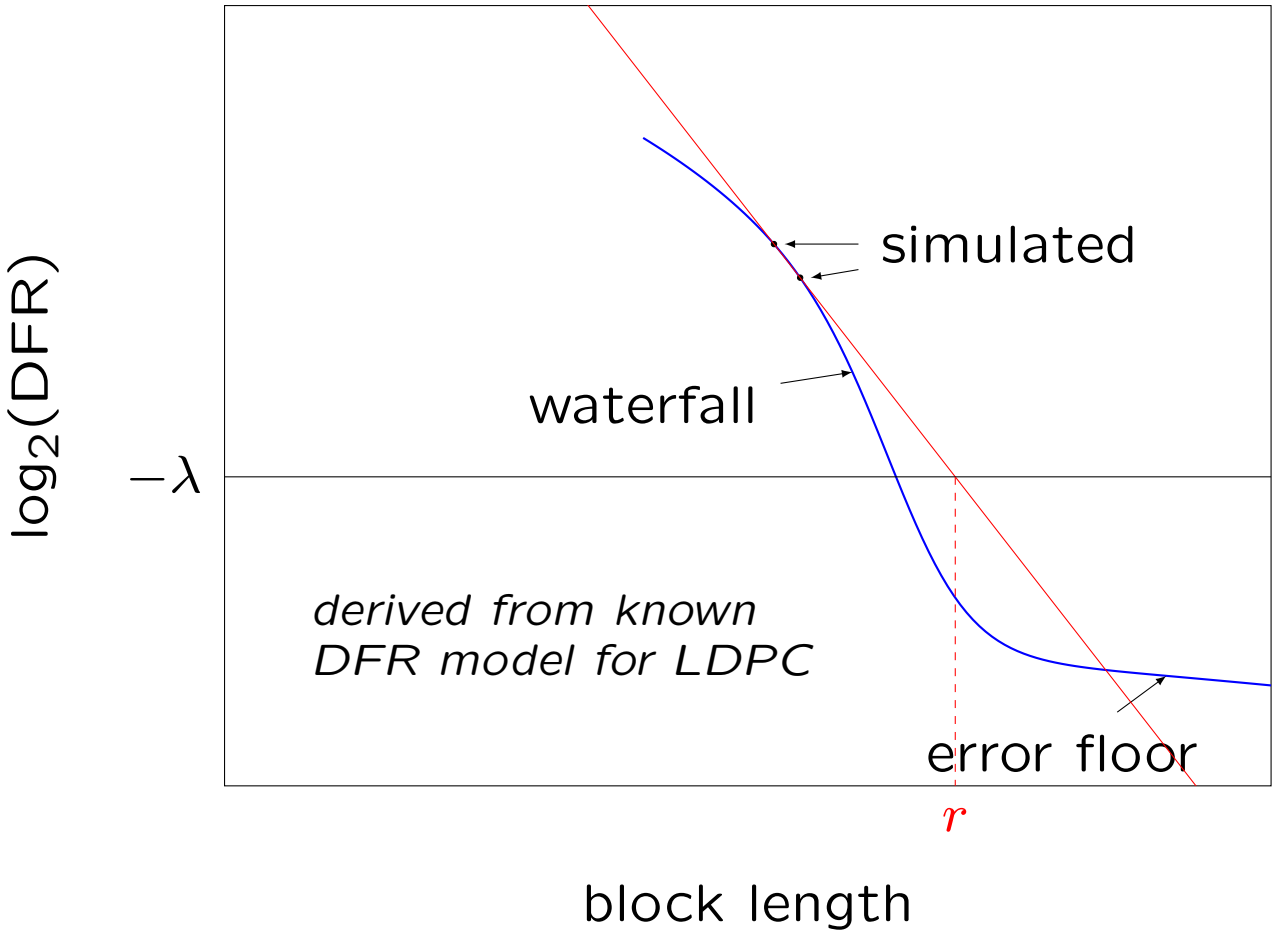
BIKE DFR Estimates (1)



Fix the row weight w and the error weight t

The failure rate can be measured for small block length r

BIKE DFR Estimates (2)



There are models for the waterfall region allowing extrapolation: asymptotic [Tillich, 2018], Markovian [Sendrier, Vasseur, 2019]

There are works in progress to model the error floor region

🚲

Roadmap

Introduction: BIKE Security & DFR

I. Coordinate Distance Spectrum

II. Weak Keys

III. m-gathering Weak Keys

IV.1 m-gathering Key Spectrum

IV.2 Bogus Positions

IV. New BIKE Decoder

V. Weak Keys and Error Floors: Perspectives and Open Questions

I. Coordinate Distance Spectrum

Coordinate Distance Spectrum and Multiplicity

[Guo, Johansson, Stankovski, Asiacrypt 2016]

Distance spectrum, $h \in \mathcal{R} = \mathbf{F}_2[x]/(x^r - 1)$

$$\mathsf{Sp}(h) = \{\mathsf{d}(i,j) \mid h_i = h_j = 1\}.$$

Coordinate distance

$$\mathsf{d}(i,j) = \min(i - j \bmod r, j - i \bmod r), \quad 0 \le \mathsf{d}(i,j) \le \lfloor r/2 \rfloor,$$

Multiplicity

$$\mu(\delta, h) = \left|\{(i,j) \mid h_i = h_j = 1, 0 \le i \le j < r, \mathsf{d}(i,j) = \delta\}\right|.$$

Spectrum Model

Model: draw $d(d-1)/2$ elements uniformly in $\{1,\ldots,(r-1)/2\}$ with repetition.

Ex: $(r,d) = (12323, 71)$ $h \in \mathcal{R}$ uniform of weight d

	Model	Simulation
$\lvert\mathrm{Sp}(h)\rvert$	2045.1	2054.4
$\mu = 1$	1660.4	1675.4
$\mu = 2$	334.8	332.1
$\mu = 3$	44.98	42.64
$\mu = 4$	4.53	3.99
$\mu = 5$	0.36	0.29
$\mu = 6$	0.024	0.017

$$\frac{r-1}{2} = 6161$$

$$\frac{d(d-1)}{2} = 2485$$

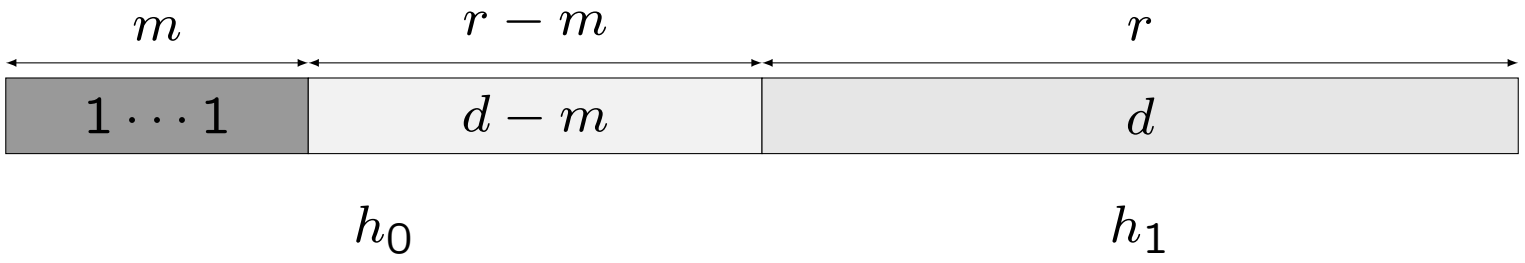
(each row gives the average number of distances with multiplicity μ)

II. Weak Keys

Known Weak Keys

Parameter $m > 0$, $m < d$, typical value 10 to 30

Type I: [Drucker, Gueron, Kostic, 2019]



$+$ blockwise rotation $+$ isomorphisms $x \mapsto x^i$, $0 \leq i < r$

$\rightarrow$ multiplicities $m - 1$, $m - 2$, ... appear once

Type II: [Vasseur, 2021] one distance has multiplicity $\geq m$

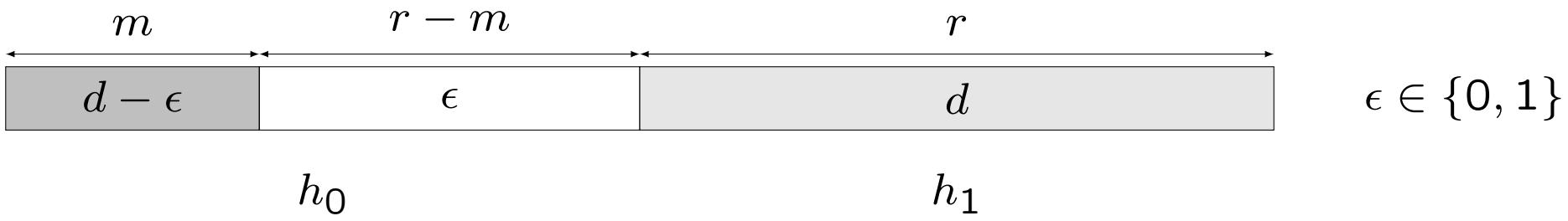
Type III: [Vasseur, 2021] one inter-block distance has multiplicity $\geq m$

No indication that $\mathrm{DFR}(\mathcal{W}) \cdot \mathrm{density}(\mathcal{W}) \geq 2^{-\lambda}$ if $\mathcal{W}$ is one of those sets

m-gathering Weak Keys

[Wang, Wang, Wang, Crypto 2023]

h_0 and h_1 of weight d in $\mathbf{F}_2[x]/(x^r - 1)$ a BIKE secret key



$\epsilon \in \{0, 1\}$

Generalize by applying rotations and isomorphisms $x \mapsto x^i$, $0 < i < r$

e.g. $(m, \epsilon) = (4000, 1)$, density is $2^{-87.28}$ and the DFR is $2^{-29.33}$

$\rightarrow$ contribution to the average DFR is $\geq 2^{-117}$

$\rightarrow$ successful attack with 2^{117} Decaps queries

The same paper shows a second attack in 2^{98} with message reuse
It is canceled by binding the key to the message as in Kyber

[Drucker, Gueron, Kostic, 2021], Binding BIKE Errors to a Key Pair, LNCS 12716

III. m-gathering Weak Keys

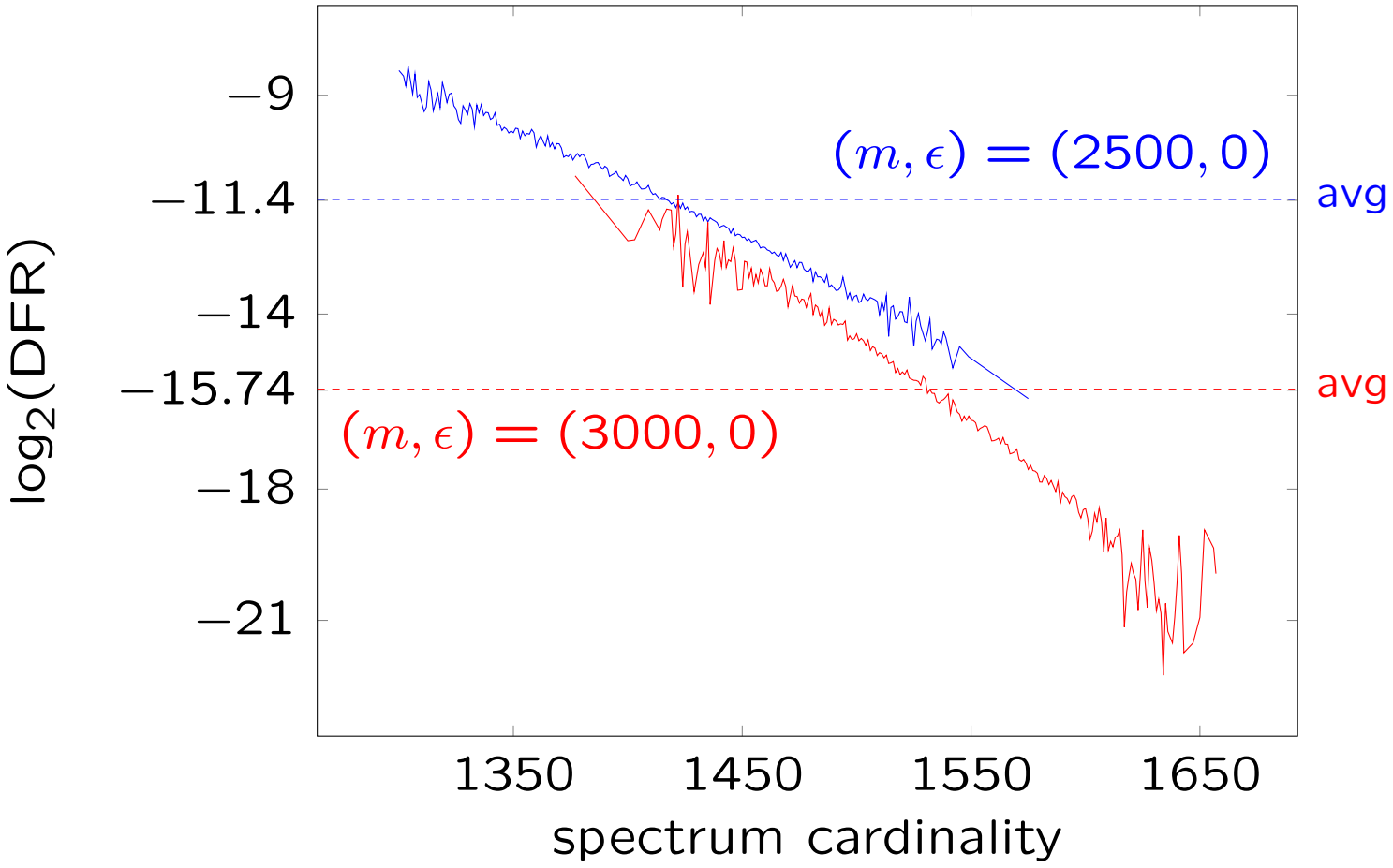
Observations on m-gathering Failures

1. Failure probability is strongly correlated with the spectrum cardinality.

2. The bit-flipping decoder features *bogus positions*: positions which are correct but look like errors.

IV.1 m-gathering Key Spectrum

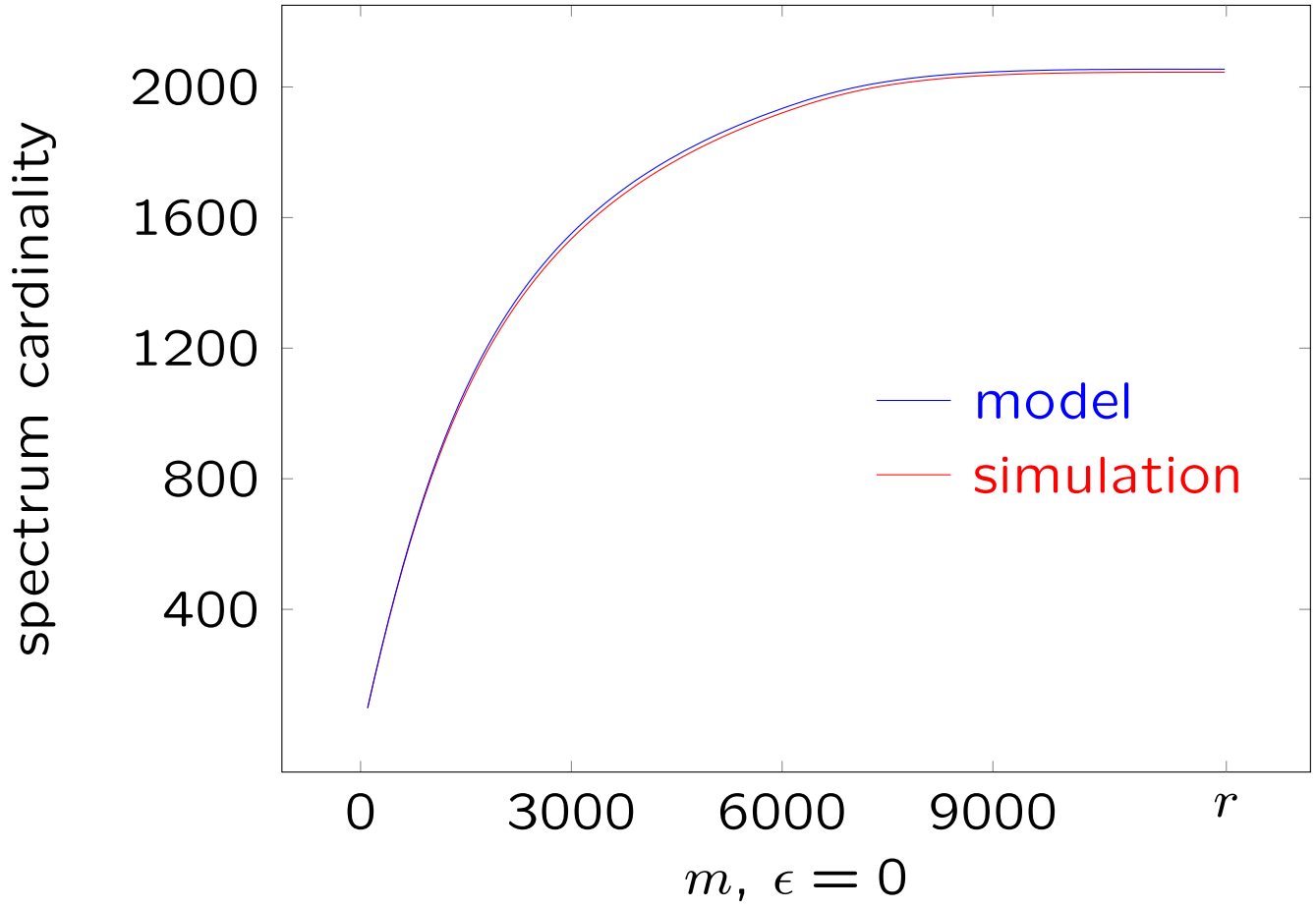
Spectrum Cardinality vs DFR for m-gathering Keys



(the spectrum cardinality of a typical element of $\mathcal{R}$ is 2050)

Spectrum Model for m-gathering Keys

Model: (for $\epsilon = 0$, $m < r/2$ before applying $x \mapsto x^i$)
draw $d(d-1)/2$ elements in $\{1, \ldots, m-1\}$ with repetition according to the distribution: $\Pr(\delta) = \dfrac{2(m-\delta)}{m(m-1)}, 1 \leq \delta < m$



Distance Mutiplicity for m-gathering Keys

	(m, ϵ)						
	(2000, 0)		(4000, 0)		(6000, 0)		uniform
	model	simul	model	simul	model	simul	model
$\mu = 1$	571.0	579.5	1 135.5	1 154.6	1 459.2	1 479.2	1 660.4
$\mu = 2$	364.1	374.2	418.1	421.7	373.6	372.5	334.8
$\mu = 3$	192.6	195.7	120.9	117.7	73.92	70.45	44.98
$\mu = 4$	85.99	84.37	28.65	26.20	11.87	10.54	4.53
$\mu = 5$	33.01	30.52	5.73	4.80	1.60	1.29	0.36
$\mu = 6$	11.07	9.39	0.99	0.74	0.19	0.13	0.024
$\mu = 7$	3.29	2.49	0.15	0.097	0.019	0.012	0.0014
$\mu = 8$	0.88	0.57	0.020	0.011	0.0017	0.0008	0.000071
$\mu > 0$	1 262.2	1 277.0	1 710.1	1 725.8	1 920.4	1 934.1	2 045.1

Expected number of distances of multiplicity μ for various gathering parameters

IV.2 Bogus Positions

Bit-Flipping Decoding

Input: $s \in \mathbf{F}_2^r$, $H \in \mathbf{F}_2^{r \times n}$

$\tilde{e} \leftarrow 0^n$; $\tilde{s} \leftarrow s$

while $\tilde{s} \neq 0$ **do**

$T \leftarrow \text{THRESHOLD}(context)$

for $j = 0, \ldots, n-1$ **do**

$\sigma_j \leftarrow \mathtt{ctr}(H, \tilde{s}, j)$

for $j = 0, \ldots, n-1$ **do**

if $\sigma_j \geq T$ **then**

$\tilde{e}_j \leftarrow \tilde{e}_j \oplus 1$

$\tilde{s} \leftarrow \tilde{s} - \mathsf{col}(H, j)$

return $\tilde{e}$

$\mathtt{ctr}(H, \tilde{s}, j)$ number of unsatisfied equations involving position j

a.k.a. counter of j

Binomial Model for Counters – Threshold for BIKE

Binomial model* for the counter value:

$$\begin{cases} \text{correct position:} & \sigma_j \sim \mathsf{Bin}(d, \pi_0) \\ \text{error position:} & \sigma_j \sim \mathsf{Bin}(d, \pi_1) \end{cases}$$

Typically $d = 71$, $\pi_0 \approx 0.4$ and $\pi_1 \approx 0.6$, with an exact value depending on the parameters and on the syndrome weight S.

For given parameters (r, d, t), the optimal threshold ([Chaulet 2017]) is well approximated by an affine function

e.g. $(r, d, t) = (12323, 71, 134) \to T = f_t(S) = 0.006258 \cdot S + 11.094$.

*$\sigma \sim \mathsf{Bin}(d, \pi) \Leftrightarrow \mathsf{Pr}(\sigma = i) = \binom{d}{i} \pi^i (1-\pi)^{d-i}$

BIKE Bit-Flipping Decoding

Input: $s \in \mathbf{F}_2^r$, $H \in \mathbf{F}_2^{r \times n}$

$\tilde{e} \leftarrow 0^n$; $\tilde{s} \leftarrow s$

for $i = 1, \ldots,$ NbIter **do** ▷ NbIter $= 7$ for security level 1

  $T \leftarrow f_t(|\tilde{s}|) + \delta$ ▷ $\delta = 3$, t is the initial error weight

  for $j = 0, \ldots, n-1$ **do**

    $\sigma_j \leftarrow \mathtt{ctr}(H, \tilde{s}, j)$

  for $j = 0, \ldots, n-1$ **do**

    if $\sigma_j \geq T$ **then**

      $\tilde{e}_j \leftarrow \tilde{e}_j \oplus 1$

      $\tilde{s} \leftarrow \tilde{s} - \mathrm{col}(H, j)$

return $\tilde{e}$

The real BIKE decoder is more complex and features so-called black and gray steps which do not essentially change what is coming next

m-gathering Bogus Positions

In-depth monitoring of m-gathering decoding failure instances reveals the following fact:

> *When the decoding fails, we observe at the second (sometimes the third) iteration that two things happen concomitantly*
>
> - *a few correct positions (a dozen or so) have an extremely high counter value*[*]*, the bogus positions;*
> - *the syndrome weight, and thus the threshold, is much lower than expected (given the remaining error weight at that point).*
>
> *The bogus positions are thus flipped, become errors and irreparably spoil the decoding process.*

Presumably, bogus positions are related to the existence of high multiplicity distances in the key spectrum

[*]near to impossible in the binomial model for counters

IV. New BIKE Decoder

New BIKE Decoder

For BIKE parameter (r, d, t)

Input: $s \in \mathbf{F}_2^r$, $H \in \mathbf{F}_2^{r \times n}$

$\tilde{e} \leftarrow 0^n$; $\tilde{s} \leftarrow s$

for $i = 1, \ldots,$ NbIter **do**

$T \leftarrow \max(T_i + \delta, f_t(|\tilde{s}|))$

for $j = 0, \ldots, n-1$ **do**

$\sigma_j \leftarrow \mathtt{ctr}(H, \tilde{s}, j)$

for $j = 0, \ldots, n-1$ **do**

if $\sigma_j \geq T$ **then**

$\tilde{e}_j \leftarrow \tilde{e}_j \oplus 1$

$\tilde{s} \leftarrow \tilde{s} - \mathsf{col}(H, j)$

return $\tilde{e}$

$$T_1 = f_t(|s|)$$
$$T_4 = (d+1)/2$$
$$T_2 = (2T_1 + T_4)/3$$
$$T_3 = (T_1 + 2T_4)/3$$
$$T_i = T_4,\ i > 4$$

$f_t(x) = 0.006258 \cdot x + 11.094, \delta = 3,$ NbIter $= 7$ (security level 1)

New Decoder DFR – Waterfall

r	10 620	10 650	10 680	10 700
#samples	$4.4 \cdot 10^9$	$13.3 \cdot 10^9$	$56.2 \cdot 10^9$	$38.2 \cdot 10^9$
#failures	16 222	7 756	3 183	870
$\log_2$(DFR)	-18.04	-20.71	-23.46	-25.39

"Waterfall DFR extrapolation":

- Level 1, $(r, d, t) = (12\,323, 71, 134)$, $\delta = 3$, NbIter $= 7 \rightarrow 2^{-181}$
- Level 3, $(r, d, t) = (24\,659, 103, 199)$, $\delta = 5$, NbIter $= 7 \rightarrow 2^{-266}$
- Level 5, $(r, d, t) = (40\,973, 137, 264)$, $\delta = 6$, NbIter $= 7 \rightarrow 2^{-343}$

New Decoder DFR – m-gathering Simulation (1/4)

(m, ϵ)	$(1600, 0)$	$(1700, 0)$	$(1800, 0)$	$(1900, 0)$	$(2000, 0)$	$(2100, 0)$
#samples	$9.5 \cdot 10^9$	$9.6 \cdot 10^9$	$9.6 \cdot 10^9$	$9.6 \cdot 10^9$	$15.7 \cdot 10^9$	$9.4 \cdot 10^9$
#failures	79913	32596	13153	5293	3383	763
$\log_2$(DFR)	-16.86	-18.16	-19.48	-20.79	-22.15	-23.56
$\log_2$(density)	-188.41	-182.15	-176.26	-170.69	-165.41	-160.40
sum	-205.27	-200.31	-195.74	-191.48	-187.56	-183.96

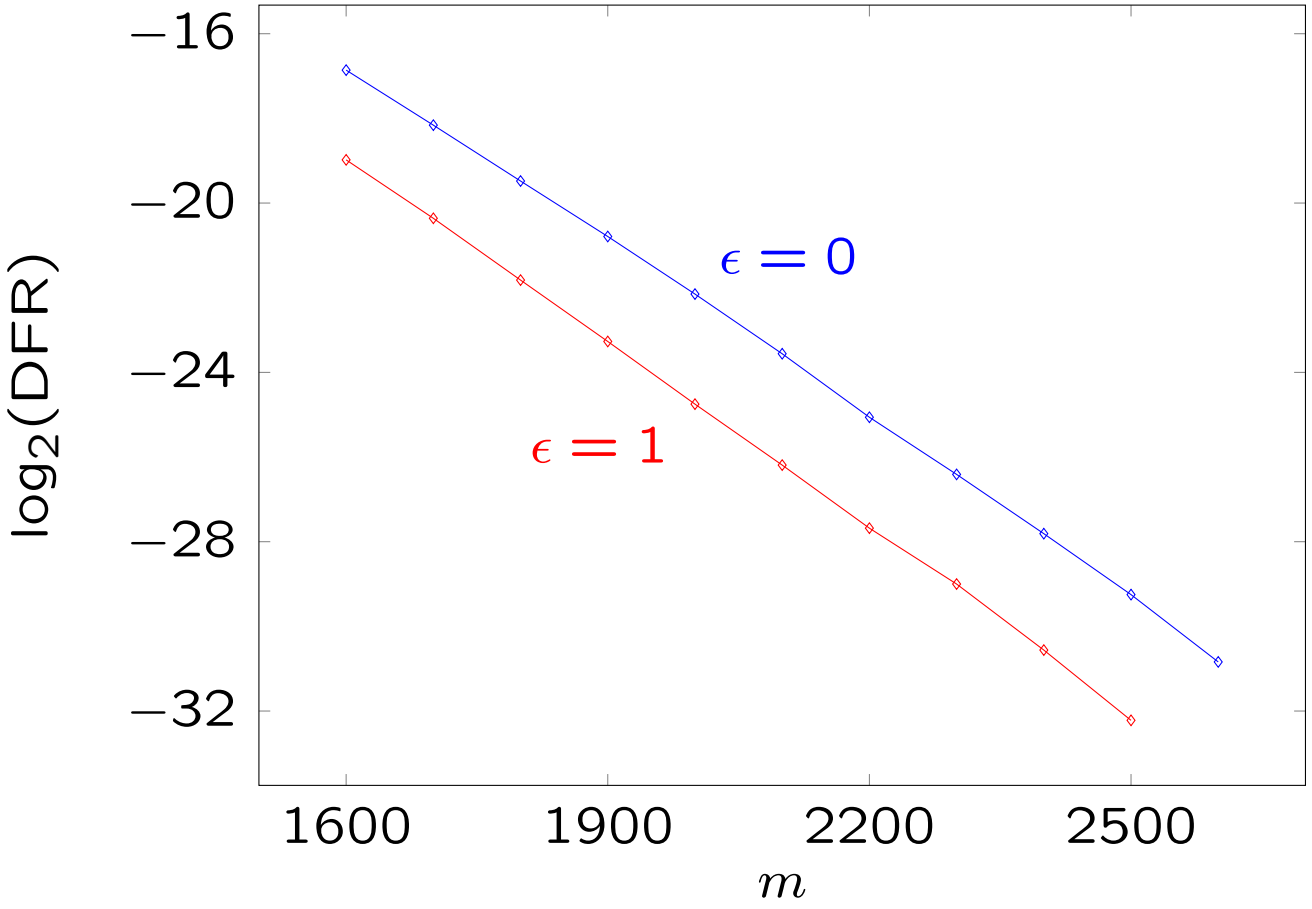
(m, ϵ)	$(2200, 0)$	$(2300, 0)$	$(2400, 0)$	$(2500, 0)$	$(2600, 0)$	$(2700, 0)$
#samples	$9.4 \cdot 10^9$	$15.8 \cdot 10^9$	$19.1 \cdot 10^9$	$25.5 \cdot 10^9$	$59.8 \cdot 10^9$	$7.9 \cdot 10^9$
#failures	270	177	81	40	31	0
$\log_2$(DFR)	-25.06	-26.41	-27.81	-29.25	-30.84	–
$\log_2$(density)	-155.62	-151.06	-146.70	-142.51	-138.50	-134.63
sum	-180.68	-177.47	-174.51	-171.76	-169.34	–

New Decoder DFR – m-gathering Simulation (2/4)

(m, ϵ)	$(1600, 1)$	$(1700, 1)$	$(1800, 1)$	$(1900, 1)$	$(2000, 1)$
#samples	$9.3 \cdot 10^9$	$9.3 \cdot 10^9$	$9.3 \cdot 10^9$	$9.3 \cdot 10^9$	$13.6 \cdot 10^9$
#failures	17916	6904	2524	921	482
$\log_2$(DFR)	-18.98	-20.36	-21.82	-23.27	-24.75
$\log_2$(density)	-179.47	-173.31	-167.52	-162.05	-156.86
sum	-198.45	-193.67	-189.34	-185.31	-181.61

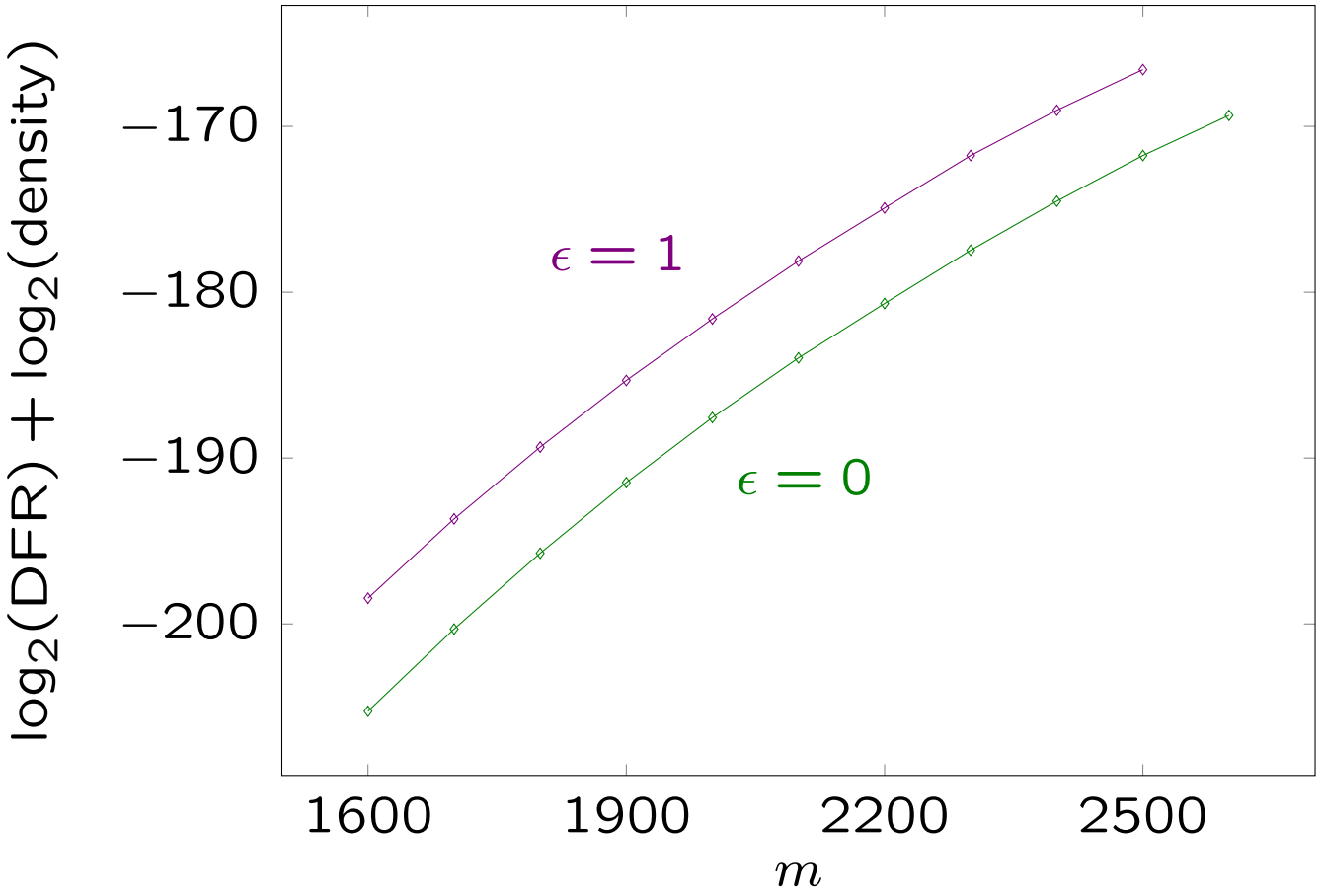
(m, ϵ)	$(2100, 1)$	$(2200, 1)$	$(2300, 1)$	$(2400, 1)$	$(2500, 1)$
#samples	$10.1 \cdot 10^9$	$20.1 \cdot 10^9$	$25.2 \cdot 10^9$	$25.3 \cdot 10^9$	$59.9 \cdot 10^9$
#failures	132	94	47	16	12
$\log_2$(DFR)	-26.19	-27.68	-29.00	-30.56	-32.22
$\log_2$(density)	-151.93	-147.24	-142.76	-138.47	-134.36
sum	-178.12	-174.91	-171.76	-169.03	-166.58

New Decoder DFR – m-gathering Simulation (3/4)



Old decoder: DFR $= 2^{-29.33}$ for $(m, \epsilon) = (4000, 1)$, density $2^{-87.28}$

New Decoder DFR – m-gathering Simulation (4/4)



Old decoder: for $(m, \epsilon) = (2500, 1)$, $\mathrm{DFR} \cdot \mathrm{density} = 2^{-145.8}$

New Decoder – Features

- Larger thresholds
- Similar algorithmic cost
- Simpler logic, and hopefully simpler analysis
- Decoding failure rate:
  - Same behavior in the waterfall region: fast exponential decrease
  - Better behavior against weak keys

V. Weak Keys & Error Floors

Perspectives and Open Questions

Weak Keys and the New Decoder

- weak keys have abnormally high multiplicities*
- higher multiplicities trigger more decoding failures
- bogus positions occur with weak keys and trigger decoding failures
- bogus positions are handled by increasing the thresholds
- higher thresholds reduce the weak key DFR by orders of magnitude

*multiplicities of values in the coordinate distance spectrum of the secret key

Error Floor Estimates

- distance to the closest near-codeword* is the dominant factor for the error floor estimation
- multiplicities are meaningful to estimate error floors

  relates to Tanner graphs and near-codewords (a.k.a. trapping sets)
- new models for counters†, straying from the typical case as the error pattern gets closer to a near-codeword

Work in progress, Markovian model taking into account near-codewords:

[Arpin, Billingsley, Lau, Perlner, Robinson, Tillich, Vasseur]

*a small set (of linear cardinality) of low weight words with low weight syndromes

†number of unsatisfied parity equations involving a position

New Model for Counters

$\mathbf{e} \in \mathbf{F}_2^n$ the error pattern (Hamming weight t at start)

$\mathbf{x} \in \mathbf{F}_2^n$ the closest near-codeword (Hamming weight $d = w/2$)

u the number of error positions in $\mathbf{x}$ ($u = |\mathbf{e} \star \mathbf{x}| \geq 1$)

For a position in the support of $\mathbf{x}$: *

$$\begin{cases} \text{correct position:} & \sigma_j \sim \mathsf{Bin}(d-u, \pi_0) + \mathsf{Bin}(u, \pi_1) \\ \text{error position:} & \sigma_j \sim \mathsf{Bin}(d-u+1, \pi_1) + \mathsf{Bin}(u-1, \pi_0) \end{cases}$$

For other positions:

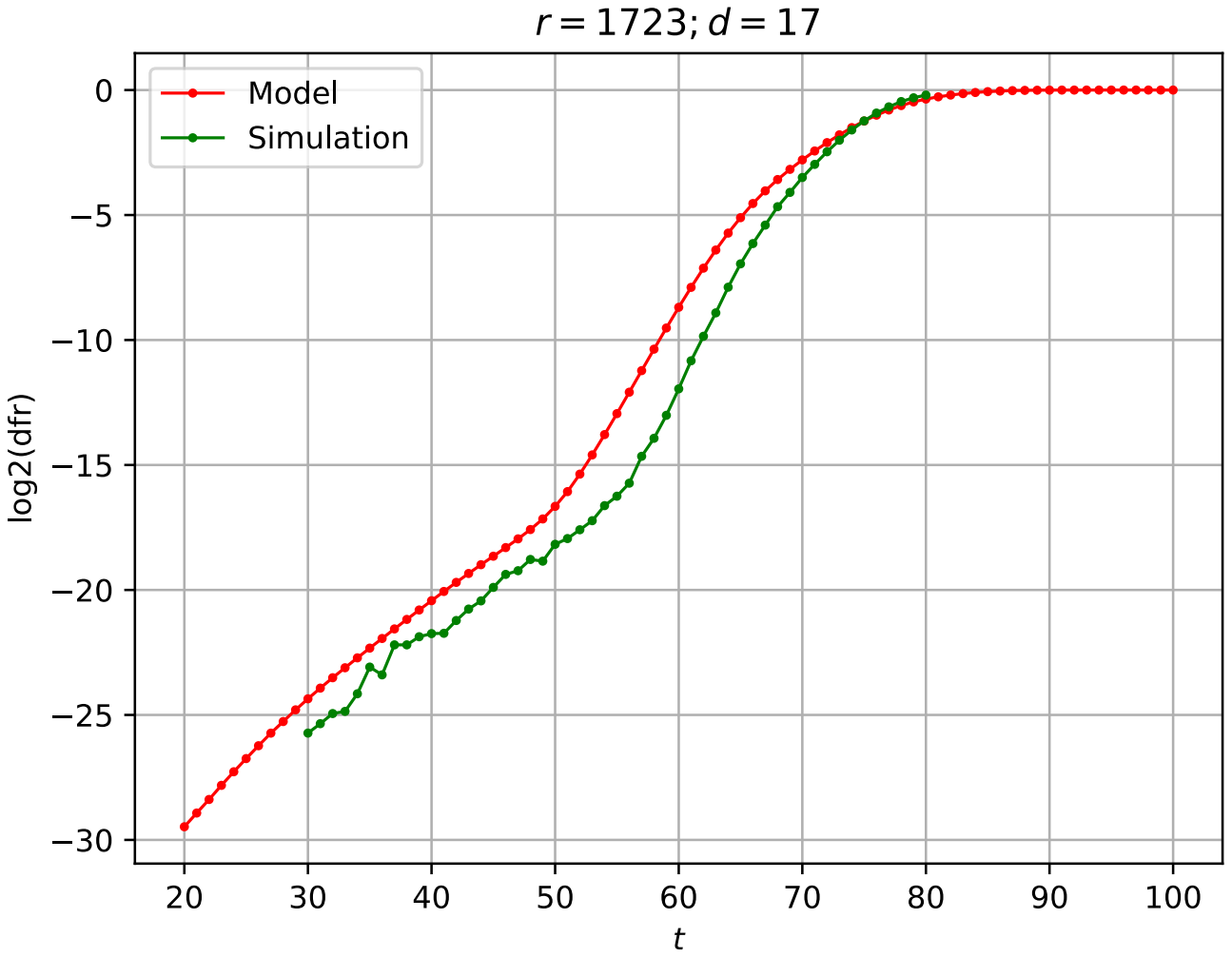
$$\begin{cases} \text{correct position:} & \sigma_j \sim \mathsf{Bin}(d, \pi_0) \\ \text{error position:} & \sigma_j \sim \mathsf{Bin}(d, \pi_1) \end{cases}$$

Recall that typically $d = 71$, $\pi_0 \approx 0.4$, $\pi_1 \approx 0.6$

counter = number of unsatisfied parity equations

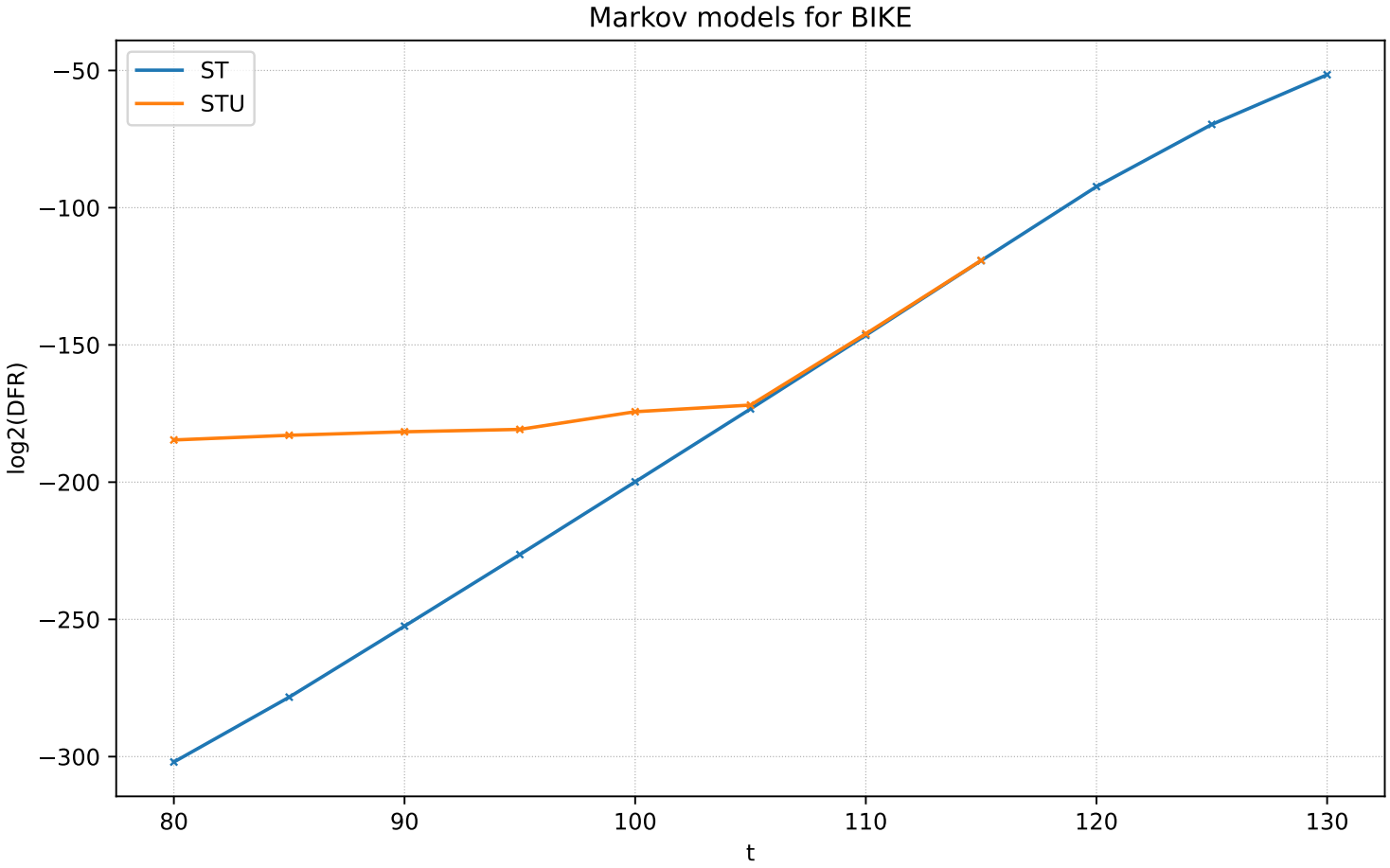
*[Arpin, Billingsley, Lau, Perlner, Robinson]

Markovian Model Matching the Error Floor (Perfect Keys)



[Tillich, Vasseur, 2024] model matches simulation for a toy example

Markovian Models (for Sequential Bit-Flipping)



[Sendrier, Vasseur, 2019] state : $(|\mathbf{e}|, |\mathbf{s}|)$

[*, ongoing] state : $(|\mathbf{e}|, |\mathbf{s}|, u)$

* [Arpin, Billingsley, Lau, Perlner, Robinson, Tillich, Vasseur]

Weak Keys vs. Error Floors – Connections & Perspectives

- Some correct positions may have higher counter values (and some erroneous positions may have lower counter values)

  consistent with bogus positions

- Near-codewords exist for all keys, their impact on error floors (and thus on decoding failures) depends on multiplicities

  consistent with a higher DFR for weak keys

Open Questions:

- Do all weak keys have untypical multiplicities? Converse?
- Are weak keys dominant for the average DFR/Error Floor?

Conclusion

Conclusions – BIKE DFR

Summing up (λ the security parameter, *e.g.* $\lambda = 128$):

- Waterfall

  DFR decreases exponentially in $O(\lambda^2)$ (block size)

- Error Floor (preliminary, work in progress)

  The new Markovian model predicts the error floor

  DFR seems to decrease exponentially in $O(\lambda)$ (error weight)

Perspectives to reduce the DFR

- improve the threshold schedule (with guaranties from the model)

- filter out keys with untypical spectrum (weak keys)

- increase the parameters