Security for Dilithium and Falcon in the QROM

September 14, 2021, NIST Postquantum Crypto Seminar

Carl A. Miller

(Not for public distribution.)

Question: What do security proofs get us?

Goal for Talk

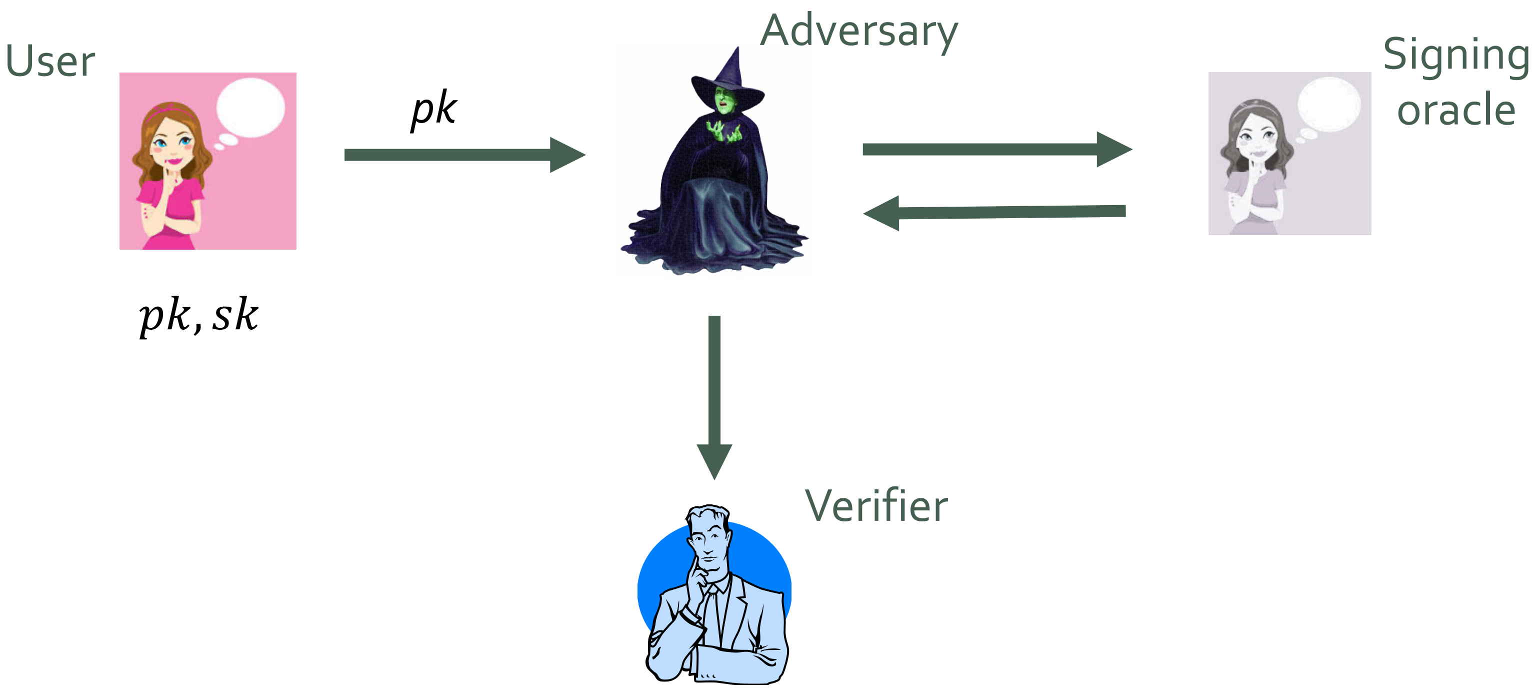
Identify all the underlying assumptions for the security of Falcon and Dilithium.

We’ll focus just on “theoretical” security.
(Side-channel attacks are out of scope.)

Quick Review of Models

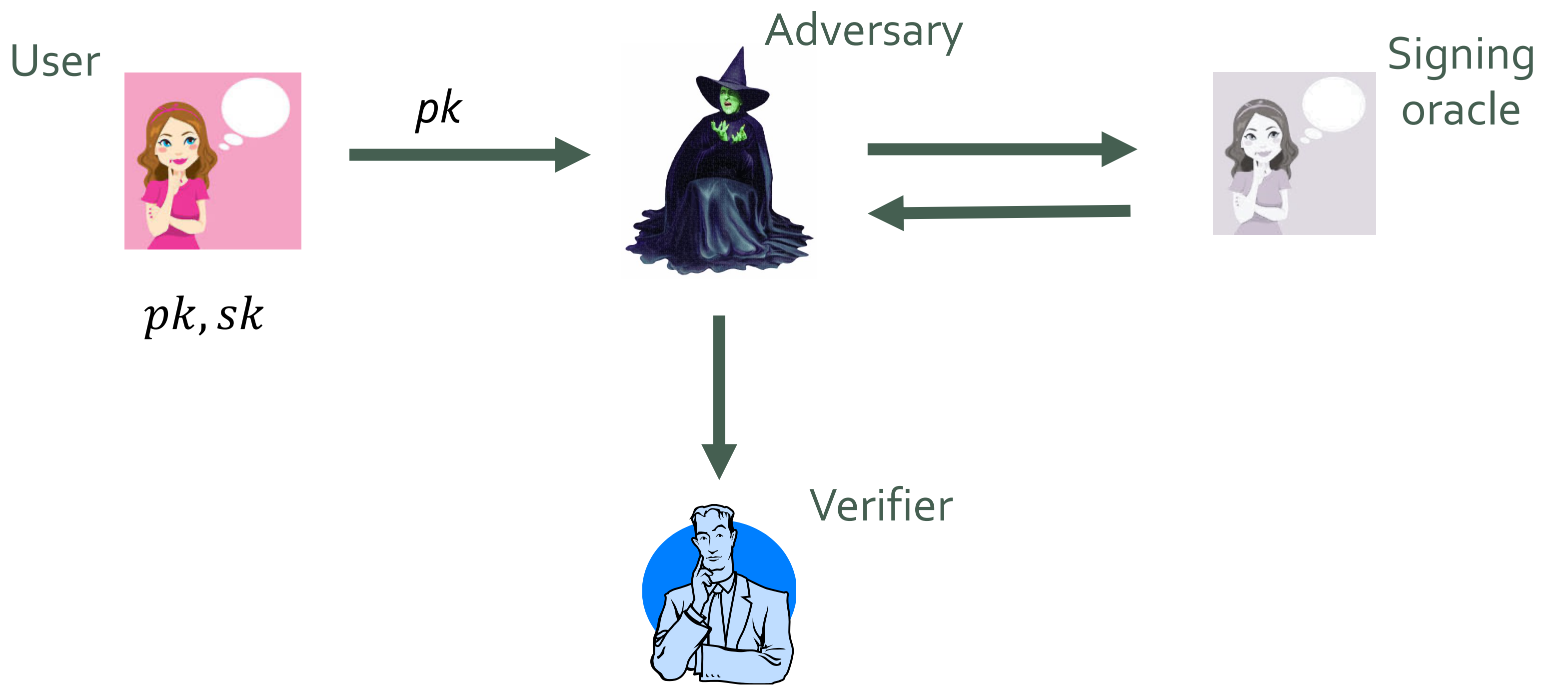
Security Models for Signatures

Goal: Prove that Adversary cannot forge a signature to any message *other* than those signed by the oracle. **(EUF-CMA)**



Security Models for Signatures

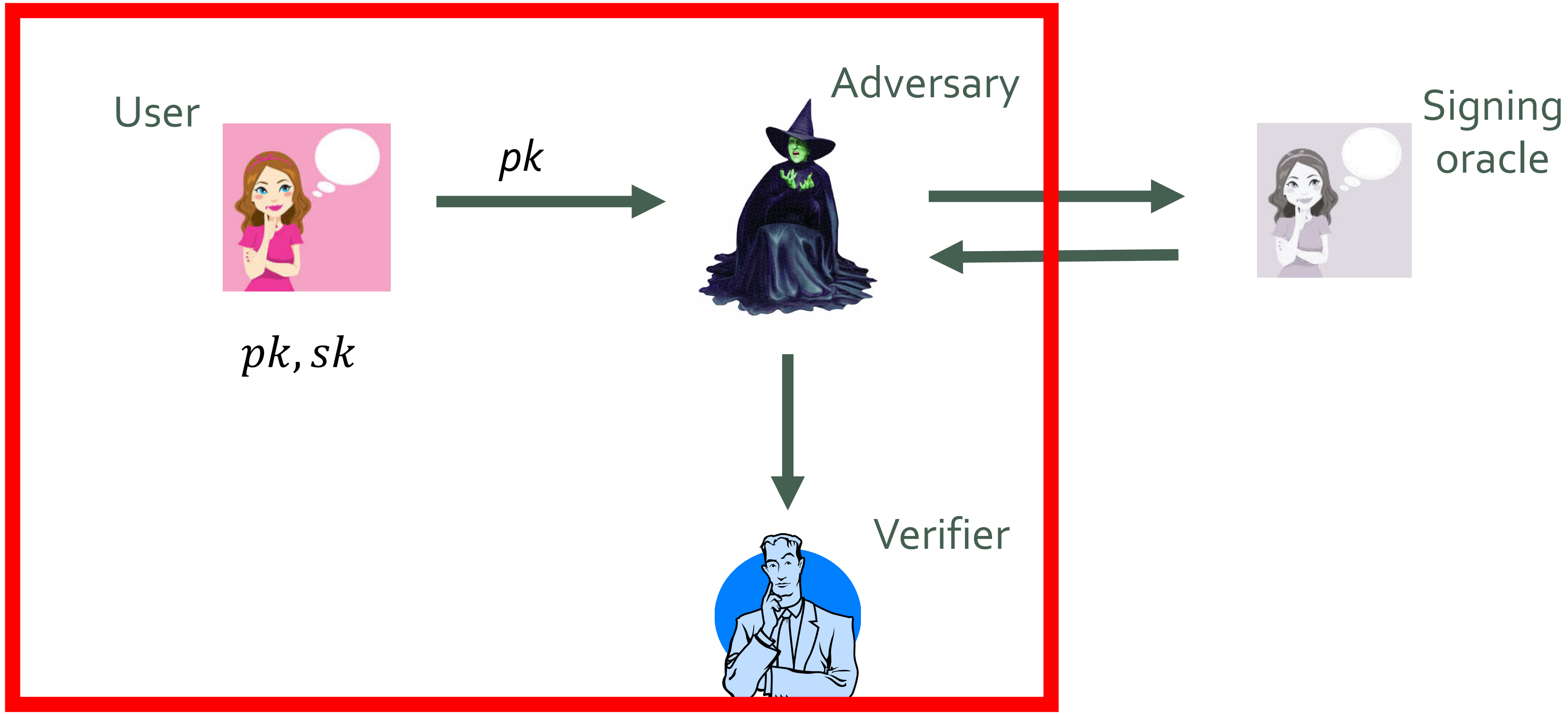
Better yet, prove that Adversary can neither sign a new message nor create a new signature for an old message. **(SUF-CMA)**



Security Models for Signatures

Security can be divided into 2 parts:

(1) Prove that forging is impossible without the signing oracle. **(EUF-NMA)**

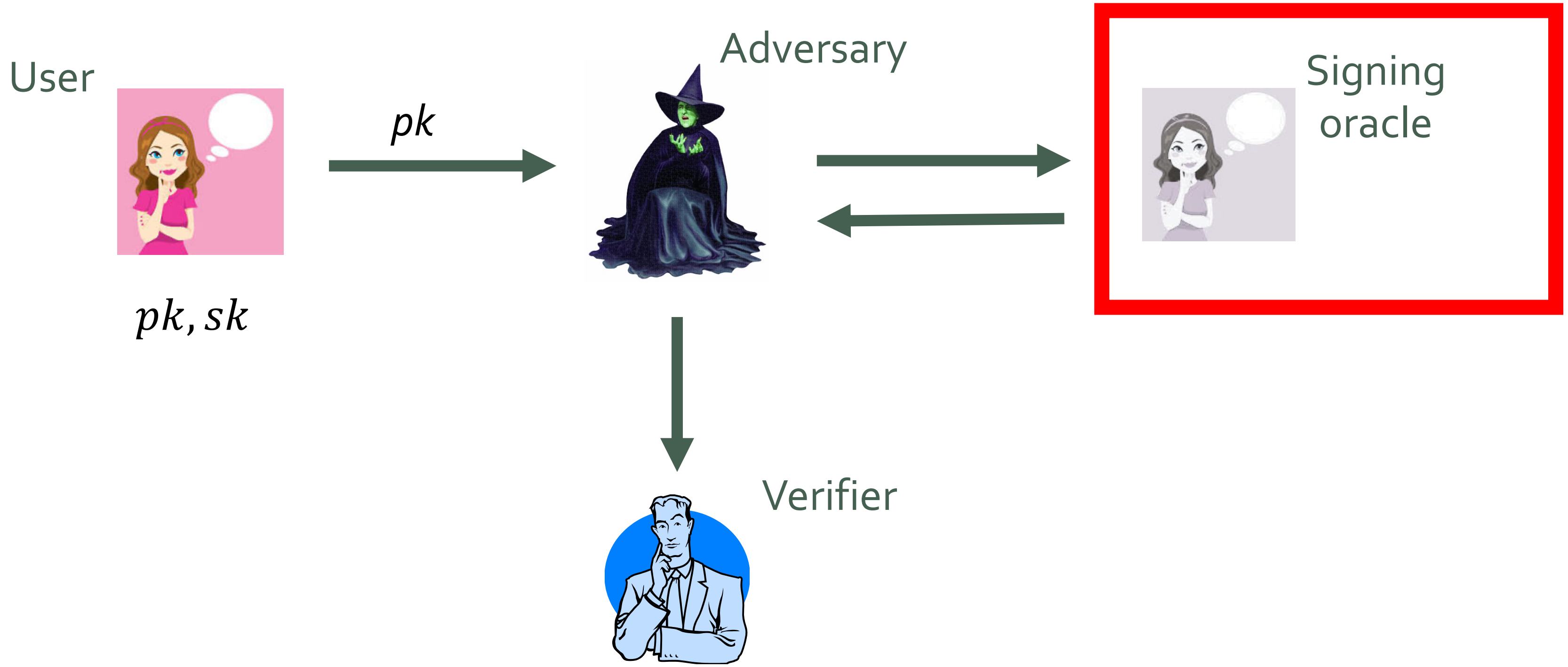


Security Models for Signatures

Security can be divided into 2 parts:

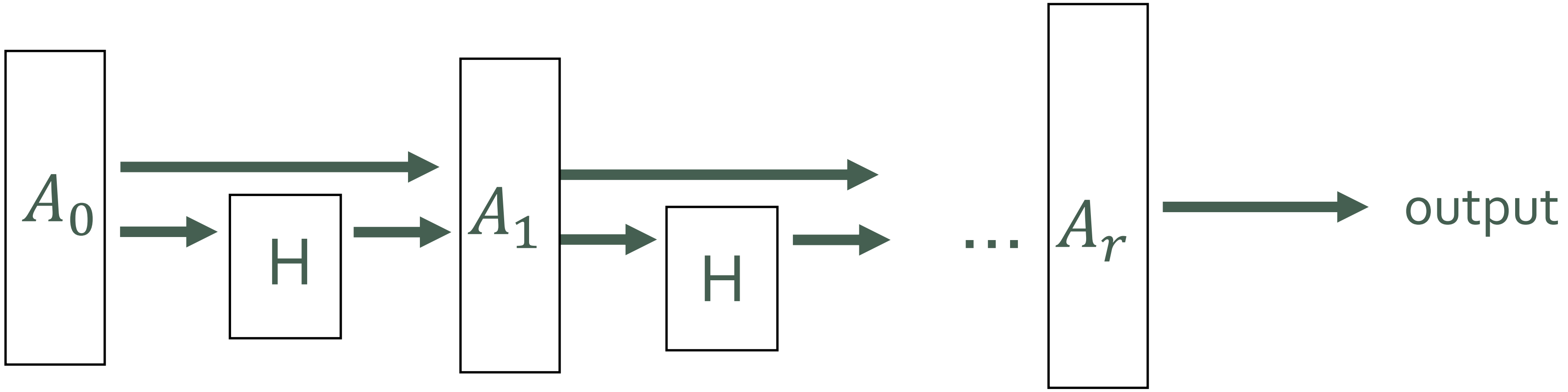
(1) Prove that forging is impossible without the signing oracle. **(EUF-NMA)**

(2) Prove that the signing oracle does not help.



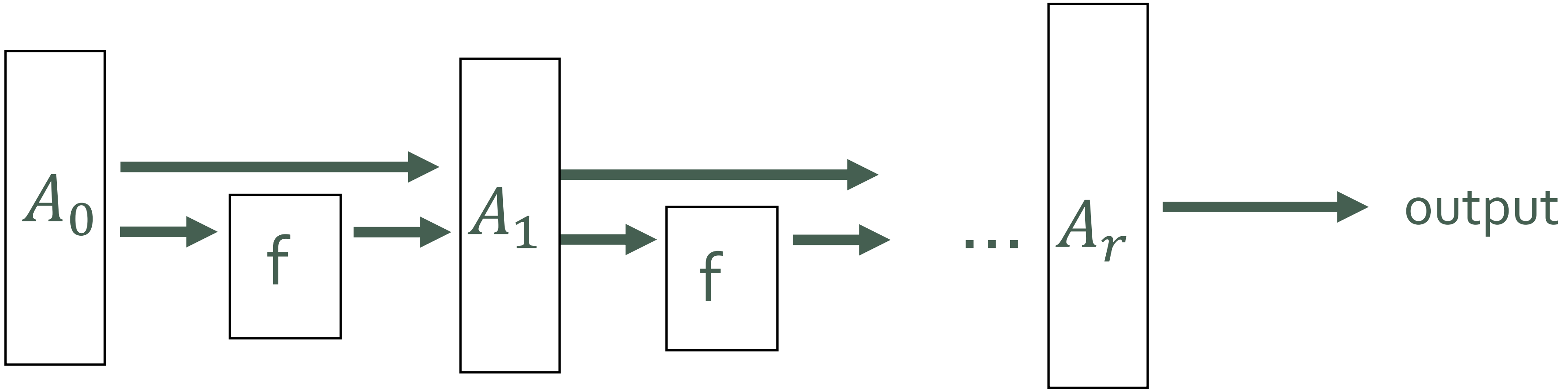
The QROM

Suppose that a probabilistic algorithm A uses a hash function H.



The QROM

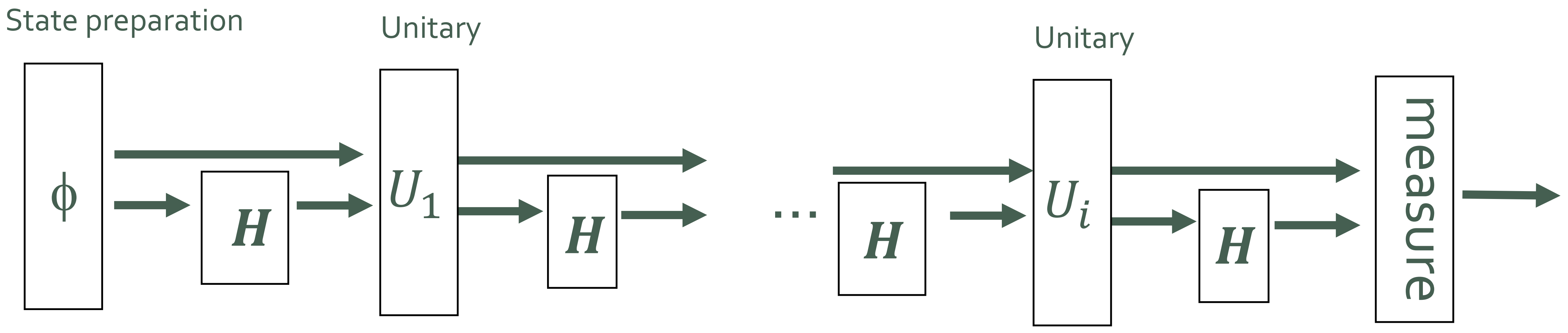
Suppose that a probabilistic algorithm A uses a hash function H.
In the ROM, a truly random function f is chosen and each instance of H is replaced by f.



The QROM

Now suppose B is a quantum algorithm, where $\boldsymbol{H}$ denotes

$$\boldsymbol{H}\,(\,|x\rangle|y\rangle) = |(\,|x\rangle|y \oplus \mathrm{H}(x)\rangle\rangle.$$



The QROM

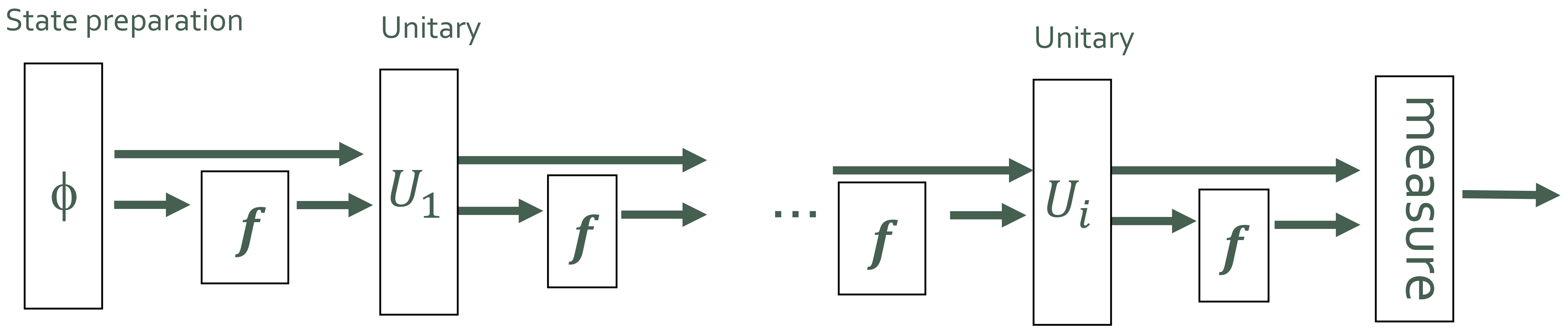
Now suppose B is a quantum algorithm, where $\boldsymbol{H}$ denotes

$$\boldsymbol{H}\,(\,|x\rangle|y\rangle) = |(\,|x\rangle|y \oplus \mathrm{H}(x)\rangle\rangle).$$

The QROM replaces $\boldsymbol{H}$ with

$$\boldsymbol{f}\,(\,|x\rangle|y\rangle) = |(\,|x\rangle|y \oplus \mathrm{f}(x)\rangle\rangle)$$

D. Boneh et al., "Random Oracles in a Quantum World". (2011)

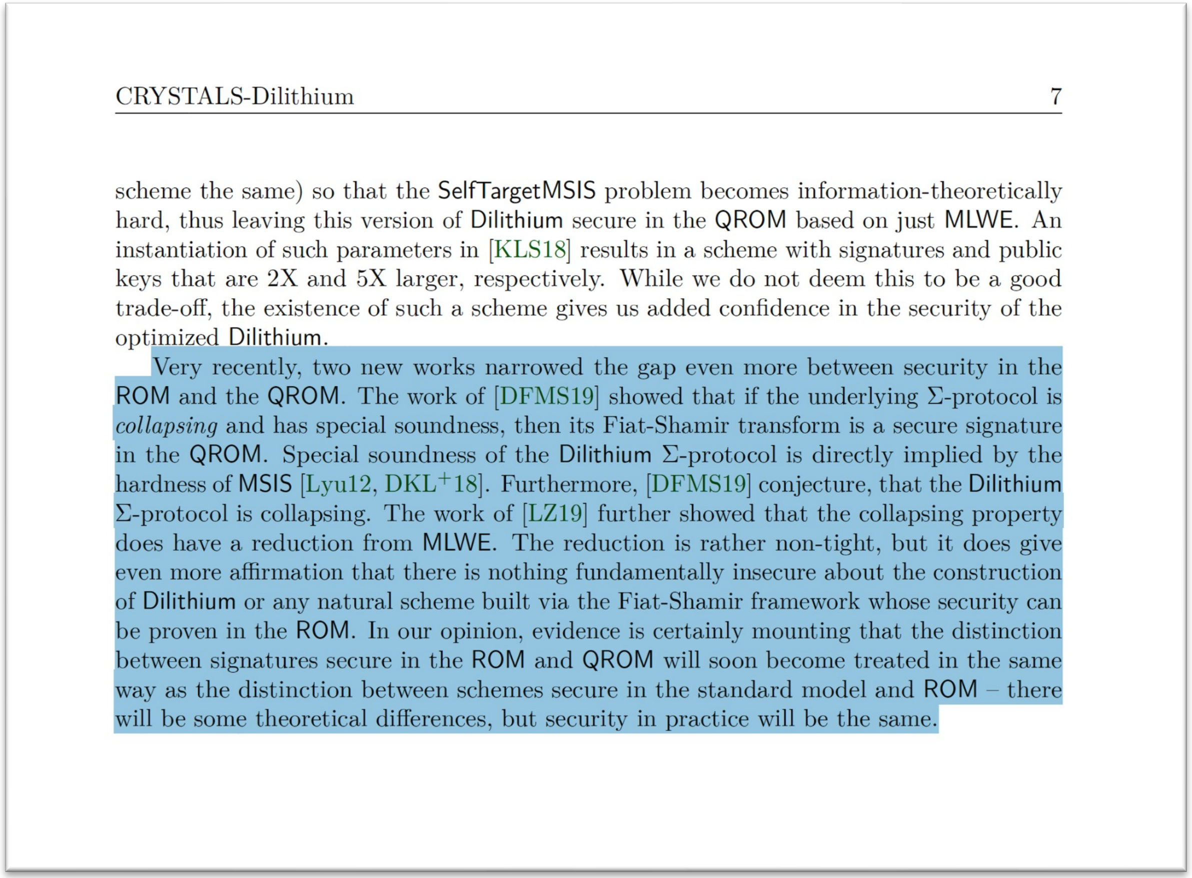


Security Proofs for Dilithium and Falcon

Sources

The Round 3 description of Dilithium has a detailed discussion of security.

I'm ignoring this paragraph (from section 1) for now, since it seems speculative.



CRYSTALS-Dilithium 7

scheme the same) so that the SelfTargetMSIS problem becomes information-theoretically hard, thus leaving this version of Dilithium secure in the QROM based on just MLWE. An instantiation of such parameters in [KLS18] results in a scheme with signatures and public keys that are 2X and 5X larger, respectively. While we do not deem this to be a good trade-off, the existence of such a scheme gives us added confidence in the security of the optimized Dilithium.

Very recently, two new works narrowed the gap even more between security in the ROM and the QROM. The work of [DFMS19] showed that if the underlying Σ-protocol is *collapsing* and has special soundness, then its Fiat-Shamir transform is a secure signature in the QROM. Special soundness of the Dilithium Σ-protocol is directly implied by the hardness of MSIS [Lyu12, DKL$^+$18]. Furthermore, [DFMS19] conjecture, that the Dilithium Σ-protocol is collapsing. The work of [LZ19] further showed that the collapsing property does have a reduction from MLWE. The reduction is rather non-tight, but it does give even more affirmation that there is nothing fundamentally insecure about the construction of Dilithium or any natural scheme built via the Fiat-Shamir framework whose security can be proven in the ROM. In our opinion, evidence is certainly mounting that the distinction between signatures secure in the ROM and QROM will soon become treated in the same way as the distinction between schemes secure in the standard model and ROM – there will be some theoretical differences, but security in practice will be the same.

Sources

The Falcon description says less about security, but it seems like the proof can be put together using these papers.

C. Gentry, et al. "Trapdoors for Hard Lattices and New Cryptographic Constructions." (2008)

D. Boneh et al., "Random Oracles in a Quantum World". (2011)

EUF-CMA Security Arguments

Falcon

- NTRU-SIS
- Floating point precision
- QROM and other hash assumptions

Dilithium

- SelfTargetMSIS
- MLWE
- QROM and other hash assumptions

EUF-CMA Security Arguments

Falcon	Dilithium
NTRU-SIS	SelfTargetMSIS
Floating point precision	MLWE
QROM and other hash assumptions	QROM and other hash assumptions

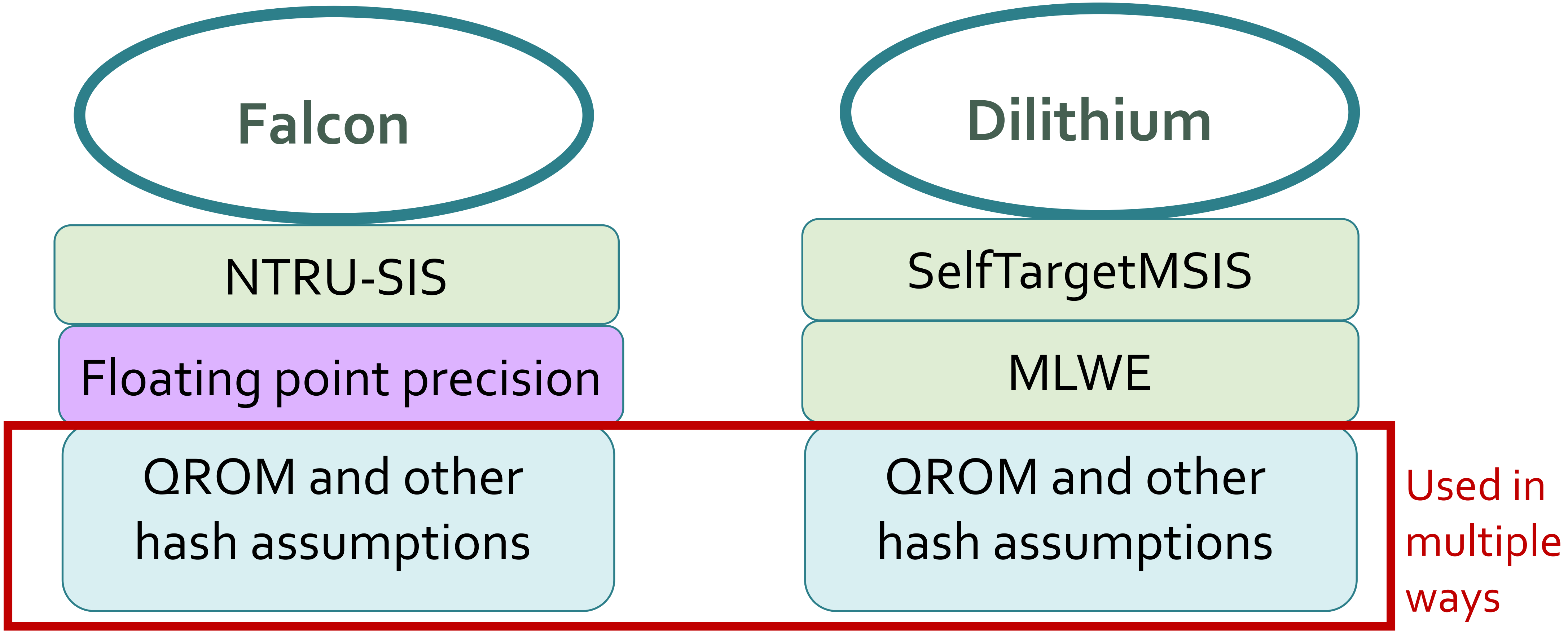
Proving EUF-NMA security

EUF-CMA Security Arguments

Falcon	Dilithium
NTRU-SIS	SelfTargetMSIS
Floating point precision	MLWE
QROM and other hash assumptions	QROM and other hash assumptions

Proving that the signing oracle does not help

EUF-CMA Security Arguments



NTRU-SIS (Falcon)

Let $R_q = \mathbb{Z}_{12289}[X]/(X^{1024} + 1)$.
Forging a single signature is equivalent to finding a solution

$$s_1, s_2 \in R_q,$$

with small Euclidean norm, to a random equation of the form

$$s_1 + s_2 h = c$$

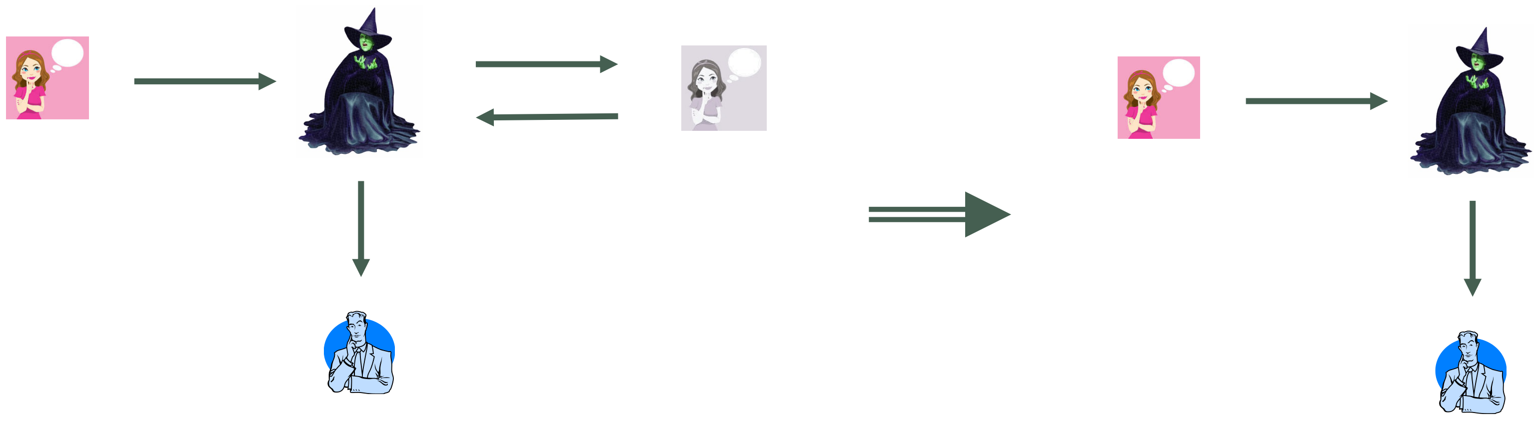
$h = gf^{-1}$, where g, f have random Gaussian coefficients

Uniformly random

Note: The authors actually state a different problem: namely, compute (f', g') with small coefficients such that $h = (g')(f')^{-1}$. Is that equivalent?

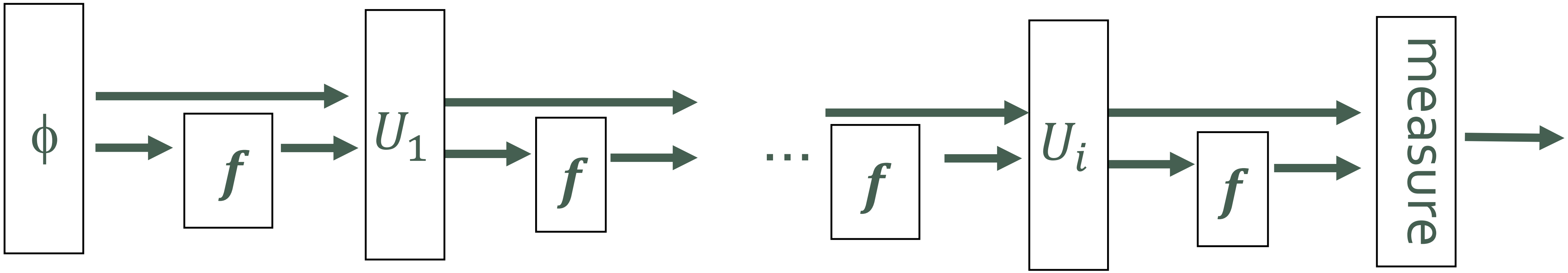
Floating Point Precision (Falcon)

For any EUF-CMA attack strategy, there is a EUF-NMA attack strategy that succeeds with equal probability.



But, that assumes infinite floating point precision.
The authors argue that (if $\leq 2^{64}$ queries are assumed) 53 bits of floating precision is sufficient to still maintain security.

The QROM Assumption



Simple.

Well-studied.

Not actually true.

Other Hash Assumptions

The protocols use hashes and extended output functions, and make assumptions about them. E.g., Dilithium says:

[5]To simplify the concrete security bound, we assume that ExpandA produces a uniform matrix $\mathbf{A} \in R_q^{k\times\ell}$, $\mathsf{ExpandMask}(K,\cdot)$ is a perfect pseudo-random function, and CRH is a perfect collision-resistant hash function.

Question: Can all such assumptions be derived from the QROM assumption?

MLWE (Dilithium)

The MLWE Problem. For integers m, k, and a probability distribution $D : R_q \to [0, 1]$ we say that the advantage of algorithm A in solving the decisional $\mathsf{MLWE}_{m,k,D}$ problem over the ring R_q is

$$\begin{aligned}\mathrm{Adv}^{\mathsf{MLWE}}_{m,k,D} := \big|&\Pr[b = 1 \mid \mathbf{A} \leftarrow R_q^{m\times k}; \mathbf{t} \leftarrow R_q^m; b \leftarrow \mathsf{A}(\mathbf{A}, \mathbf{t})] \\ &- \Pr[b = 1 \mid \mathbf{A} \leftarrow R_q^{m\times k}; \mathbf{s}_1 \leftarrow D^k; \mathbf{s}_2 \leftarrow D^m; b \leftarrow \mathsf{A}(\mathbf{A}, \mathbf{A}\mathbf{s}_1 + \mathbf{s}_2)]\big| .\end{aligned}$$

Here, $R_q = \mathbb{Z}_{8380417}[X]/(X^{256} + 1)$, and D is a uniform distribution over all elements of R_q that have small coefficients.

SelfTargetMSIS (Dilithium)

Let $B_\tau \subseteq R_q$ be the set of elements whose coefficients are from $\{-1,0,1\}$ and which have exactly τ nonzero coefficients.

The SelfTargetMSIS Problem. Suppose that $\mathsf{H} : \{0,1\}^* \to B_\tau$ is a cryptographic hash function. To an algorithm A we associate the advantage function

$$\mathrm{Adv}^{\mathsf{SelfTargetMSIS}}_{\mathsf{H},m,k,\gamma}(\mathsf{A}) :=$$

$$\Pr\left[\begin{array}{l} 0 \le \|\mathbf{y}\|_\infty \le \gamma \\ \wedge\, \mathsf{H}(\mu \,\|\, [\,\mathbf{I} \mid \mathbf{A}\,] \cdot \mathbf{y}) = c \end{array} \;\middle|\; \mathbf{A} \leftarrow R_q^{m \times k}; \left(\mathbf{y} := \begin{bmatrix} \mathbf{r} \\ c \end{bmatrix}, \mu\right) \leftarrow \mathsf{A}^{|\mathsf{H}(\cdot)\rangle}(\mathbf{A})\right].$$

SelfTargetMSIS (Dilithium)

The SelfTargetMSIS Problem. Suppose that $\mathsf{H} : \{0,1\}^* \to B_\tau$ is a random function. To an algorithm A we associate the advantage function

$$\mathrm{Adv}^{\mathsf{SelfTargetMSIS}}_{\mathsf{H},m,k,\gamma}(\mathsf{A}) :=$$

$$\Pr\left[\begin{array}{l} 0 \le \|\mathbf{y}\|_\infty \le \gamma \\ \wedge\, \mathsf{H}(\mu \parallel [\,\mathbf{I} \mid \mathbf{A}\,] \cdot \mathbf{y}) = c \end{array} \;\middle|\; \mathbf{A} \leftarrow R_q^{m\times k};\ \left(\mathbf{y} := \begin{bmatrix}\mathbf{r} \\ c\end{bmatrix}, \mu\right) \leftarrow \mathsf{A}^{|\mathsf{H}(\cdot)\rangle}(\mathbf{A})\right].$$

Basically the adversary is trying to solve

$$H(\boldsymbol{A}w) = w_k \text{ and } \|w\|_\infty \le \gamma$$

But, they are also allowed a salt μ and an additive factor v:

$$H(\mu \,\|v + \boldsymbol{A}w) = w_k \text{ and } \|w\|_\infty, \|v\|_\infty \le \gamma$$

SelfTargetMSIS (Dilithium)

The SelfTargetMSIS Problem. Suppose that $\mathsf{H} : \{0,1\}^* \to B_\tau$ is a random function. To an algorithm A we associate the advantage function

$$\mathrm{Adv}^{\mathsf{SelfTargetMSIS}}_{\mathsf{H},m,k,\gamma}(\mathsf{A}) :=$$

$$\Pr\left[\begin{array}{l} 0 \le \|\mathbf{y}\|_\infty \le \gamma \\ \wedge\, \mathsf{H}(\mu \,\|\, [\,\mathbf{I} \mid \mathbf{A}\,] \cdot \mathbf{y}) = c \end{array} \;\middle|\; \mathbf{A} \leftarrow R_q^{m \times k};\ \left(\mathbf{y} := \begin{bmatrix}\mathbf{r}\\ c\end{bmatrix}, \mu\right) \leftarrow \mathsf{A}^{|\mathsf{H}(\cdot)\rangle}(\mathbf{A})\right].$$

The authors say that – although it's not always explicit – Fiat-Shamir signatures typically rely on complicated assumptions like this one. True?