# Improved side channel attacks and security estimation on Kyber

Russell Chiu [1] Dana Dachman-Soled [1] Santosh Ghosh [2] Hari Kailad [1] Hunter Kippen [1]
Rishub Nagpal [3] Avery Parker [1] Robert Primas [2]

[1]University of Maryland

[2]Intel Labs, Hillsboro, USA

[3]Graz University of Technology, Austria

## Introduction

CRYSTALS Kyber is one of the first "quantum-resistant" key encapsulation mechanisms (KEM) originally submitted to the NIST competition, now standardized by NIST as ML-KEM [4]. Therefore, understanding the side channel security of Kyber is critical.

In prior work, a framework was introduced to estimate the security of **LWE** with Hints [1]. In this framework, the **LWE** problem is first reduced to **DBDD**, an intermediate problem between **LWE** and **uSVP**. The **DBDD** problem could then be reduced further by integrating side information, or hints. With a further reduction to **uSVP**, one can then estimate the BKZ blocksize to solve the uSVP instance.

Prior attacks [3] have focused on obtaining side channel information from the inverse NTT applied during decryption. In this work, we introduce a new model, which allows us to embed the Kyber **LWE** problem as a **DBDD** instance. This allows us to estimate security with side channel information on the NTT and perform attacks.

## Embedding the Kyber LWE Instance

Consider a single Kyber LWE equation $\mathbf{a} \cdot \mathbf{s} + \mathbf{e} = \mathbf{b}$ over $\mathcal{R}_q$. In the NTT domain, this can be written as

$$\hat{\mathbf{a}} \circ \hat{\mathbf{s}} + \hat{\mathbf{e}} = \hat{\mathbf{b}}$$

Structure of the Kyber NTT: Allows us to split $\hat{\mathbf{s}}$ into the even and odd coordinates, $\hat{\mathbf{s}}_E$ and $\hat{\mathbf{s}}_O$.

$$\hat{\mathbf{s}}_E = \mathbf{s}_E \mathbf{V}_{half}$$

$$\hat{\mathbf{s}}_O = \mathbf{s}_O \mathbf{V}_{half}$$

Let $\hat{\mathbf{u}}$ be a ciphertext in the NTT space. Then the product is equal to

$$\hat{\mathbf{s}} \circ \hat{\mathbf{u}} = \hat{\mathbf{s}}\mathbf{U}$$

Here, for each product $(\hat{s}_{2i}X + \hat{s}_{2i+1})(\hat{u}_{2i}X + \hat{u}_{2i+1})$, we let

$$\mathbf{U}_i = \begin{pmatrix} \hat{u}_{2i} & \hat{u}_{2i+1} \cdot \zeta^{2 \cdot \mathrm{br}(i)+1} \\ \hat{u}_{2i+1} & \hat{u}_{2i} \end{pmatrix}$$

$$\mathbf{U} = \begin{pmatrix} \mathbf{U_1} & & \\ & \ddots & \\ & & \mathbf{U_{128}} \end{pmatrix}$$

Then we obtain the following (two) equations for the even and odd respectively (even coordinates only displayed here):

$$\mathbf{s}_E \mathbf{V}_{half} \mathbf{\Pi}_{\mathbf{U}_E} - (\hat{\mathbf{s}} \circ \hat{\mathbf{u}})(\mathbf{U}^{\sim})_E = \mathbf{0}$$

We can then transform this to a nonstandard LWE instance by converting to a block matrix.

$$(\mathbf{s}_E \mid\mid (\hat{\mathbf{s}} \circ \hat{\mathbf{u}})) \begin{bmatrix} \mathbf{V}_{half} \mathbf{\Pi}_{\mathbf{U}_E} \\ -(\mathbf{U}^{\sim})_E \end{bmatrix}$$

By row reducing and factoring out the projection, we get the following.

$$(\mathbf{s}_E \mathbf{A}_E + (\hat{\mathbf{s}} \circ \hat{\mathbf{u}})) \mathbf{\Pi}_{\mathbf{U}_E} = \mathbf{0}$$

Note this looks like an LWE instance – however, we do not have the guarantee of a unique short vector anymore, as the coordinates of the error are no longer very short!

## Hint Integration

Let $\mathbf{x} = (\mathbf{s}_E \mid\mid (\hat{\mathbf{s}} \circ \hat{\mathbf{u}}))$.

- Perfect Hints.

$$\langle \mathbf{x}, \mathbf{v} \rangle = \gamma$$

  Exact guesses on NTT coordinates can be modeled as perfect hints where $\mathbf{v}$ is a unit basis vector.
- Approximate Hints.

$$\langle \mathbf{x}, \mathbf{v} \rangle \approx \gamma$$

  Noisy side channel information on an NTT coordinate can be modeled as an approximate hint.
- Short Vector Hints. Knowledge of $\mathbf{v}$ such that $\mathbf{v} \in \Lambda$ from the DBDD instance is short.

Pathological Short Vectors.

If only the first 64 coordinates are nonzero, note that the NTT of

$$\Phi_{32}(n) = \prod_{i=0}^{32} x - \zeta^{2 \cdot \mathrm{br}(i)+1}$$

is zero.

Integrating these pathological NTT structure based short vectors into the instance and projecting on them improves BKZ estimates. Reduces predicted BKZ blocksize by around 40.

## Preliminary Results

Experimental data on 64 nonzero NTT coordinates with $n$ guesses (the number of perfect coordinate hints).

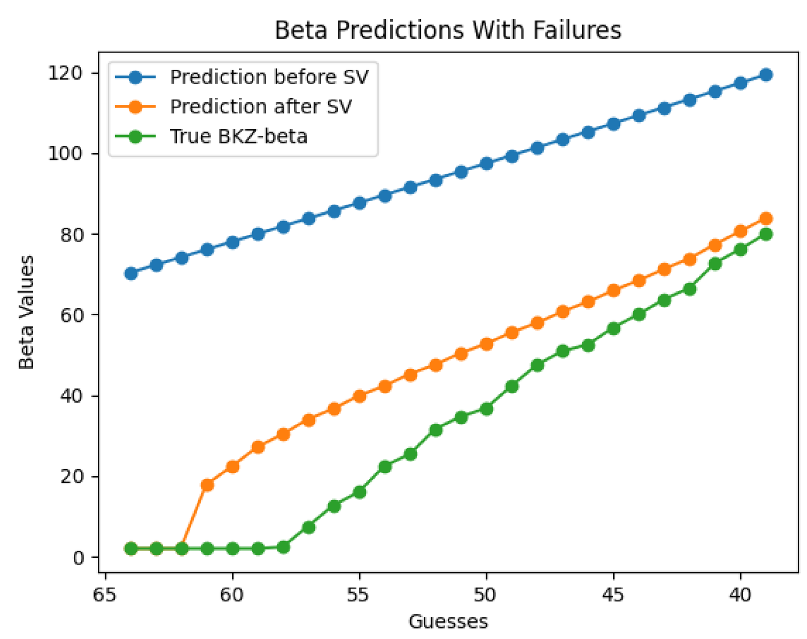


Figure 1. BKZ prediction with and without SV Hints at different guess values

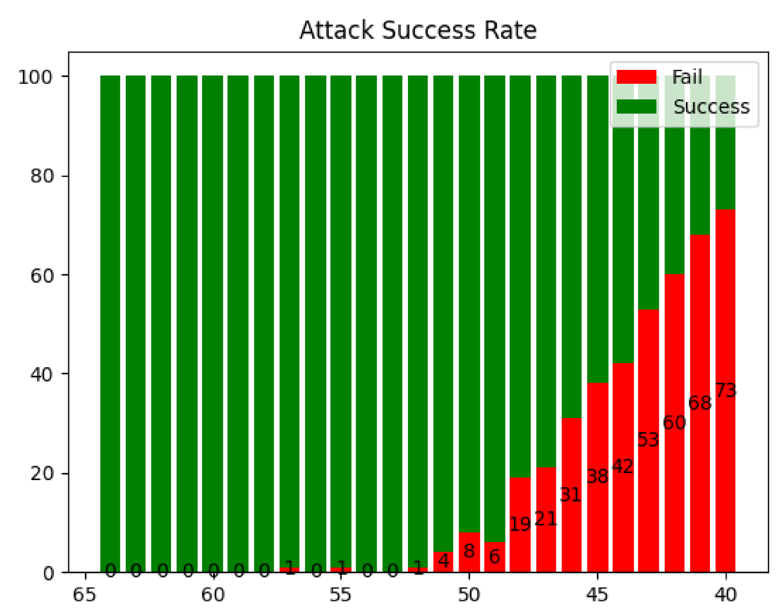


Figure 2. BKZ Success / Failure with number of guesses

## Future Work

- Test with different locations for nonzero values.
- BKZ is effectively able to emulate the algebraic attack from [2]. First 32 NTT coefficients correspond to every fourth coefficient (bit reversed order). Groups of 4 independent short vector hints span the 4 nonzero coordinates – projecting on these short vectors will set them to zero, effectively recovering the shortest solutions to the $1 \times 4$ linear system given in [2].

## References

[1] Dana Dachman-Soled, Léo Ducas, Huijing Gong, and Mélissa Rossi. LWE with side information: Attacks and concrete security estimation. Cryptology ePrint Archive, Paper 2020/292, 2020. `https://eprint.iacr.org/2020/292`.

[2] Dana Dachman-Soled, Huijing Gong, Mukul Kulkarni, and Aria Shahverdi. (in)security of ring-lwe under partial key exposure. *Journal of Mathematical Cryptology*, 15(1):72–86, 2021.

[3] Mike Hamburg, Julius Hermelink, Robert Primas, Simona Samardjiska, Thomas Schamberger, Silvan Streit, Emanuele Strieder, and Christine van Vredendaal. Chosen ciphertext k-trace attacks on masked CCA2 secure kyber. Cryptology ePrint Archive, Paper 2021/956, 2021. `https://eprint.iacr.org/2021/956`.

[4] National Institute of Standards and Technology. Module-lattice-based key-encapsulation mechanism standard. Federal Information Processing Standards Publication (FIPS) NIST FIPS, 2023. `https://doi.org/10.6028/NIST.FIPS.203.ipd`.

## Acknowledgements

⋆ Supported in part by NSF grant #CNS-2154705 and by Intel through the Intel Labs Crypto Frontiers Research Center.