Decoding QC-MPDC codes

Nicolas Sendrier   Valentin Vasseur

QC-MDPC [MTSB13]

- McEliece-like public-key encryption scheme with a quasi-cyclic structure
  - Reasonable key sizes
  - Reduction to generic hard problems over quasi-cyclic codes
- Promising code-based key exchange mechanism proposed to the NIST call for standardization of quantum safe cryptography
  - “BIKE”
  - “QC-MDPC KEM”

$\mathbf{H} = (\mathbf{H}_0|\mathbf{H}_1) \leftarrow \mathbb{F}_2^{r\times n}$ quasi-cyclic
Weight of the lines of $\mathbf{H}$: w
$\mathbf{G} = (\mathbf{S}\mathbf{H}_1^T|\mathbf{S}\mathbf{H}_0^T) \in \mathbb{F}_2^{r\times n}$ generator matrix of $\mathbf{H}$ with $\mathbf{S}$ a dense circulant block

$\xrightarrow{\quad\mathbf{G}\quad}$

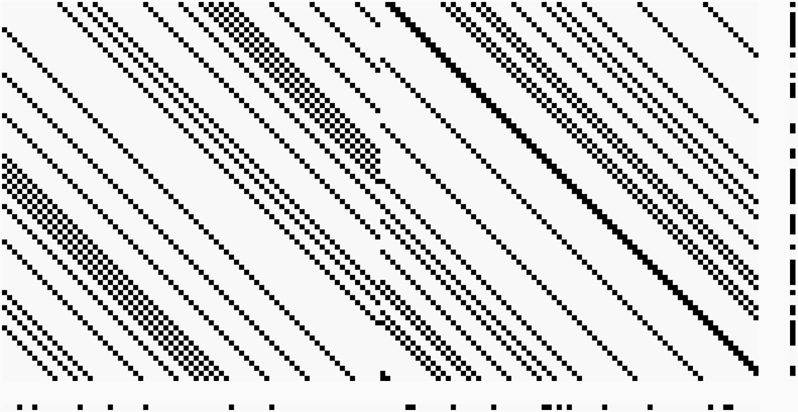
$\mathrm{m} \leftarrow \{0,1\}^r$
$\mathrm{c} = \mathrm{m}\mathbf{G}$
$\mathrm{e} \in \{0,1\}^n$
$|\mathrm{e}| = t$

$\xleftarrow{\quad \mathrm{y} = \mathrm{c} + \mathrm{e} \quad}$

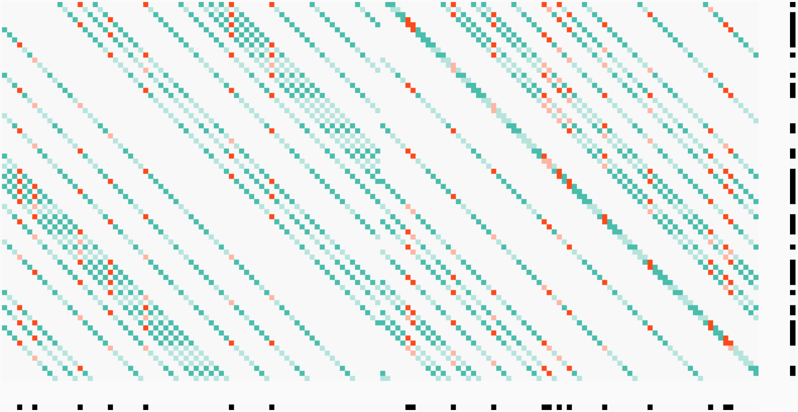
$\mathrm{m} = \mathrm{Decode}(\mathrm{y}, \mathbf{H})$

Parameters (BIKE): $r, d, t \in \mathbb{N}$, $n = 2r$, $w = 2d \sim t \sim \sqrt{n}$

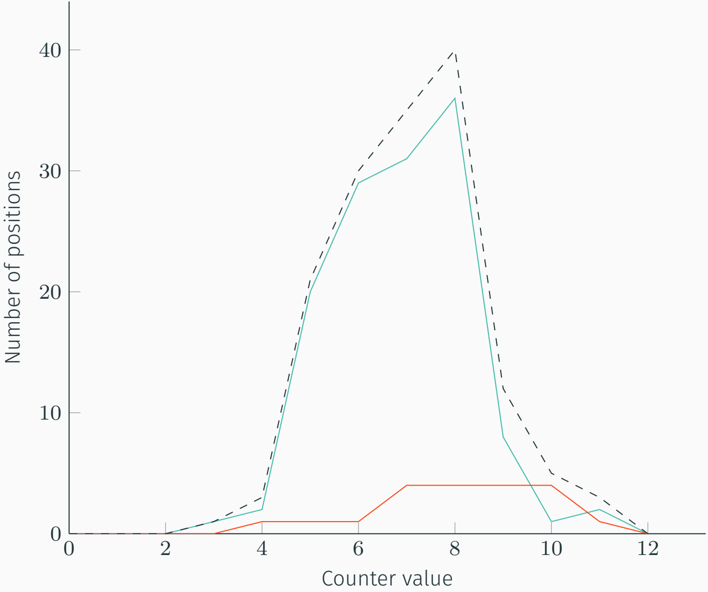
n	r	w	t	security
20 326	10 163	142	134	128
39 706	19 853	206	199	192
65 498	32 749	274	264	256



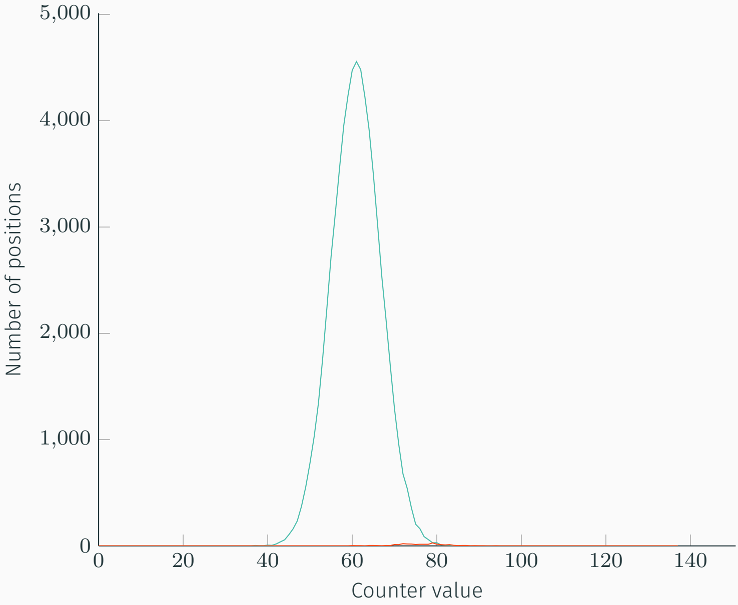
Classic decoding algorithm (*bit-flipping*)



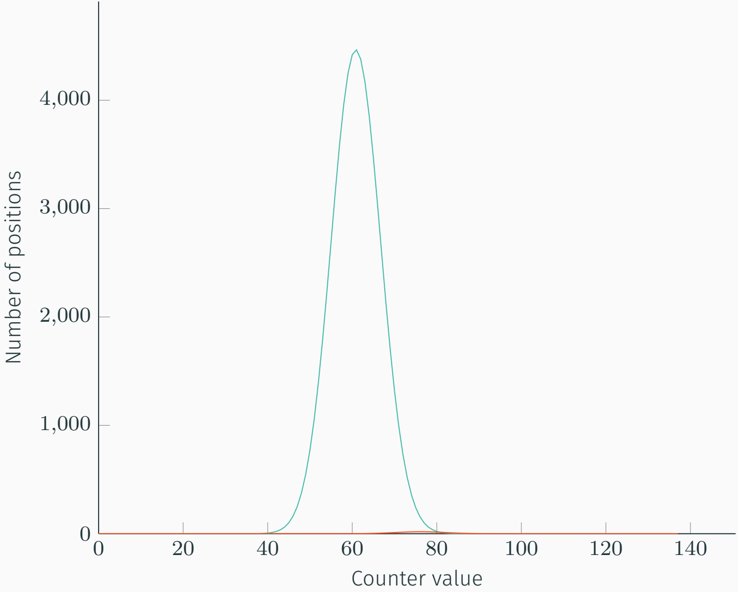
Counter distribution (our example)



Counters distributions (an example of BIKE $n = 65\,498$, $w = 274$, $t = 264$)



Counters distributions (average for BIKE $n = 65\,498$, $w = 274$, $t = 264$)



Counters distributions (formulas, see [Cha17])

- Write: $E_\ell = \left|\{i \in \{0, \ldots, r-1\} \,\middle|\, |\mathrm{eq}_i \cap \mathrm{e}| = \ell\}\right|$ and $X = \sum_{\ell \text{ odd}} (\ell - 1) E_\ell$,
- $\Pr\left(\left|\mathrm{s} \cap \mathrm{h}_j\right| = \sigma | j \notin \mathrm{e}\right) = f_{d,\pi_0}(\sigma)$ with $\pi_0 = \frac{\bar{\sigma}_{\mathrm{corr}}}{d} = \frac{(w-1)|\mathrm{s}|-X}{d(n-|\mathrm{e}|)}$,
- $\Pr\left(\left|\mathrm{s} \cap \mathrm{h}_j\right| = \sigma | j \in \mathrm{e}\right) = f_{d,\pi_1}(\sigma)$ with $\pi_1 = \frac{\bar{\sigma}_{\mathrm{err}}}{d} = \frac{|\mathrm{s}|+X}{d|\mathrm{e}|}$,

$$f_{n,p}(i) = \binom{n}{i} p^i (1-p)^{n-i} .$$

Classic decoding algorithm (*bit-flipping*)

while the syndrome is non-zero **do**
  choose a threshold
  for each position **do**
    if its counter is over the threshold **then**
      flip it

Motivation for decoding improvements and analysis

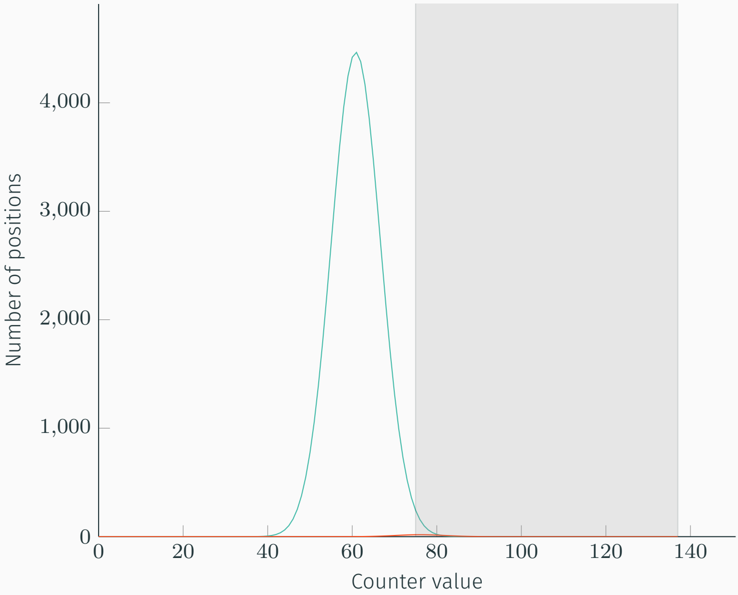
- [GJS16]: attack correlation between faulty error patterns and the secret key
- The best bit flipping improvement so far still have a decoding failure rate (DFR) around 2^{-30} ([Cho16; Cha17])
- A typical target DFR would be smaller than 2^{-64}
- ([ELPS18] extends [GJS16] to a timing attack requiring about 2^{28} samples)

Improving the decoder with a two-stage decoder

First stage: Grey decoding
choose a set *G* of positions with a high density of errors
while the syndrome weight is not below a certain target weight **do**
    choose a threshold
    for each position in *G* **do**
        if its counter is over the threshold **then**
            flip it

Second stage: Step-by-step decoding
while the syndrome is non-zero **do**
    choose *i* such that the *i*-th equation is unverified
    choose a position *j* in that equation
    if its counter is over half the weight of the column **then**
        flip it

First stage: Grey zone



Improving the decoder with a two-stage decoder

First stage: Grey decoding
choose a set G of positions with a high density of errors
while the syndrome weight is not below a certain target weight **do**
    choose a threshold
    for each position in G **do**
        if its counter is over the threshold **then**
            flip it

Second stage: Step-by-step decoding
while the syndrome is non-zero **do**
    choose i such that the i-th equation is unverified
    choose a position j in that equation
    if its counter is over half the weight of the column **then**
        flip it

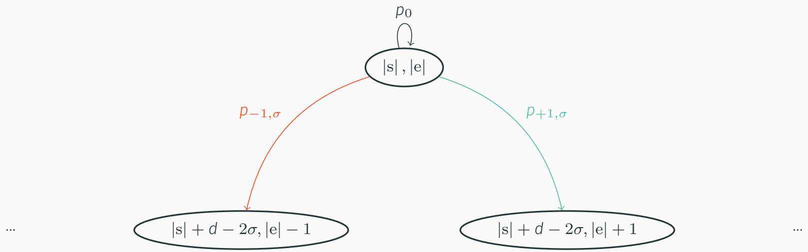
Analysis of the step-by-step decoder (model)

- Markovian process with state $(|s|, |e|)$
- Infinite number of iterations
- Counters distributions:
  - Write: $E_\ell = \left|\{i \in \{0, \ldots, r-1\} \mid |\mathrm{eq}_i \cap \mathrm{e}| = \ell\}\right|$ and $X = \sum_{\ell \text{ odd}} (\ell - 1) E_\ell$,
  - $\Pr\left(\left|\mathrm{s} \cap \mathrm{h}_j\right| = \sigma | j \notin \mathrm{e}\right) = f_{d,\pi_0}(\sigma)$ with $\pi_0 = \frac{\bar{\sigma}_{\mathrm{corr}}}{d} = \frac{(w-1)|\mathrm{s}| - X}{d(n - |\mathrm{e}|)}$,
  - $\Pr\left(\left|\mathrm{s} \cap \mathrm{h}_j\right| = \sigma | j \in \mathrm{e}\right) = f_{d,\pi_1}(\sigma)$ with $\pi_1 = \frac{\bar{\sigma}_{\mathrm{err}}}{d} = \frac{|\mathrm{s}| + X}{d|\mathrm{e}|}$,

$$f_{n,p}(i) = \binom{n}{i} p^i (1-p)^{n-i} .$$

- Different choices for X
  - $X = 0$ (the most pessimistic)
  - Expected value for X when e is uniform of weight $|\mathrm{e}|$ knowing $|\mathrm{s}|$ (close to our observations)
  - Expected value for X when e is uniform of weight $|\mathrm{e}|$ (the most optimistic)

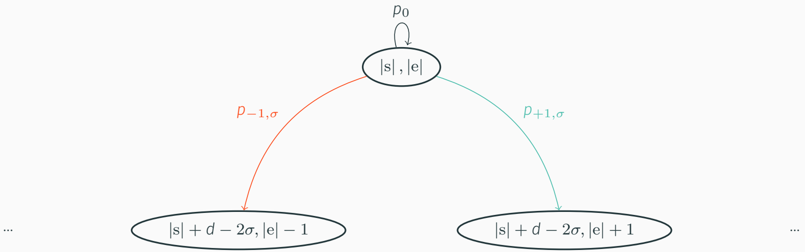
Markovian model



When we flip the column h_j, $s \leftarrow s + h_j$

$$\left|s + h_j\right| = |s| + \left|h_j\right| - 2\left|h_j \cap s\right| = |s| + d - 2\sigma$$

Markovian model: transition probabilities



$$p_0 = \Pr\left(\left|s \cap h_j\right| < 0.5d\right) = \sum_{\sigma < 0.5d} (1-\alpha) f_{d,\pi_0}(\sigma) + \alpha f_{d,\pi_1}(\sigma)$$

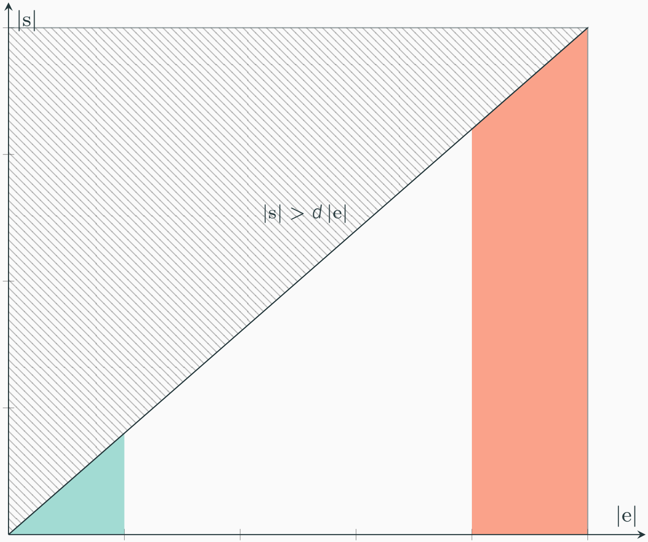
and if $\sigma \geq 0.5d$:

$$p_{+1,\sigma} = \Pr\left(\left|s \cap h_j\right| = \sigma, j \notin e\right) = (1-\alpha) f_{d,\pi_0}(\sigma)$$
$$p_{-1,\sigma} = \Pr\left(\left|s \cap h_j\right| = \sigma, j \in e\right) = \alpha f_{d,\pi_1}(\sigma)$$

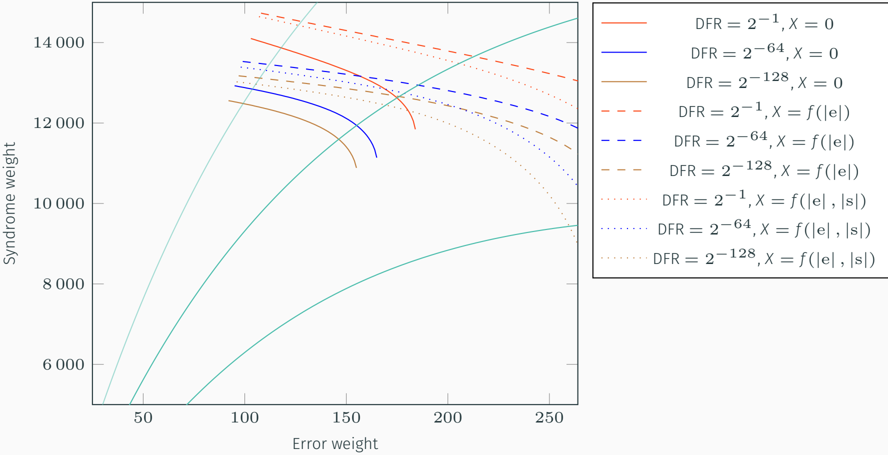
where

$$\alpha = \Pr(j \in e) = \frac{\sum_{j \in e} \left|s \cap h_j\right|}{|s|\, w} = \frac{|s| + X}{|s|\, w}.$$

Conditions

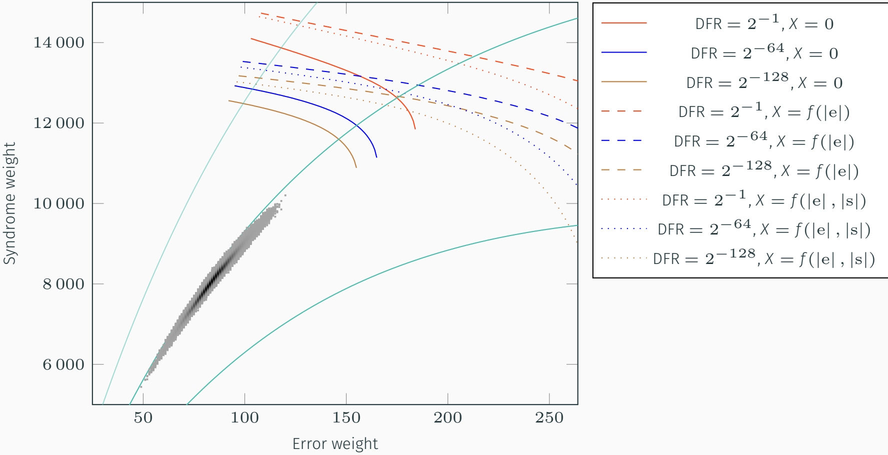


Results



14 000
12 000
10 000
8 000
6 000
50
100
150
200
250
Syndrome weight
Error weight
DFR = 2^{-1}, $X = 0$
DFR = 2^{-64}, $X = 0$
DFR = 2^{-128}, $X = 0$
DFR = 2^{-1}, $X = f(|e|)$
DFR = 2^{-64}, $X = f(|e|)$
DFR = 2^{-128}, $X = f(|e|)$
DFR = 2^{-1}, $X = f(|e|, |s|)$
DFR = 2^{-64}, $X = f(|e|, |s|)$
DFR = 2^{-128}, $X = f(|e|, |s|)$

Results



Syndrome weight
Error weight
14 000
12 000
10 000
8 000
6 000
50
100
150
200
250
DFR = 2^{-1}, $X = 0$
DFR = 2^{-64}, $X = 0$
DFR = 2^{-128}, $X = 0$
DFR = 2^{-1}, $X = f(|e|)$
DFR = 2^{-64}, $X = f(|e|)$
DFR = 2^{-128}, $X = f(|e|)$
DFR = 2^{-1}, $X = f(|e|, |s|)$
DFR = 2^{-64}, $X = f(|e|, |s|)$
DFR = 2^{-128}, $X = f(|e|, |s|)$

Conclusion

- As it is now, the two-stage decoding process greatly decreases the DFR
- We provide an analysis of the second stage
- We expect the whole process to have a DFR below 2^{-64}