On Tight Quantum Security of HMAC and NMAC in the Quantum Random Oracle Model

Akinori Hosoyamada (NTT Corporation / Nagoya University)

Tetsu Iwata (Nagoya University)

@Crypto 2021

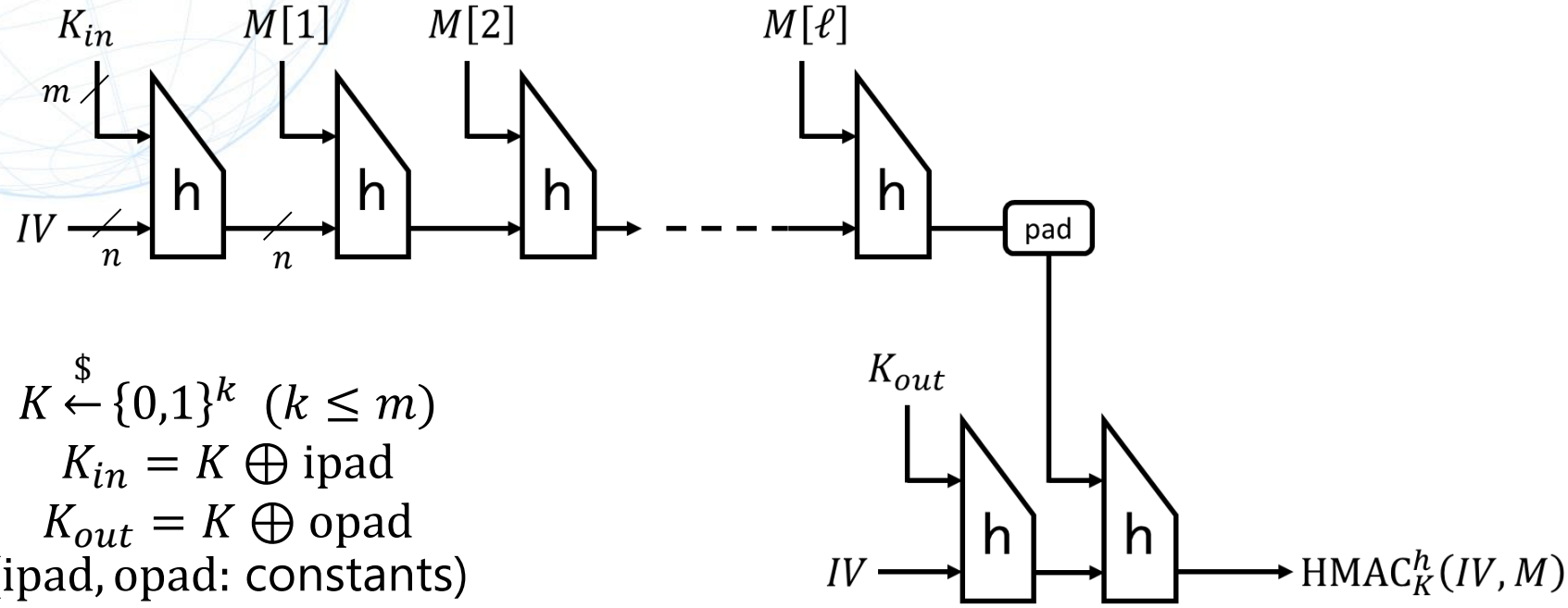
Introduction

Types of MACs

- Block cipher based
  - CBC-MAC, PMAC,…
- Wegman-Carter & polynomial
  - GMAC, Poly1305,…
- Hash based
  - **HMAC/NMAC**, keyed-sponge,…

HMAC (Hash-based MAC)

- Most basic approach to convert Merkle-Damgård hash → MAC
- Standardized in FIPS PUB 198 / Used in TLS, SSH, IPsec,…



$K \xleftarrow{\$} \{0,1\}^k \quad (k \leq m)$

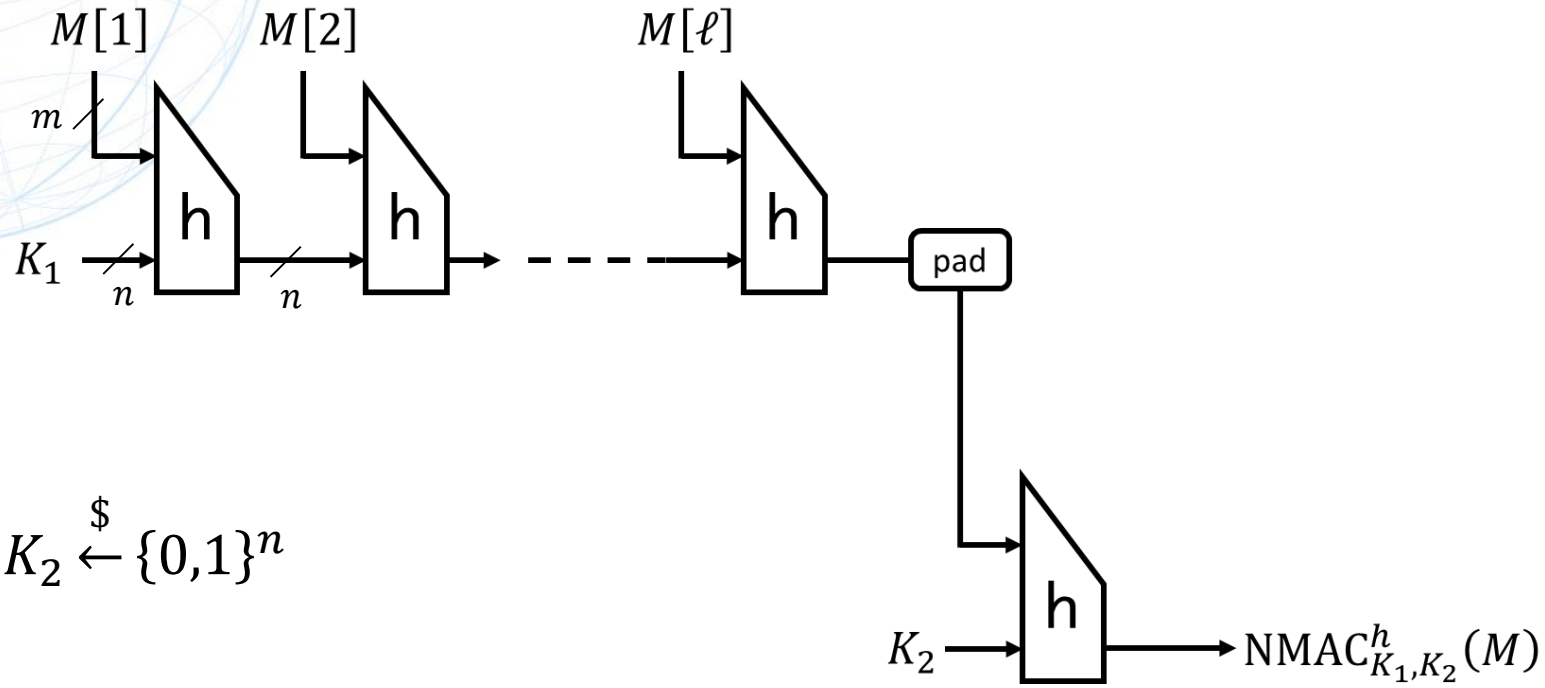
$K_{in} = K \oplus \mathrm{ipad}$

$K_{out} = K \oplus \mathrm{opad}$

(ipad, opad: constants)

NMAC

- Two-key variant of HMAC



$$K_1, K_2 \xleftarrow{\$} \{0,1\}^n$$

Security of HMAC/NMAC

- Security against classical attacks:

  Tight security bound…$O(2^{n/2})$ (n: output length) [GPR14]

- Security against quantum query attacks:

  Secure up to $O(2^{n/5})$ (or $O(2^{n/8})$) queries[*1] (in the standard model) [SY17]

  Trivial attack…$O(2^{n/3})$ queries

[GPR14] Gazi, P., Pietrzak, K., Rybár, M.: The exact prf-security of NMAC and HMAC. (CRYPTO 2014)
[SY17] Song, F., Yun, A.: Quantum security of NMAC and related constructions – PRF domain extension against quantum attacks. (CRYPTO2017)

*1 These security bounds are not explicitly provided in [SY17], but we can reasonably deduce the corresponding security in the QROM is $O(2^{n/5})$ (or $O(2^{n/8})$)

Security of HMAC/NMAC

- Security against classical attacks:

  Tight security bound…$O(2^{n/2})$ (n: output length) [GPR14]

- Security against quantum query attacks:

  Secure up to $O(2^{n/5})$ (or $O(2^{n/8})$) queries[*1] (in the standard model) [SY17]

  Trivial attack…$O(2^{n/3})$ queries

Q. Can we show the tight quantum security bound?

[GPR14] Gazi, P., Pietrzak, K., Rybár, M.: The exact prf-security of NMAC and HMAC. (CRYPTO 2014)
[SY17] Song, F., Yun, A.: Quantum security of NMAC and related constructions – PRF domain extension against quantum attacks. (CRYPTO2017)

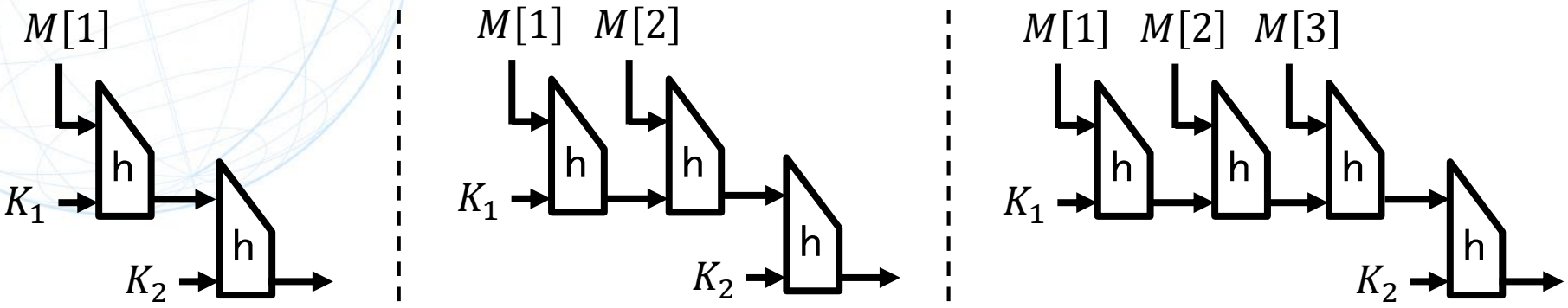
*1 These security bounds are not explicitly provided in [SY17], but we can reasonably deduce the corresponding security in the QROM is $O(2^{n/5})$ (or $O(2^{n/8})$)

Results

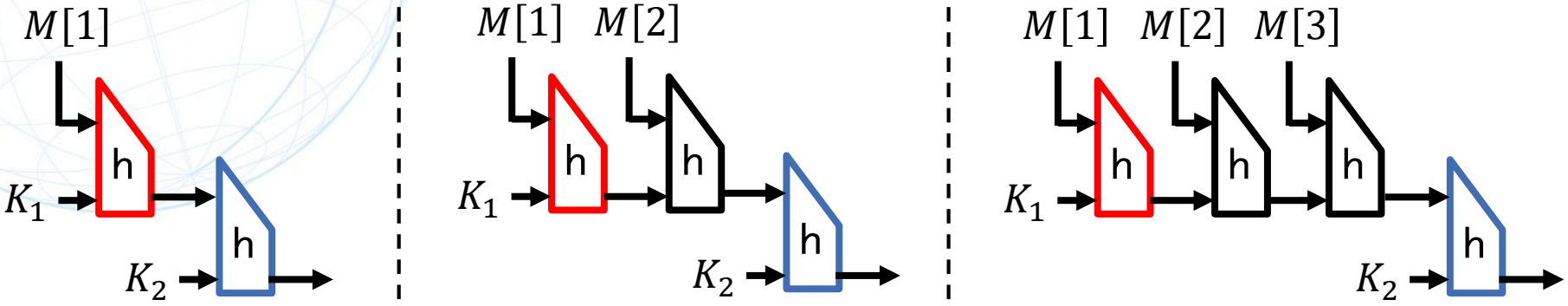
- Tight quantum security bound in the QROM (h : QRO)
- HMAC/NMAC are indistinguishable from a RF against quantum query attacks as long as $(q_h + Q) \cdot \ell^{\frac{3}{5}} \geq 2^{n/3}$
  - q_h: max. num. of quantum queries to h
  - Q: max. num. of quantum queries to the keyed oracle of HMAC/NMAC
  - ℓ: maximum message length
- Tight when ℓ is not exponentially large
- Compressed oracle technique
- Hardest part: Proving the prob. of a bad event is low
  - We introduce a new idea to capture the uncertainty of outputs of a random function that the adversary cannot observe

Rough Overview of the Proof

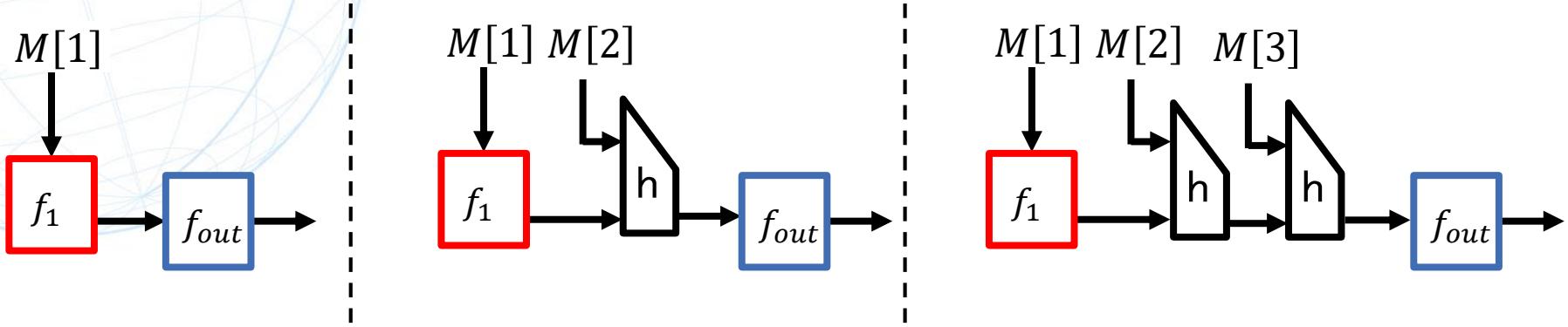
Game G_0 (NMAC)



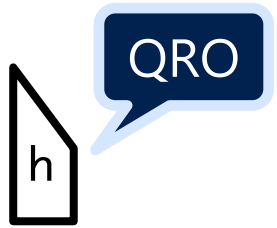
Game G_0 (NMAC)



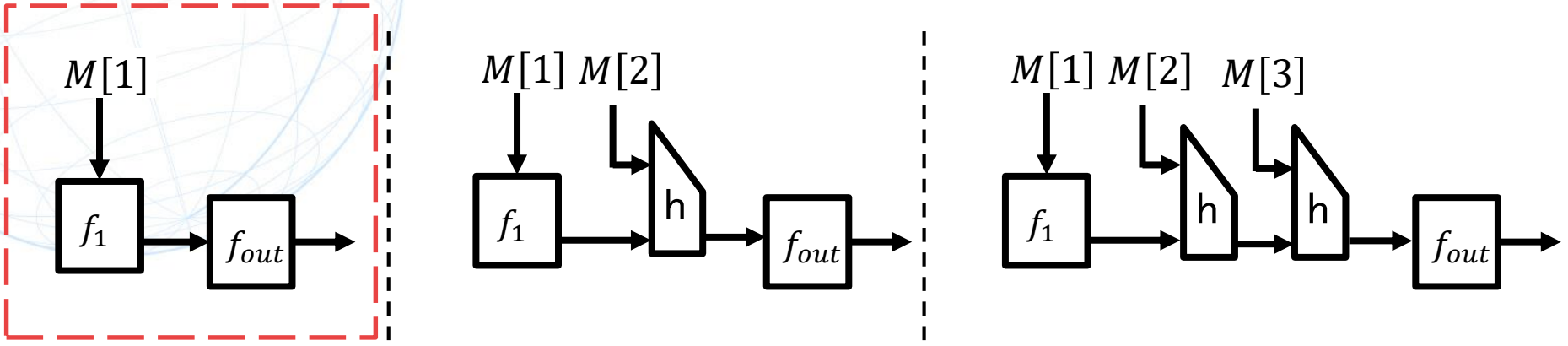
Game G_1

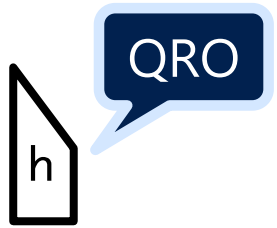


$$h(K_1,\cdot) \Rightarrow f_1$$
$$h(K_2,\cdot) \Rightarrow f_{out}$$

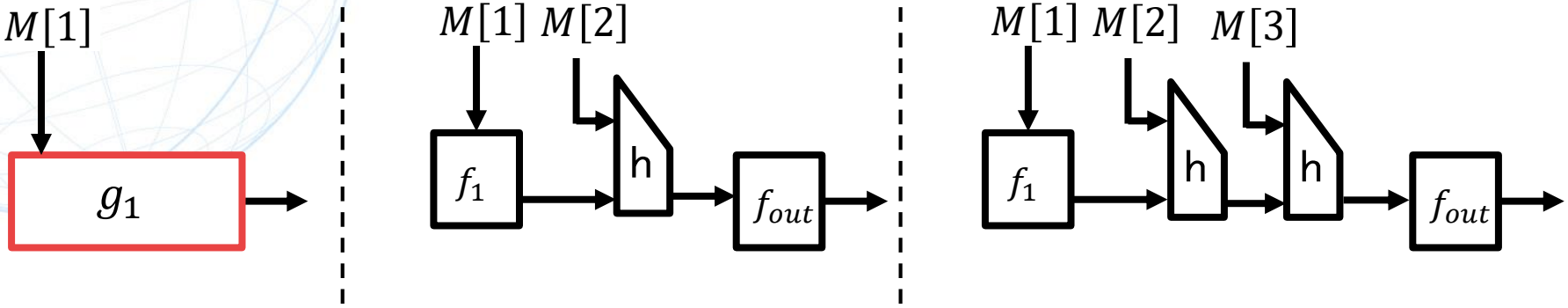


Game G_1

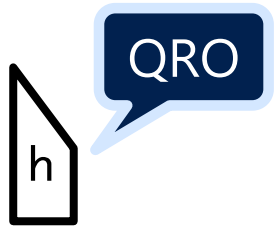




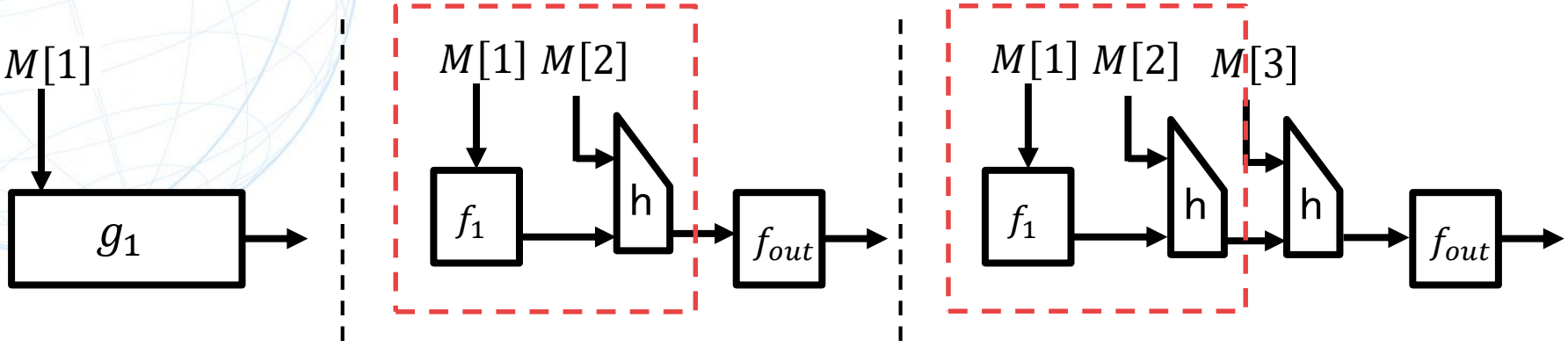
Game G'_1

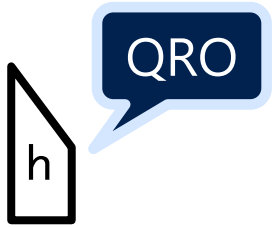


$f_{out} \circ f_1 \Rightarrow g_1$

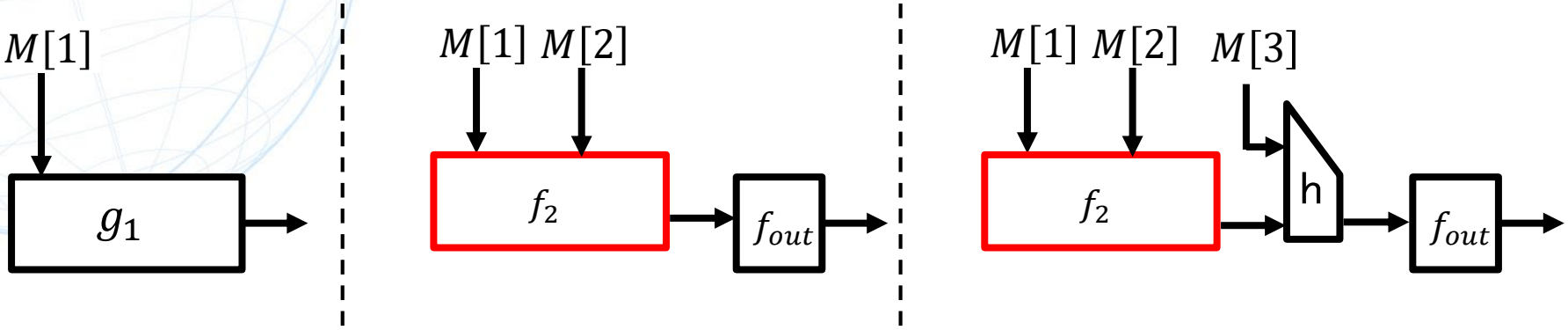


Game G'_1

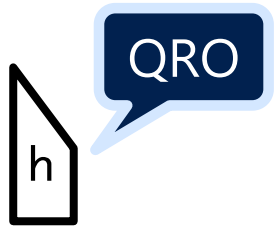




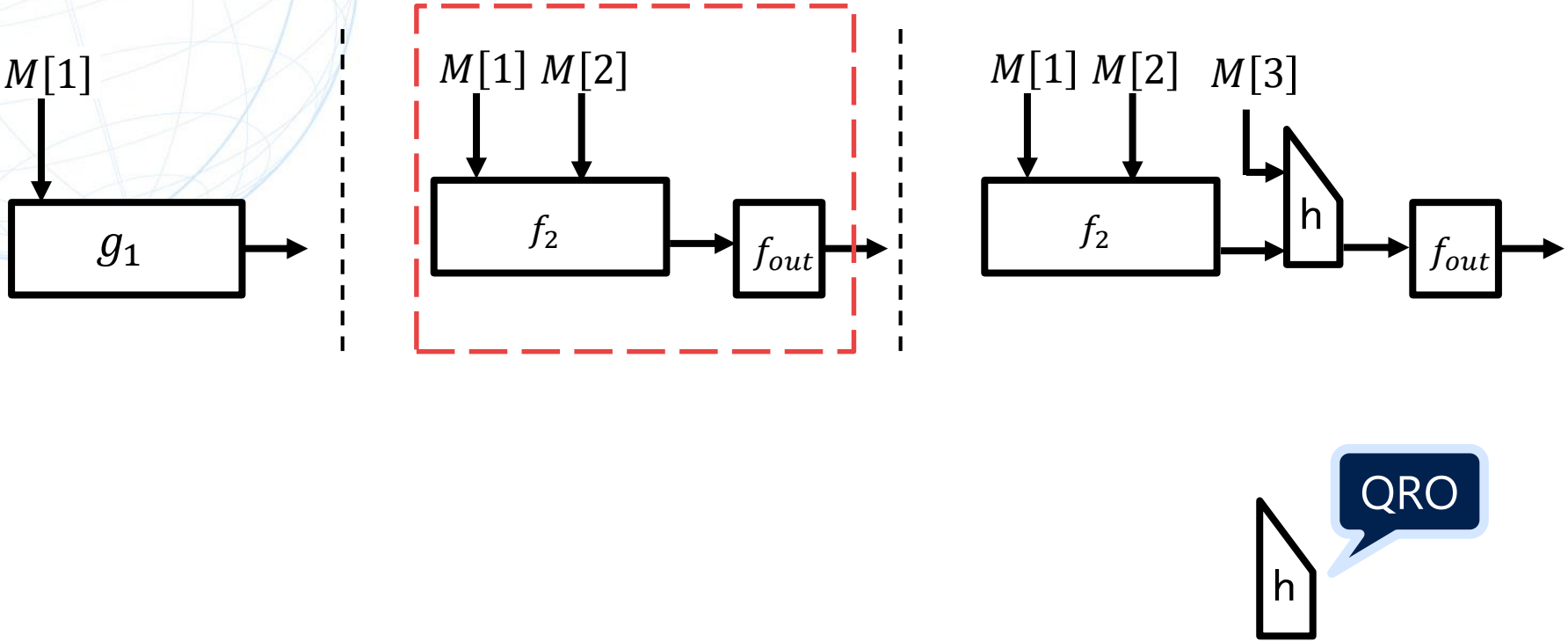
Game G_2



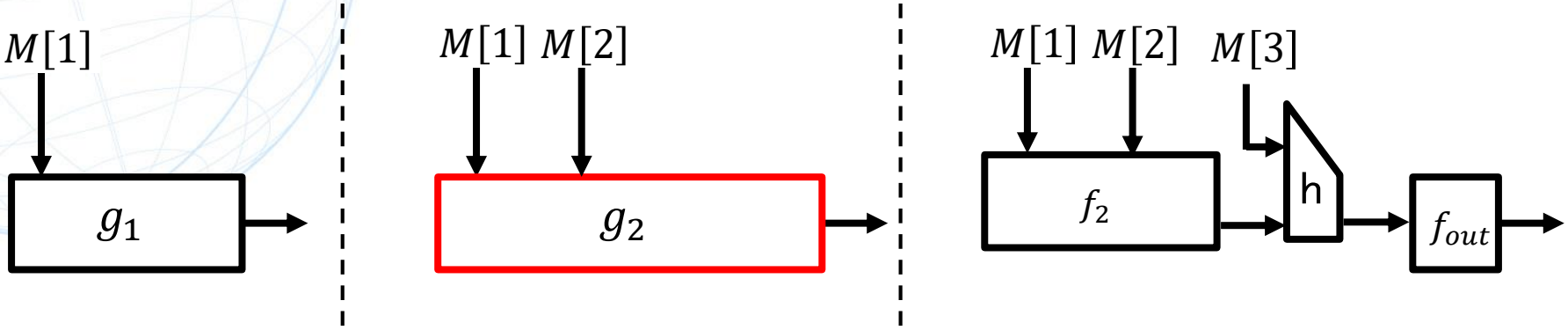
$$h(f_1(\cdot),\cdot) \Rightarrow f_2(\cdot,\cdot)$$



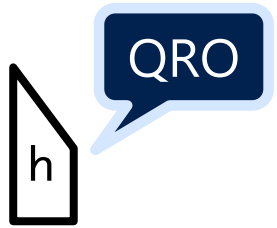
Game G_2



Game G'_2



$$f_{out} \circ f_2 \Rightarrow g_2$$



Game transitions

- From G_i to G'_i: $f_{out} \circ f_i \Rightarrow g_i$
- From G'_i to G_{i+1}: $h(f_i(\cdot),\cdot) \Rightarrow f_{i+1}(\cdot,\cdot)$
- $f_{out}, f_1, f_2, \ldots, g_1, g_2, \ldots$ are independent random functions

- G'_ℓ: The ideal game (random function)

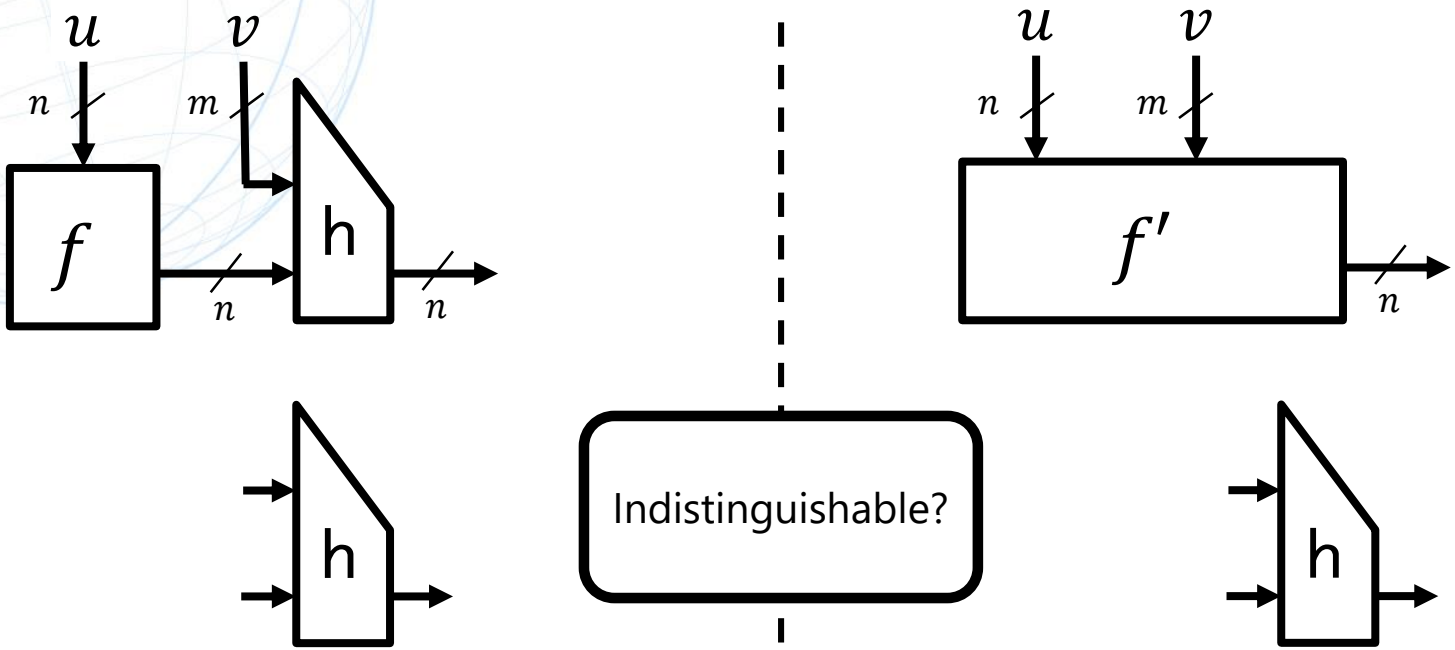
Game transitions

- From G_i to G'_i: $f_{out} \circ f_i \Rightarrow g_i$ The hardest part
- From G'_i to G_{i+1}: $h(f_i(\cdot),\cdot) \Rightarrow f_{i+1}(\cdot,\cdot)$
- $f_{out}, f_1, f_2, \ldots, g_1, g_2, \ldots$ are independent random functions

- G'_ℓ: The ideal game (random function)

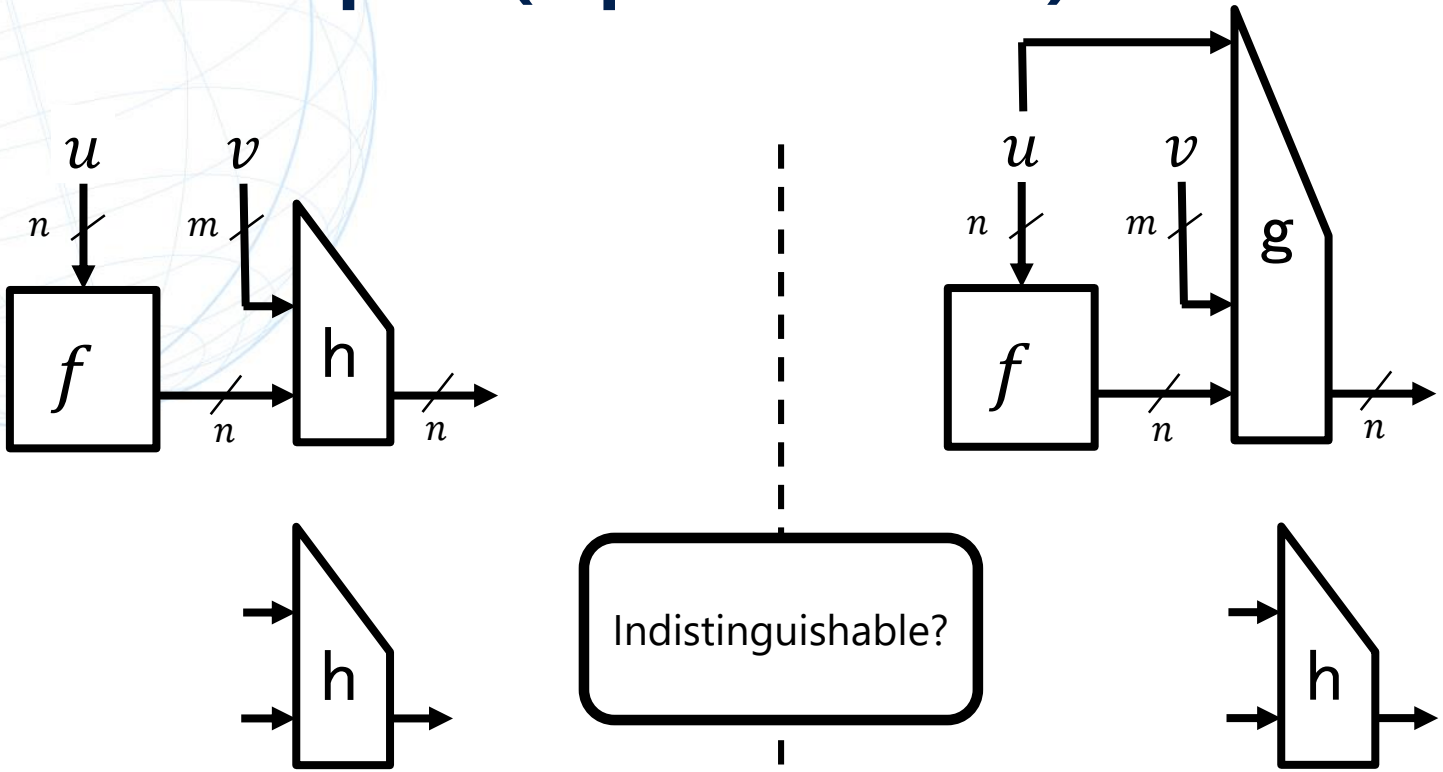
The hardest part: Classical proof

The hardest part (simplified)



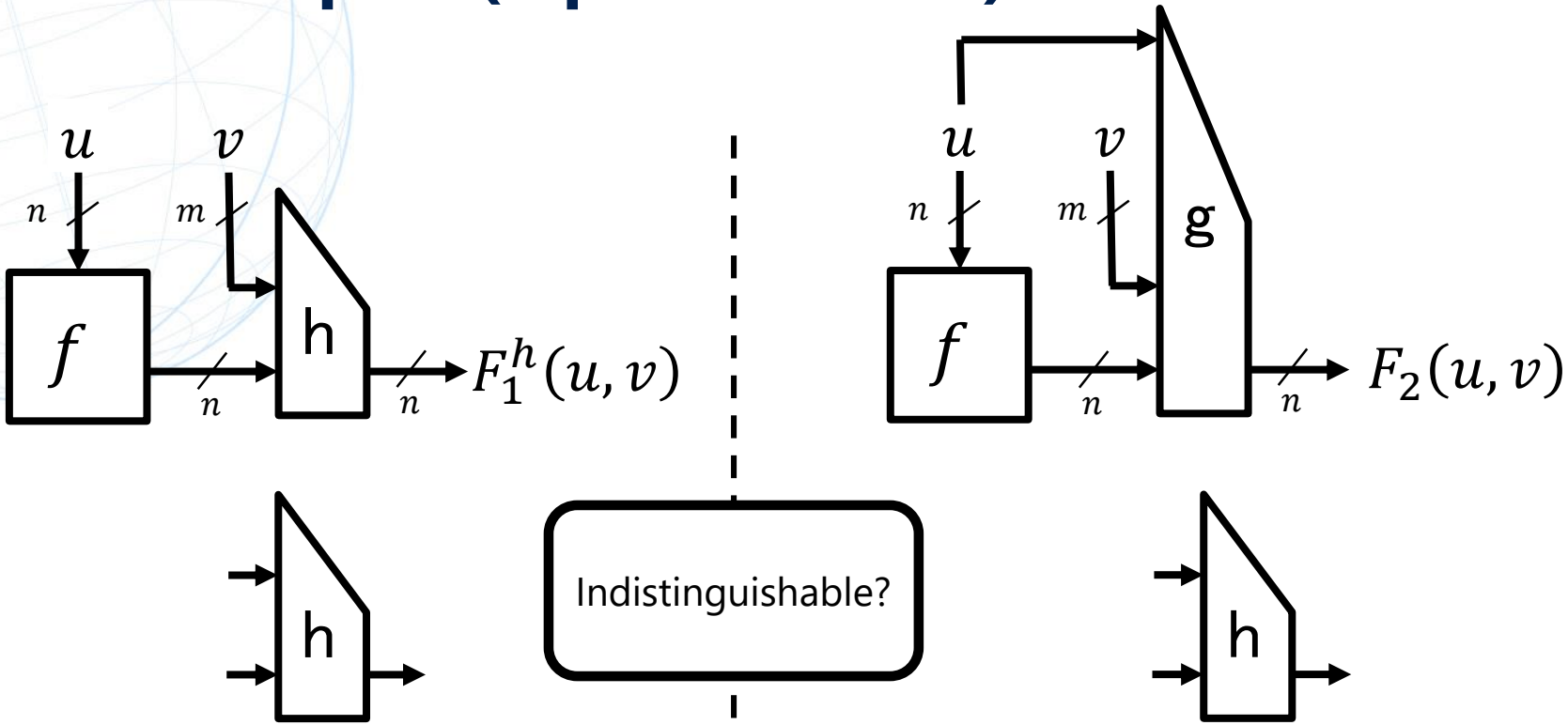
f, f', h are independent random functions

The hardest part (equivalent ver.)



f, g, h are independent random functions

The hardest part (equivalent ver.)



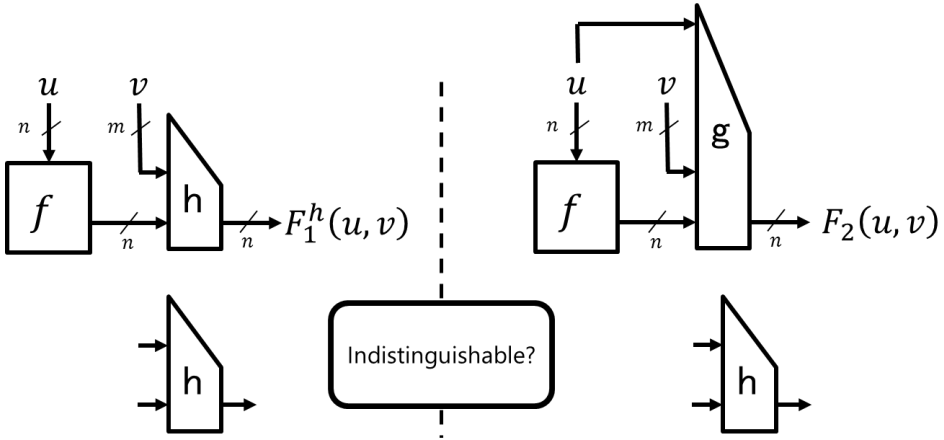
f, g, h are independent random functions

Classical proof idea

- (F_1^h, h) and (F_2, h) are indistinguishable if **coll** and **hit** do not occur

coll : a new output of f collides with a previous input to h

hit : a new direct query to h collides with a previous output of f

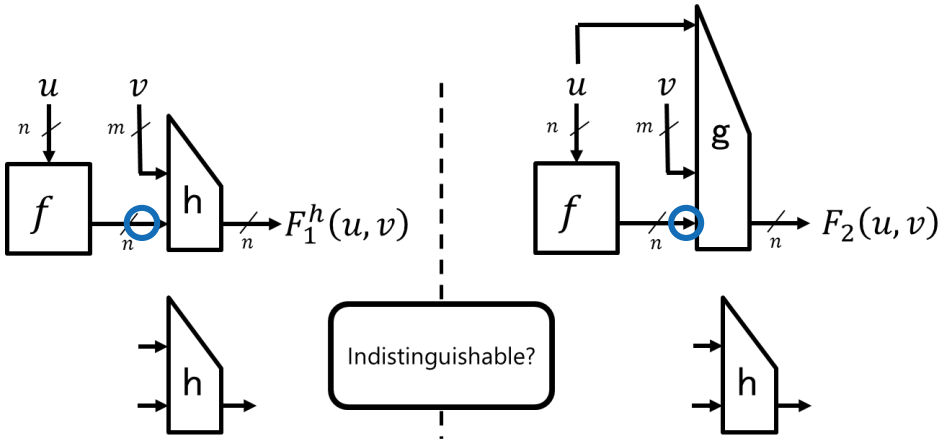


Classical proof idea

- (F_1^h, h) and (F_2, h) are indistinguishable if **coll** and **hit** do not occur

coll : a new output of f collides with a previous input to h

hit : a new direct query to h collides with a previous output of f

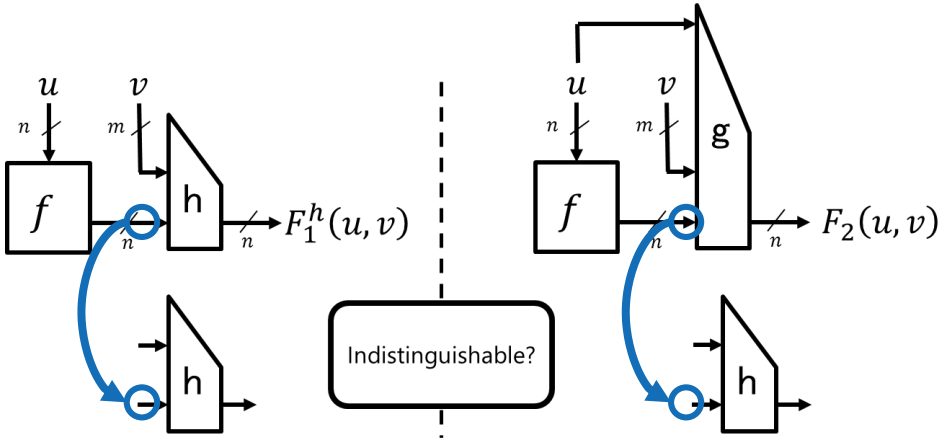


Classical proof idea

- (F_1^h, h) and (F_2, h) are indistinguishable if **coll** and **hit** do not occur

coll : a new output of f collides with a previous input to h

hit : a new direct query to h collides with a previous output of f

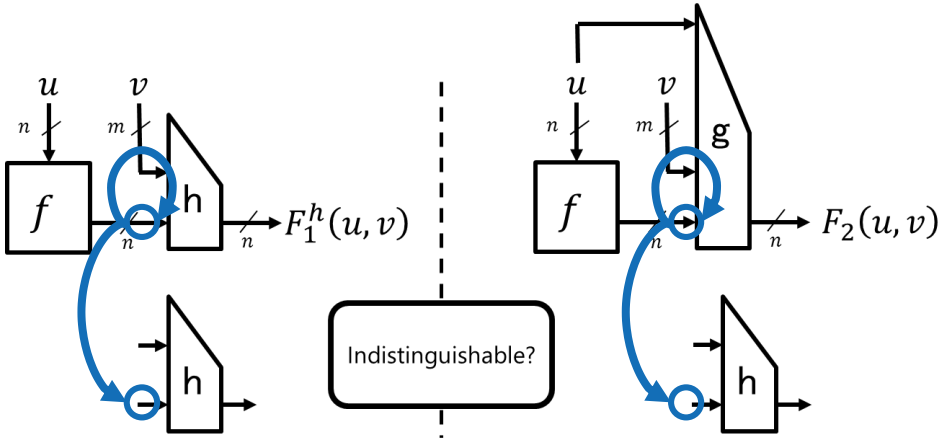


Classical proof idea

- (F_1^h, h) and (F_2, h) are indistinguishable if **coll** and **hit** do not occur

coll : a new output of f collides with a previous input to h

hit : a new direct query to h collides with a previous output of f

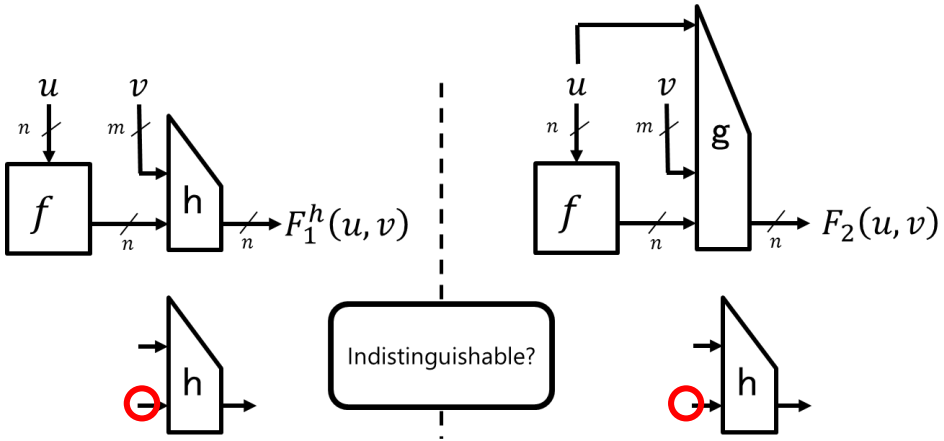


Classical proof idea

- (F_1^h, h) and (F_2, h) are indistinguishable if **coll** and **hit** do not occur

coll : a new output of f collides with a previous input to h

hit : a new direct query to h collides with a previous output of f

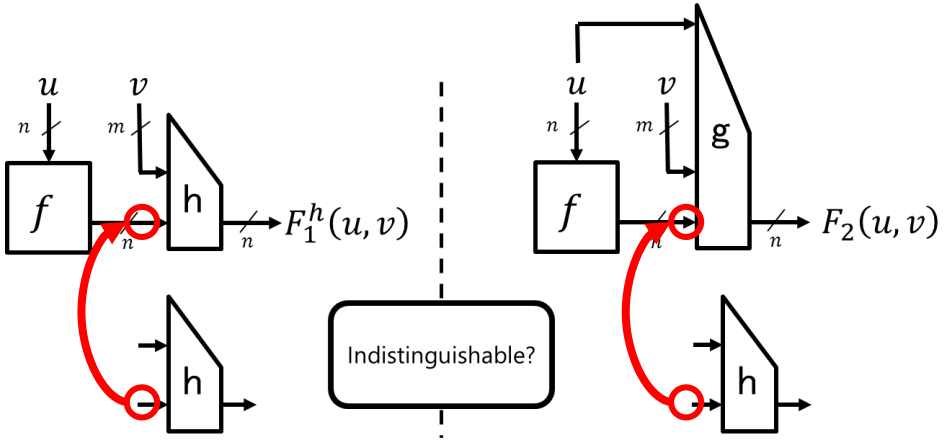


Classical proof idea

- (F_1^h, h) and (F_2, h) are indistinguishable if **coll** and **hit** do not occur

coll : a new output of f collides with a previous input to h

hit : a new direct query to h collides with a previous output of f



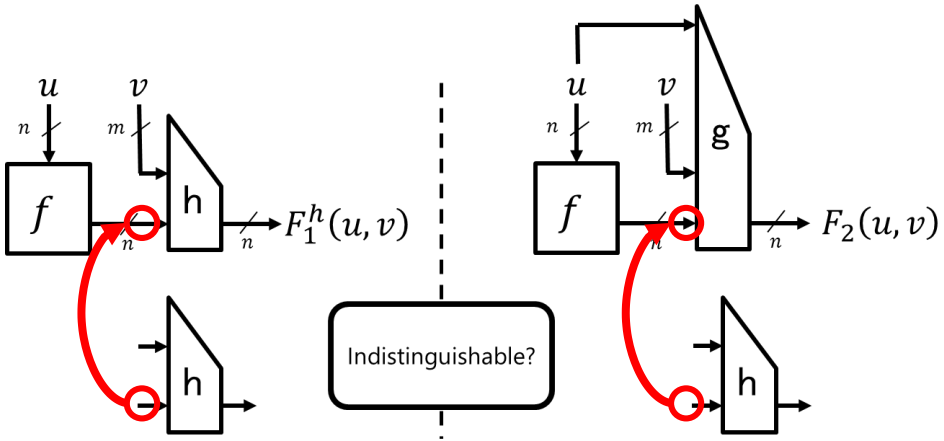
Classical proof idea

- (F_1^h, h) and (F_2, h) are indistinguishable if **coll** and **hit** do not occur

coll : a new output of f collides with a previous input to h

hit : a new direct query to h collides with a previous output of f

$\neg(\text{coll} \vee \text{hit}) \Rightarrow$ outputs of F_1^h and F_2 seem completely random



Classical proof idea

NTT

- (F_1^h, h) and (F_2, h) are indistinguishable if **coll** and **hit** do not occur

coll : a new output of f collides with a previous input to h

hit : a new direct query to h collides with a previous output of f

$\neg(\text{coll} \vee \text{hit}) \Rightarrow$ outputs of F_1^h and F_2 seem completely random

What we have to prove: $\Pr[\text{coll}]$ and $\Pr[\text{hit}]$ are small

- $\Pr[\text{coll}]$ can easily be upper bounded by using the randomness of f
- $\Pr[\text{hit}]$...the randomness of f cannot be directly used and we do not know what the adversary will query

  → some techniques needed (e.g., coefficient-H) to deal with hit

Compressed Oracle Technique

Compressed Oracle Technique

- Classical proofs often rely on the fact that queries/answers can be recorded
- However, in the quantum setting “recording queries” in naïve ways disturb quantum states → lots of classical proofs are invalid in the quantum setting
- Zhandry’s compressed oracle technique enables us to record queries of random functions to some extent [Zha19]

[Zha19] Zhandry, M.: How to record quantum queries, and applications to quantum indifferentiability. CRYPTO 2019.

Compressed Oracle Technique (cont'd)

NTT

- The compressed oracle holds databases of queries/answers in quantum superposition
- Quantum states of the adversary and the oracle look like

$$\sum_{a,b,c,x_1,\dots,y_q} \alpha_{a,b,c,x_1,\dots,y_q} \underbrace{|a,b,c\rangle}_{\text{adversary}} \otimes \underbrace{\left|(x_1,y_1),\dots,(x_q,y_q)\right\rangle}_{\text{Oracle's database}}$$

- It behaves like the classical lazy-sampling (to some extent)
  - A fresh query $x \rightarrow$ uniform superposition of y is added: $\sum_y \frac{1}{\sqrt{2^n}} |y\rangle$
  - Sometimes records are forgotten or overwritten

Proof in the Quantum setting

The first proof idea in the quantum setting

The first idea: classical proof idea + compressed oracle

The joint quantum states of

- Adversary A and the oracles $\left(F_1^h, h\right)$, and
- Adversary A and the oracles (F_2, h)

will be indistinguishable as long as coll and hit do not happen in the databases & for A's query

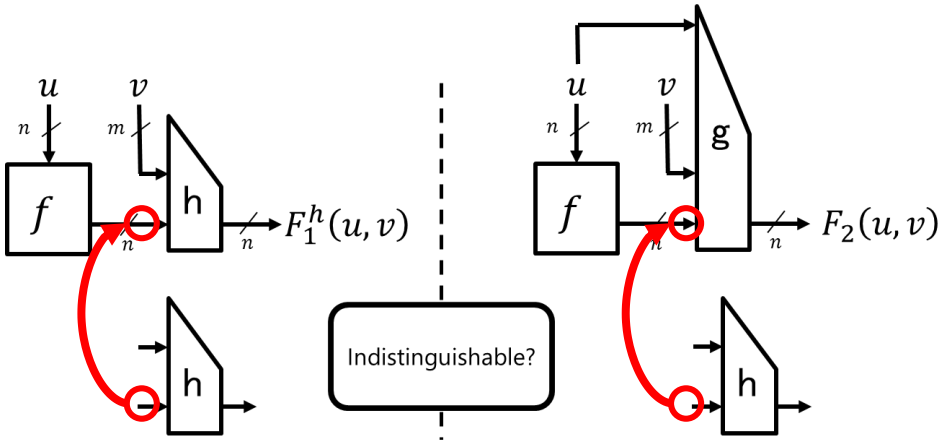
What we have to prove: "Pr[coll]" and "Pr[hit]" are small

- Pr[coll] can easily be upper bounded by using the randomness of f
- Pr[hit]… the randomness of f cannot be directly used and we do not know what the adversary will query and classical proof techniques cannot be used

  → **new technique** needed to deal with hit

How to deal with hit in the quantum setting

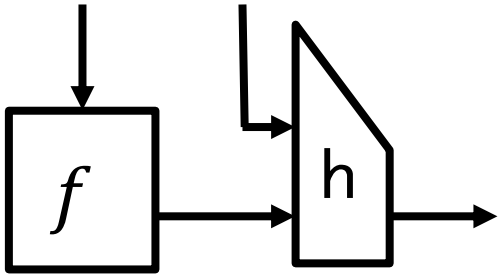
hit…the event that the adversary A succeeds to guess an output of f



How to deal with hit in the quantum setting

NTT

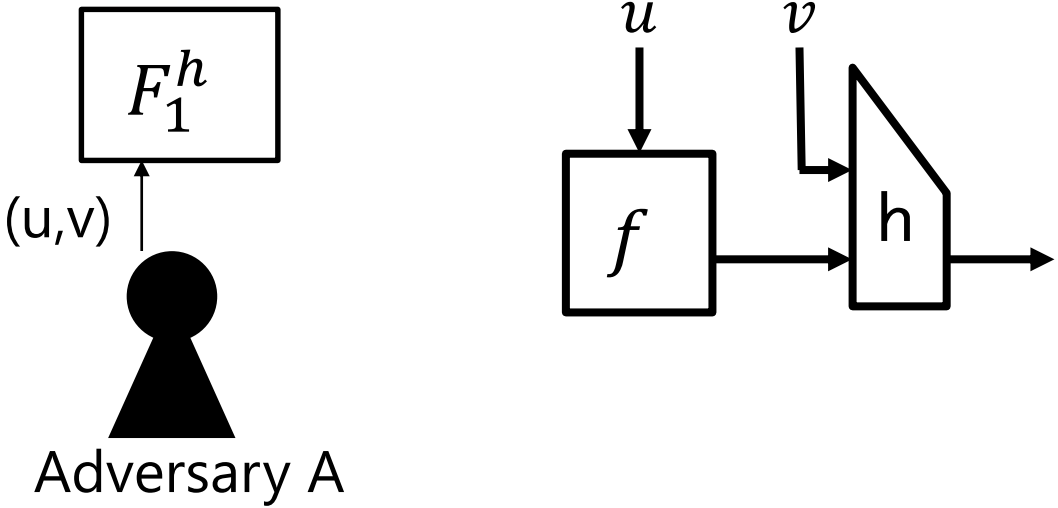
F_1^h



Adversary A

How to deal with hit in the quantum setting

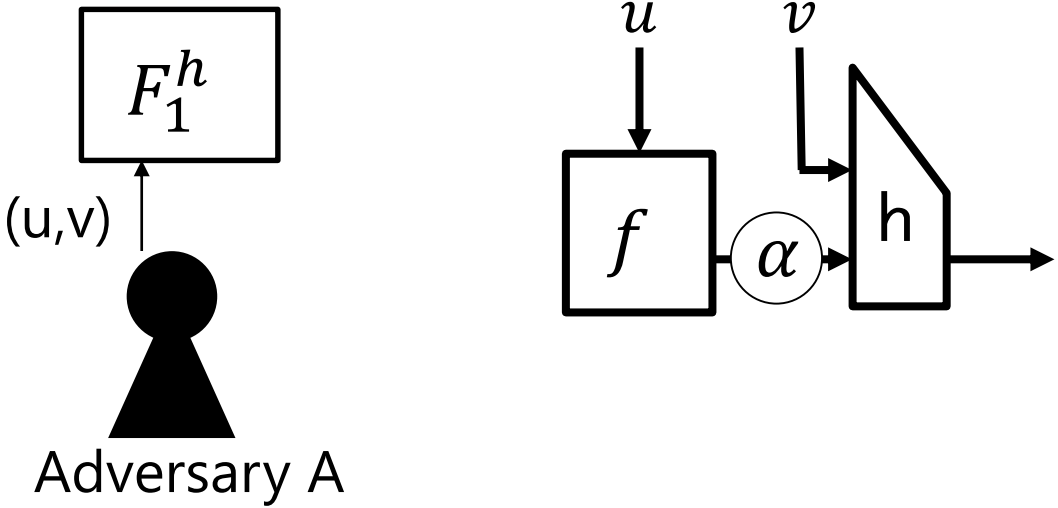
Suppose the adversary A makes a fresh query (u,v) to F_1^h



How to deal with hit in the quantum setting

Suppose the adversary A makes a fresh query (u,v) to F_1^h

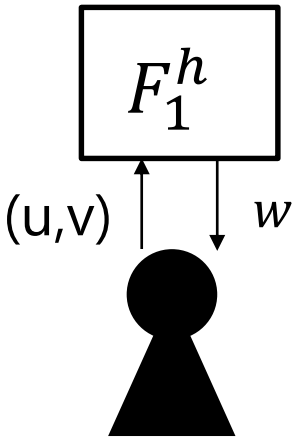
1. $\alpha = f(u)$ is sampled

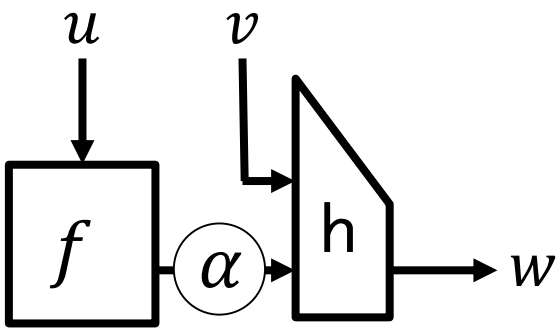


How to deal with hit in the quantum setting

Suppose the adversary A makes a fresh query (u,v) to F_1^h

1. $\alpha = f(u)$ is sampled
2. $w = h(\alpha, v)$ is sampled and returned to A





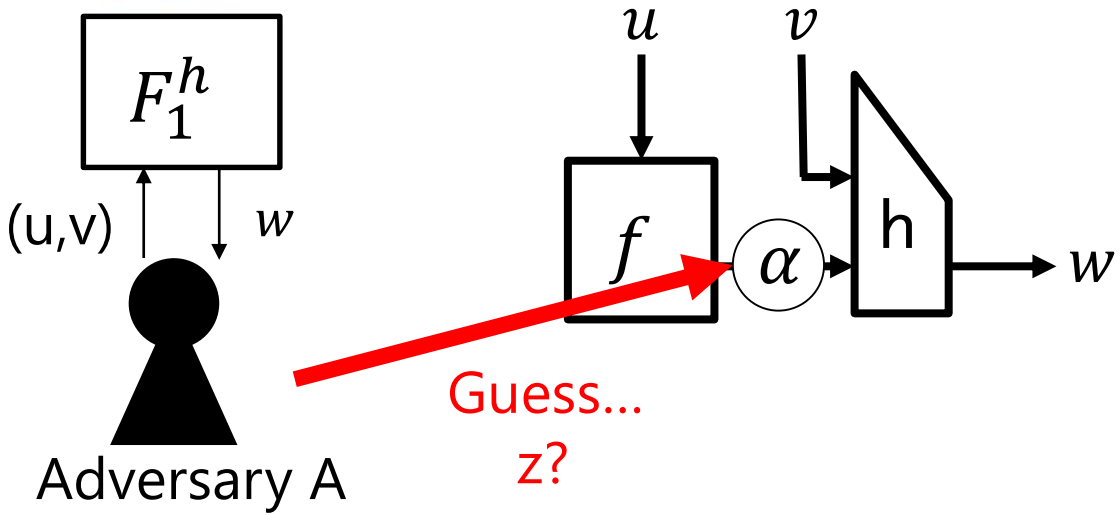
Database for f Database for h

$$|(u,\alpha)\rangle \otimes |(\alpha||v,w)\rangle$$

How to deal with hit in the quantum setting

Suppose the adversary A makes a fresh query (u,v) to F_1^h

1. $\alpha = f(u)$ is sampled
2. $w = h(\alpha, v)$ is sampled and returned to A



Database for f

Database for h

$$|(u,\alpha)\rangle \otimes |(\alpha||v,w)\rangle$$

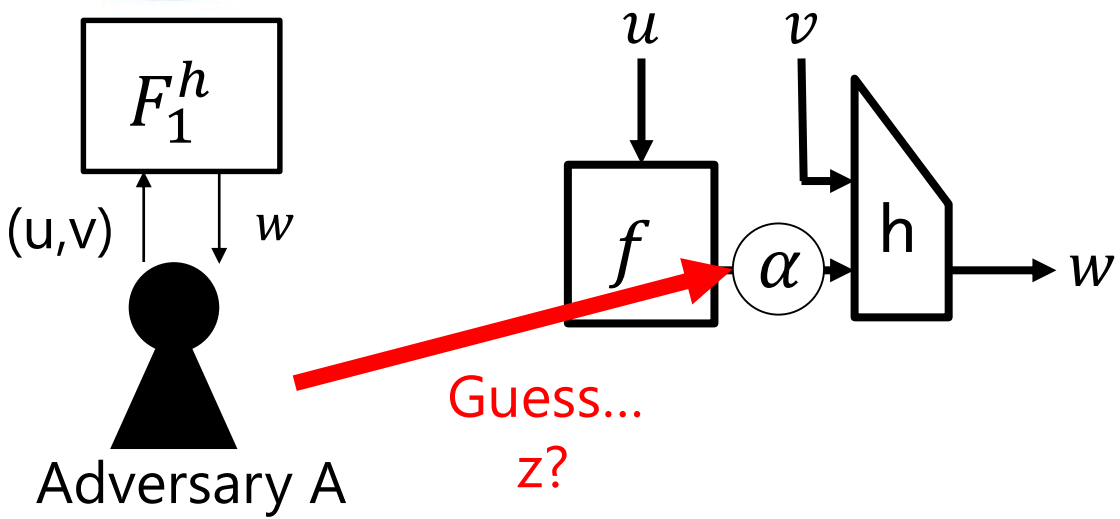
How to deal with hit in the quantum setting

NTT

Suppose the adversary A makes a fresh query (u,v) to F_1^h

1. $\alpha = f(u)$ is sampled
2. $w = h(\alpha, v)$ is sampled and returned to A

The adversary knows w but does not know anything about $\alpha = f(u)$



Database for f Database for h

$$|(u,\alpha)\rangle \otimes |(\alpha||v,w)\rangle$$

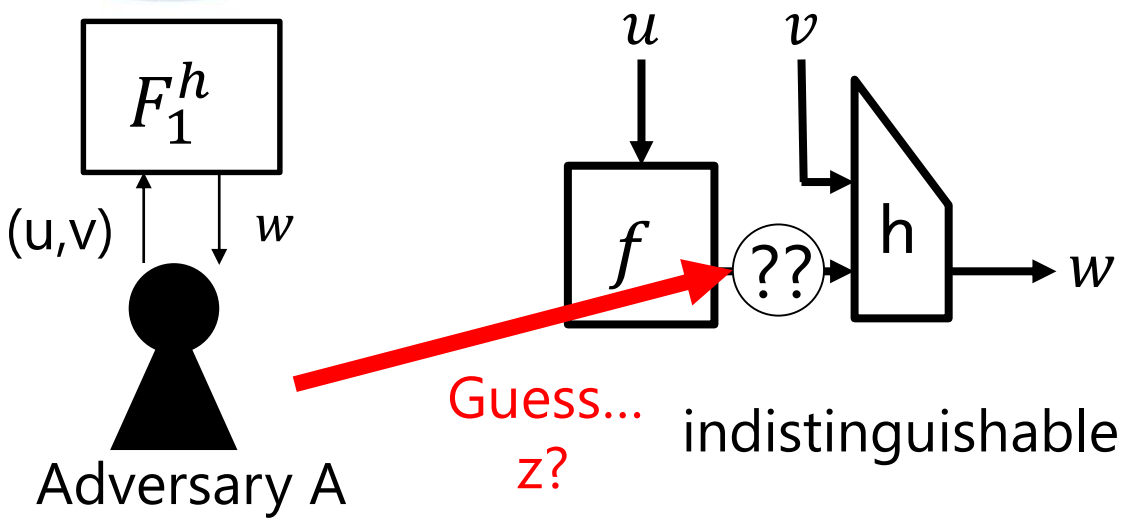
How to deal with hit in the quantum setting

NTT

Suppose the adversary A makes a fresh query (u,v) to F_1^h

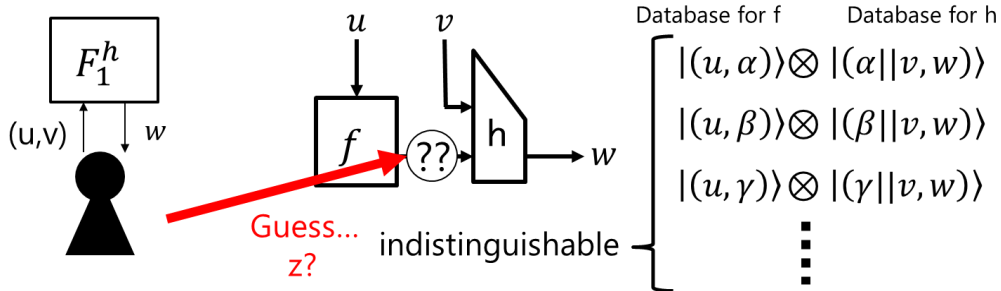
1. $\alpha = f(u)$ is sampled
2. $w = h(\alpha, v)$ is sampled and returned to A

The adversary knows w but does not know anything about $\alpha = f(u)$



How to deal with hit in the quantum setting

We say databases are *equivalent* if they are indistinguishable from the adversary



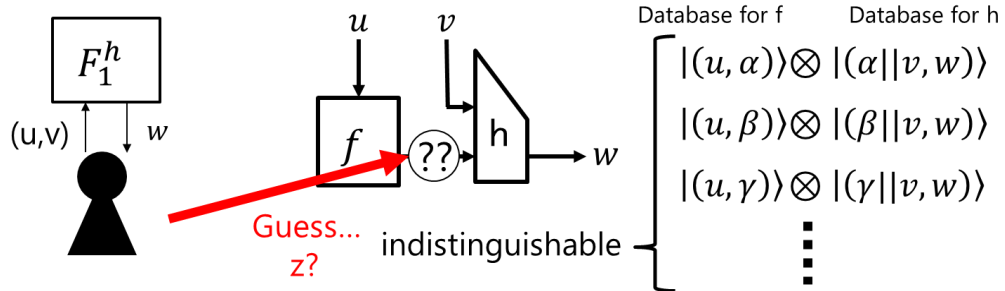
How to deal with hit in the quantum setting

We say databases are *equivalent* if they are indistinguishable from the adversary

For arbitrary z (the adversary's guess of an output of f),

$$\Pr[z \text{ is indeed an output of } f] = \frac{\#\{\text{equivalent DBs s.t. } f(u)=z \text{ for some } u\}}{\#\{\text{equivalent DBs}\}}$$

$$\approx O\left(\frac{i}{2^n}\right) \quad \text{(after making i queries)}$$



How to deal with hit in the quantum setting

NTT

We say databases are *equivalent* if they are indistinguishable from the adversary

For arbitrary z (the adversary's guess of an output of f),

$$\Pr[z \text{ is indeed an output of } f] = \frac{\#\{\text{equivalent DBs s.t. } f(u)=z \text{ for some } u\}}{\#\{\text{equivalent DBs}\}}$$

$$\approx O\left(\frac{i}{2^n}\right) \quad \text{(after making i queries)}$$

$$\Pr[\text{hit occurs at the } i\text{-th query}] \approx O\left(\frac{i}{2^n}\right)$$

How to deal with hit in the quantum setting

NTT

$|\phi_{\mathrm{hit}}\rangle$:=The vector that corresponds to "**hit** happens after q queries"

$\left|\phi_{\mathrm{hit}}^{(i)}\right\rangle$:=The vector that corresponds to "**hit** happens at the i-th query"

$$\||\phi_{\mathrm{hit}}\rangle\| \leq \sum_{1\leq i\leq q}\left\|\left|\phi_{\mathrm{hit}}^{(i)}\right\rangle\right\| \leq \sum_{1\leq i\leq q}\sqrt{\Pr[\mathrm{hit}\ \text{occurs at the}\ i\text{- th query}]} \leq q\cdot O\left(\sqrt{\frac{q}{2^n}}\right) \leq O\left(\sqrt{\frac{q^3}{2^n}}\right)$$

The norm of $|\phi_{\mathrm{hit}}\rangle$ is small as long as $q \ll 2^{\frac{n}{3}}$

$$\left(\||\phi_{\mathrm{hit}}\rangle\| = \sqrt{\Pr[\mathrm{hit}]}\right)$$

Remarks

More precisely, we show

- $|\phi_{\mathrm{hit}}\rangle$ is small as long as $q \ll 2^{\frac{n}{3}}$
- $|\phi_{\mathrm{coll}}\rangle$ is small as long as $q \ll 2^{\frac{n}{3}}$
- The quantum states are “equal” (there is an isometry) if $\neg(\mathrm{hit} \vee \mathrm{coll})$
- Quantum coefficients for equivalent databases are exactly equal

by tracing coefficient of each vector in detail

Game transitions

- From G_i to G'_i: $f_{out} \circ f_i \Rightarrow g_i$
- From G'_i to G_{i+1}: $h(f_i(\cdot),\cdot) \Rightarrow f_{i+1}(\cdot,\cdot)$
- $f_{out}, f_1, f_2, \ldots, g_1, g_2, \ldots$ are independent random functions

- G_0: The real game (NMAC)
- G'_ℓ: The ideal game (random function)

Game transitions

- From G_i to G'_i: $f_{out} \circ f_i \Rightarrow g_i$
- From G'_i to G_{i+1}: $h(f_i(\cdot),\cdot) \Rightarrow f_{i+1}(\cdot,\cdot)$ indistinguishable up to $2^{n/3}$ queries
- $f_{out}, f_1, f_2, \ldots, g_1, g_2, \ldots$ are independent random functions

- G_0: The real game (NMAC)
- G'_ℓ: The ideal game (random function)

Game transitions

- From G_i to G'_i: $f_{out} \circ f_i \Rightarrow g_i$ indistinguishable up to $2^{n/3}$ queries
- From G'_i to G_{i+1}: $h(f_i(\cdot),\cdot) \Rightarrow f_{i+1}(\cdot,\cdot)$ indistinguishable up to $2^{n/3}$ queries
- $f_{out}, f_1, f_2, \ldots, g_1, g_2, \ldots$ are independent random functions

- G_0: The real game (NMAC)
- G'_ℓ: The ideal game (random function)

Game transitions

- From G_i to G'_i: $f_{out} \circ f_i \Rightarrow g_i$ indistinguishable up to $2^{n/3}$ queries
- From G'_i to G_{i+1}: $h(f_i(\cdot),\cdot) \Rightarrow f_{i+1}(\cdot,\cdot)$ indistinguishable up to $2^{n/3}$ queries
- $f_{out}, f_1, f_2, \ldots, g_1, g_2, \ldots$ are independent random functions

- G_0: The real game (NMAC)
- G'_ℓ: The ideal game (random function)

} indistinguishable up to $2^{n/3}$ queries (when ℓ is small)

Game transitions

- From G_i to G'_i: $f_{out} \circ f_i \Rightarrow g_i$ indistinguishable up to $2^{n/3}$ queries
- From G'_i to G_{i+1}: $h(f_i(\cdot),\cdot) \Rightarrow f_{i+1}(\cdot,\cdot)$ indistinguishable up to $2^{n/3}$ queries
- $f_{out}, f_1, f_2, \ldots, g_1, g_2, \ldots$ are independent random functions

- G_0: The real game (NMAC)
- G'_ℓ: The ideal game (random function)

} indistinguishable up to $2^{n/3}$ queries (when ℓ is small)

Proof for HMAC is almost the same

Summary

Summary

- Tight quantum security bound in the QROM (h : QRO)
- HMAC/NMAC are indistinguishable from a RF against quantum query attacks as long as $(q_h + Q) \cdot \ell^{\frac{3}{5}} \geq 2^{n/3}$
  - q_h: max. num. of quantum queries to h, Q: max. num. of quantum queries to the keyed oracle of HMAC/NMAC, ℓ: maximum message length
- Tight when ℓ is not exponentially large
- Compressed oracle technique
- Hardest part: Proving the adversary's guess success prob. is low
  - We introduced "equivalent databases" to capture the uncertainty of outputs of a random function that the adversary cannot observe directly

Thank you!