Error floor prediction with Markov models for QC-MDPC codes

Sarah Arpin, Jun Bo Lau, Antoine Mesnard, **Ray Perlner**, Angela Robinson, Jean-Pierre Tillich, and Valentin Vasseur.

Crypto 2025
Santa Barbara • August 18, 2025

BIKE[2]

BIKE (**Bi**t-Flipping **K**ey **E**ncapsulation) is a code-based KEM (key encapsulation mechanism) based on QC-MDPC (Quasi-Cyclic Moderate-Density Parity-Check) codes. BIKE uses an *iterative decoder*, with a nonzero DFR (Decoding Failure Rate).

- BIKE in the NIST PQC Competition
  - Narrowly lost out to HQC in the 4^{th} round.
  - BIKE has smaller keys and ciphertexts, but BIKE's DFR has long been uncertain.
- IND-CCA security
  - BIKE's security proof for IND-CCA2 requires a DFR below $2^{-\lambda}$ for λ bits of security.
  - $2^{-\lambda}$ DFR is too low to measure – need to model for cryptographic parameters.
  - The GJS[1] key-recovery attack shows security loss is real if DFR is too high.

We model the DFR of QC-MDPC codes with dramatically improved accuracy.

[1] *A Key Recovery Attack on MDPC with CCA Security Using Decoding Errors*, Qian Guo, Thomas Johansson, and Paul Stankovski (2016).

[2] BIKE: Bit flipping key encapsulation - https://bikesuite.org

BIKE at a high level

- Parity check matrix $\mathbf{H} = [\mathbf{H}_0|\mathbf{H}_1]$ is composed of two sparse circulant blocks.
  - each column $\mathbf{h}_j$ of $\mathbf{H}$ has Hamming weight $|\mathbf{h}_j| = d$
- Public key $\mathbf{H}_0^{-1}\mathbf{H}$
- Message encoded as error vector $\mathbf{e} \in \mathbb{F}_2^{2r}$ of weight t.
- Ciphertext is $\mathbf{c} = \mathbf{H}_0^{-1}\mathbf{H}\mathbf{e}^T \in \mathbb{F}_2^r$.
- To decrypt, compute *syndrome* $\mathbf{s} = \mathbf{H}\mathbf{e}^T$ as $\mathbf{s} = \mathbf{H}_0\mathbf{c}$
- Then decode using Black-Grey-Flip (**BGF**) syndrome decoder.[3]
  - This is where decoding failures can happen.

[3]The BGF decoder: *QC-MDPC decoders with several shades of gray*, Drucker–Gueron–Kostic

Syndrome Decoding: Step-by-step

The BGF decoder used by BIKE is complicated enough to make explicit analysis challenging. Step-by-step is a simpler variant for analysis.

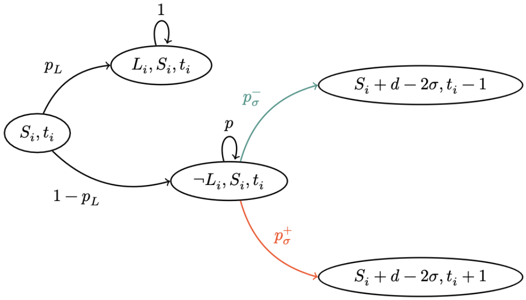
Input: A parity check matrix $\mathbf{H}$ and a syndrome vector $\mathbf{s}$.
Output: An error pattern $\mathbf{e}'$ satisfying $\mathbf{H}\mathbf{e}'^T = \mathbf{s}$.
Initialize: $\mathbf{e}' = 0$, $\Delta\mathbf{s} = \mathbf{s}$.
While $\Delta\mathbf{s} \neq 0$:

- Assign threshold $T := T(\Delta\mathbf{s})$.
- Sample a random column $\mathbf{h}_j$ of $\mathbf{H}$, with $j \in \{0, 1, ..., n-1\}$.
- Compute *counter* $\sigma = |\mathbf{h}_j \star s'|$
- If $\sigma \geqslant T$, then: Flip bit j of $\mathbf{e}'$ and set $\Delta\mathbf{s} = \Delta\mathbf{s} + \mathbf{h}_j$.
  - (A flip reduces $|\Delta\mathbf{s}|$ by $2\sigma - d$)

Once $\Delta\mathbf{s} = 0$, **return** $\mathbf{e}'$.

Markov Approach: Previous work [SV18][4]

State space: (S, t) where $S = |\Delta \mathbf{s}|$ and $t = |\Delta \mathbf{e}| = |\mathbf{e}' - \mathbf{e}|$.
L: blocked state.



- Problem: does not accurately model error floor.

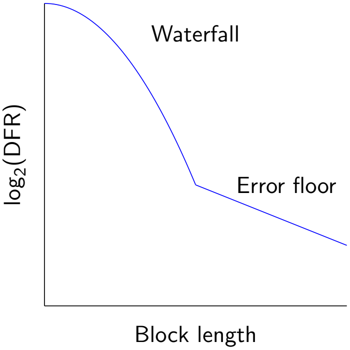
[4] *On the Decoding Failure Rate of QC-MDPC Bit-Flipping Decoders*, Nicolas Sendrier and Valentin Vasseur (2018). Figure: *Post-quantum cryptography: a study of the decoding of QC-MDPC codes*, Valentin Vasseur PhD thesis (2021).

What is an error floor?

Graphs of DFRs on a log scale for low- to moderate-density parity check codes with iterative decoders display a phenomenon:

- Initial, rapid decrease of decoding failures (**waterfall region**)
- Eventual plateau, more linear decrease (**error floor region**)

To accurately predict the DFR for higher code length (signal-to-noise ratio), one must account for the error floor region.



BIKE at Small Parameters: From [ABHLPR22][5]

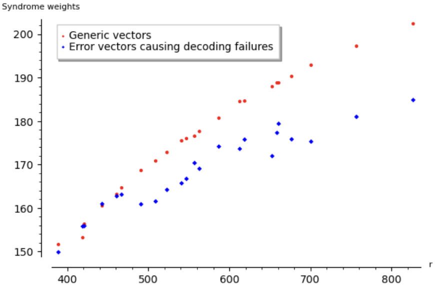


Fig. 7: Syndrome weights of random vectors with $t = 18$ (red circles) and vectors causing decoding failures (blue diamonds).

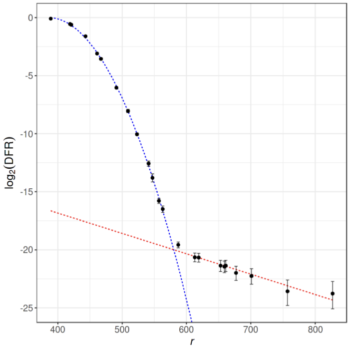


Fig. 1: Decoding failure rates as in Table 1 on a semi-log graph, with a quadratic best fit (blue) in the waterfall region $r < 587$ and a linear best fit (red) in the error floor region $r \geq 587$.

How can we get closer to an analysis of BIKE decoding failures?

[5] *A Study of Error Floor Behavior in QC-MDPC Codes*, Sarah Arpin, Tyler Raven Billingsley, Daniel Rayor Hast, Jun Bo Lau, Ray Perlner, and Angela Robinson (2022)

Near codewords

Definition

Let $\mathbf{H}$ be a parity-check matrix describing a code $\mathcal{C}$. A (u, v)-near codeword is an error vector $\mathbf{e}$ of weight u whose syndrome $\mathbf{s} = \mathbf{H}\mathbf{e}^T$ has weight v.

- McKay, Postol (2003): near codewords with small u, v and low-weight codewords cause high error floor for certain LDPC codes.
  - Basic intuition: Iterative decoders try to push $\Delta\mathbf{e}$ to 0 by decreasing $|\Delta\mathbf{s}|$
  - But $|\Delta\mathbf{s}|$ can get stuck at a local minimum ($\Delta\mathbf{e}$ is codeword or near codeword)

Marco Baldi. QC-LDPC Code-Based Cryptography (2014)
David J.C. MacKay, Michael S. Postol. Weaknesses of Margulis & Ramanujan-Margulis Low-Density Parity-Check Codes (2003)
Tom Richardson. Error floors of LDPC codes (2003)
Gerd Richter. Finding small stopping sets in the Tanner graphs of LDPC codes (2006)

The set $\mathcal{N}$ of near codewords

[Vas21][6] defines an important set of (d, d)-near codewords for QC-MDPC codes:

Definition

Let $H = [\mathbf{H_0}|\mathbf{H_1}]$ have polynomial representation $(h_0(x), h_1(x))$.

$$\mathcal{N} := \{(x^s h_0(x), 0) : s \in \{0, 1, ..., r-1\}\} \cup \{(0, x^s h_1(x)) : s \in \{0, 1, ..., r-1\}\} \subseteq \mathbb{F}_2^n.$$

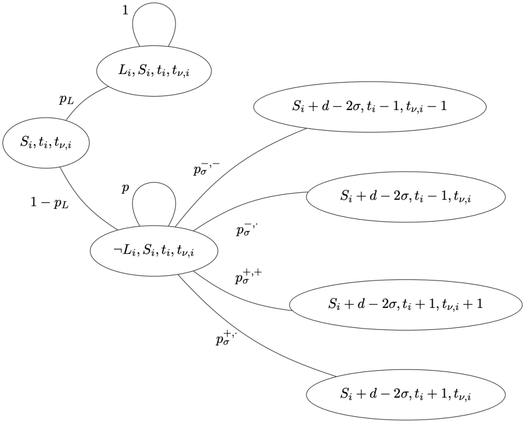
(Vectors of the form: half from a row of ${\mathbf{H_i}}^T$ and the other half 0's.)

[ABHLPR22][7] Finds convergence to $\mathcal{N}$ is dominant behavior in QC-MDPC error floors.

[6] *Post-quantum cryptography: a study of the decoding of QC-MDPC codes* Valentin Vasseur (2021)
[7] *A Study of Error Floor Behavior in QC-MDPC Codes*, Sarah Arpin,Tyler Raven Billingsley, Daniel Rayor Hast, Jun Bo Lau, Ray Perlner, and Angela Robinson (2022)

How we add the effect of near codewords to the Markov Model

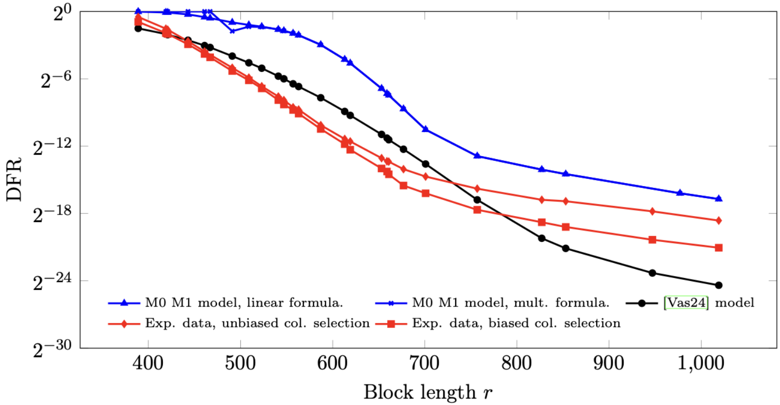
Fix a near codeword ν.
$(S_i, t_i, t_{\nu,i})$ = state at iteration i of decoder.
t_ν keeps track of overlaps with a near codeword ν.
L = blocked state.



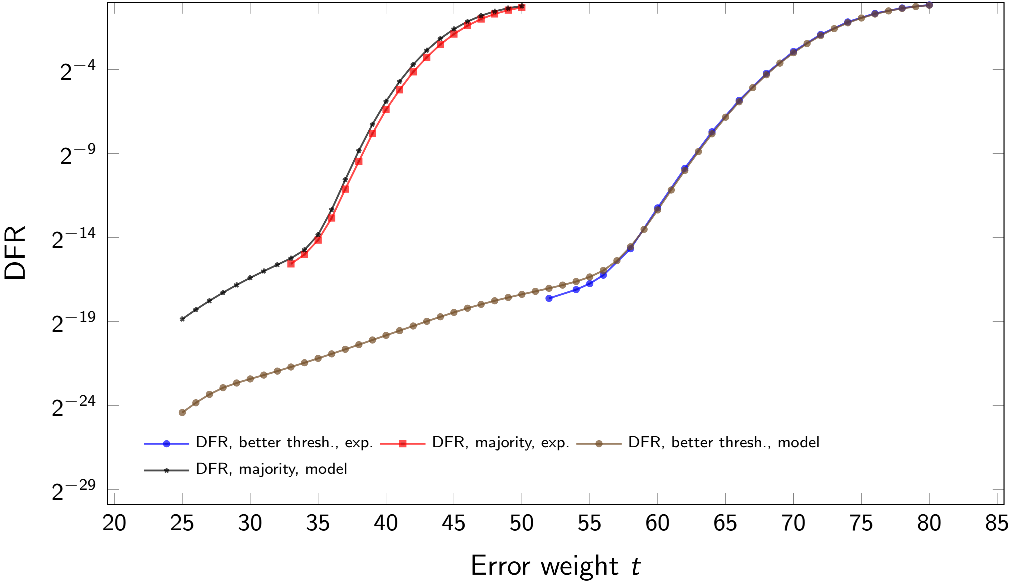
Our two Markov-based models for DFR

- Model 1:
  - Extrapolates DFR from effect of a single arbitrarily chosen $\nu \in \mathcal{N}$.
  - Retains a fudge factor $\xi = 0.955$ from [Vas21] refinement of [SV19b].
  - Uses simplified heuristics to model "average key".
  - State is $(s, t, u) = (|\Delta\mathbf{s}|, |\Delta\mathbf{e}|, |\Delta\mathbf{e} \star \nu|)$
- Model 2:
  - Models DFR directly from effect of nearest $\nu \in \mathcal{N}$ to $\Delta\mathbf{e}$.
  - Does not use ξ (equivalent to $\xi = 1$).
  - Models DFR for specific key using "key shape" info collected from its Tanner Graph.
  - State is (s, t, u, b), where b indicates which half of ν is nonzero.

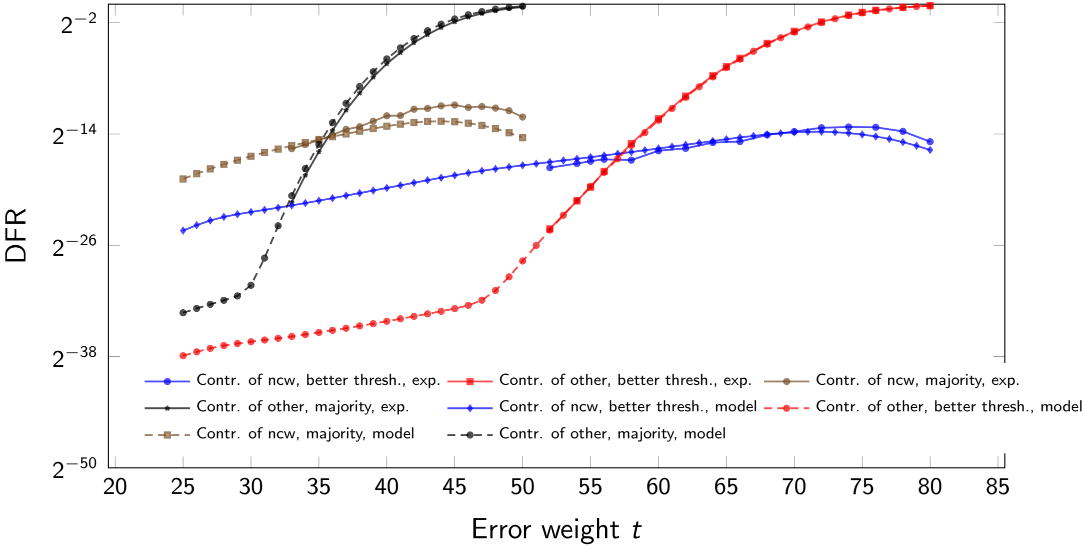
Model 1 DFR vs experiment



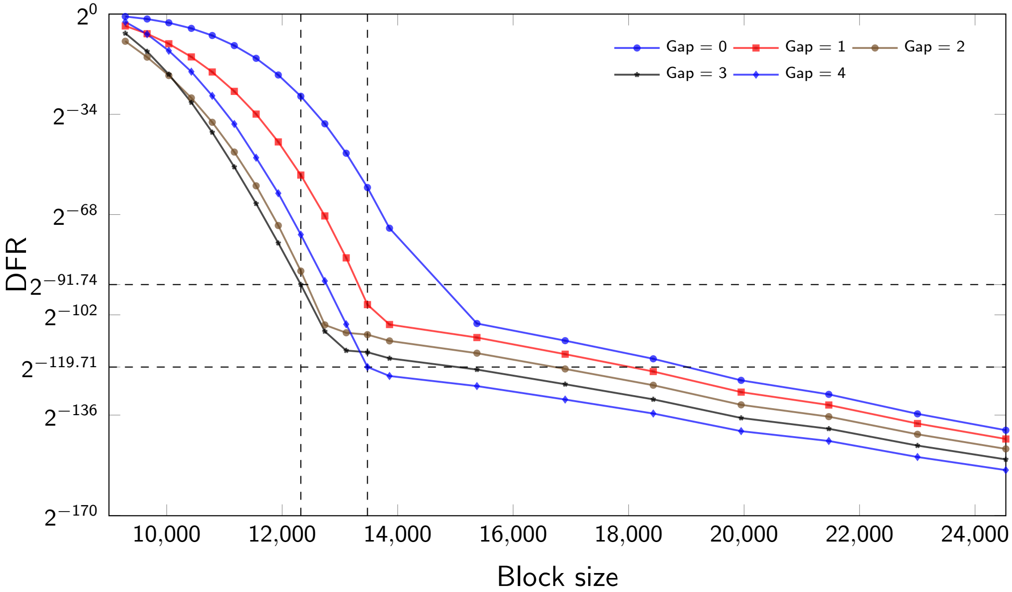
Experiment vs. Model 2



Experiment vs. Model 2 (II)



BIKE parameter 1



Conclusion

- Our techniques allow for accurate predictions of QC-MDPC DFRs, including in the error floor region.
- Our model takes key shape into account which can enable filtering out weak keys.
- We show that only a small modification (block size $+ 10\%$) is needed to make BIKE1 parameters convincingly IND-CCA2 secure.
- Future work may extend these results to parallel decoders like BGF, which seem to perform better than the step-by-step decoders we consider.

Thank you!