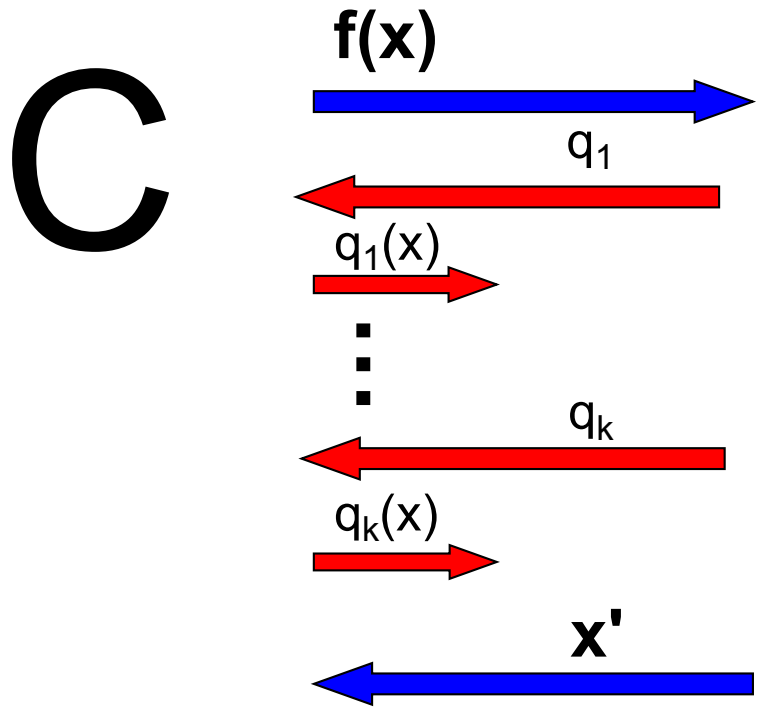
UNPROVABLE SECURITY

of Leakage Resilient Cryptography beyond the Information-theoretic Limit

Rafael Pass
Cornell Tech

Leakage-Resilient Cryptography

- Study of cryptography in the **presence of leakage** [ISW03,MR04]
- Our focus is on the **bounded-leakage model** [Maurer94,AGV09]
  - the attacker gets some arbitrary poly-time leakage on the secret of **some bounded size l(n) << n**
- Consider the simplest task: **l(n)-leakage resilient OWF**



WINS if:
(1) f(x') = f(x)
(2) $|q_1(x)|+..|q_k(x)| < l(|x|)$

i.e., inverts using **l-bounded leakage**

Leakage-Resilient OWF

Trivial leakage resilience:

- Any OWF/OWP is a **O(log n)-leakage resilient**
- Any subexp-secure OWF/OWP is **O(n^{ε})-leakage resilient** for some eps > 0

Thm [KV'09,ADW'09]: Any CRH h is a **n/2-leakage resilient OWF**

Proof idea:

- Given h(x), and leak(x), **x still has lots of entropy left.**
- If you can find find a pre-image, its unlikely to be x.
- So any attacker that succeeds in inverting h given the leakage, can be used to find collision in h.

Leakage-Resilient OWF

Trivial leakage resilience:

- Any OWF/OWP is a **O(log n)-leakage resilient**
- Any subexp-secure OWF/OWP is **$O(n^{\varepsilon})$-leakage resilient** for some eps > 0

Thm [KV'09,ADW'09]: Any CRH h is a **n/2-leakage resilient OWF**

Key point: secret needs to have (actual) entropy left after leakage
This principle is used in all proofs for "non-trivial" leakage resilience

Focus of this paper:

Can we go beyond this "information-theoretic" barrier?

(can we prove leakage-resilience also of "computationally hidden" secrets)

Open Questions

1. Can we base $O(n^{\varepsilon})$-leakage resilient OWP on polynomial hardness assumption?

2. Can we base $O(n^{\varepsilon})$-leakage resilient OWF with $<< 2^{n^{\wedge}\varepsilon}$ pre-images on polynomial hardness assumption?

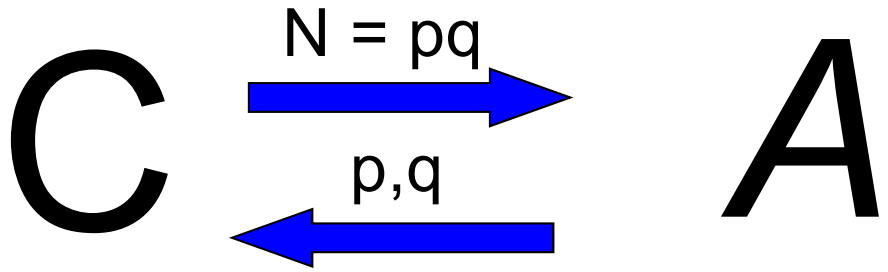
Main Result

Answer is **NO** with respect to any **black-box security reductions** from **O(1)-round intractability assumptions.**

Any security reduction R **itself** must constitute an attack on the intractability assumption

Intractability Assumptions

Following [Naor'03,DOP'05,HH'09,..GW'11,P'11], we model an **intractability assumption** as an interaction between a Challenger C and an attacker A.

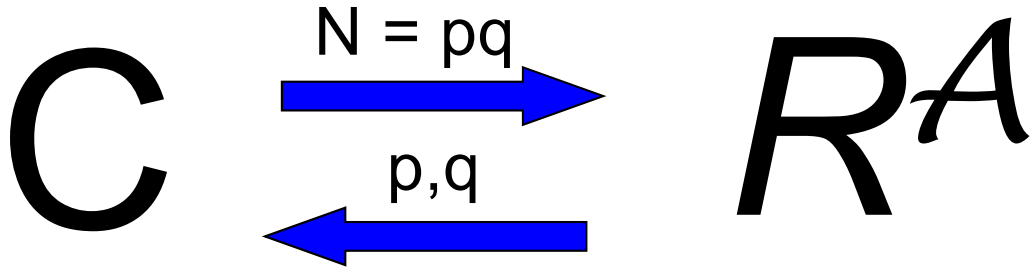


Intractability assumption (C,t) :
"no PPT can make C output 1 w.p significantly above t"

Bounded-round assumption [P'11]: C communicates in r(n)-round

- 2-round: Factoring, f is a OWF, G is a PRG, DDH, …
- O(1)-round: Enc is semantically secure, (P,V) is WH,
- l(n)-round: **f is l-leakage-resilient OWF**

Black-box Reductions



R^A breaks C whenever A breaks security of scheme

Reduction R may rewind and restart A.

Main Theorem

- Assume the existence of CRH.
- Let f be a function with **$2^{n^{\wedge}\varepsilon}$-bounded pre-image set sizes.**
- Let (C,t) be a **O(1)-round intractability assumption**

If there exists a PPT black-box reduction R for basing **$O(n^{\varepsilon})$-leakage resilient one-wayness of f** on the hardness of (C,t), then there exists a PPT attacker B that breaks (C,t)

Remark:
Our main theorem applies also to **leakage-resilience of NP-search problems**

Related Work

[Wichs'13] shows a very related black-box impossibility results for OWP:

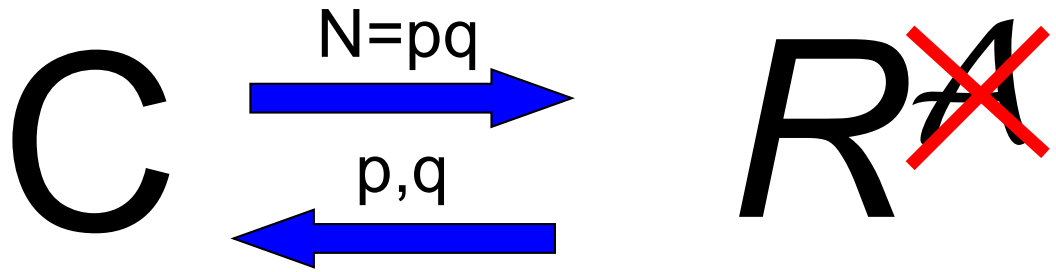
- But his result considers only **restricted classes of reduction**, which treat the **leakage query as a black-box**
- Non-black-box usage of leakage queries have proven useful in the related context of KDM-security [BHH10, MPS'16]

[Aggarwal-Maurer'11] also study l-leakage-resilience of NP search problems.

- Do no present lower-bounds;
- But relate this problem to other computational problems (e.g., optimal algorithms without leakage)

Large body of work on "meta-reductions" [BV'99]

Meta-reductions [BV'99,(Bra'79)]

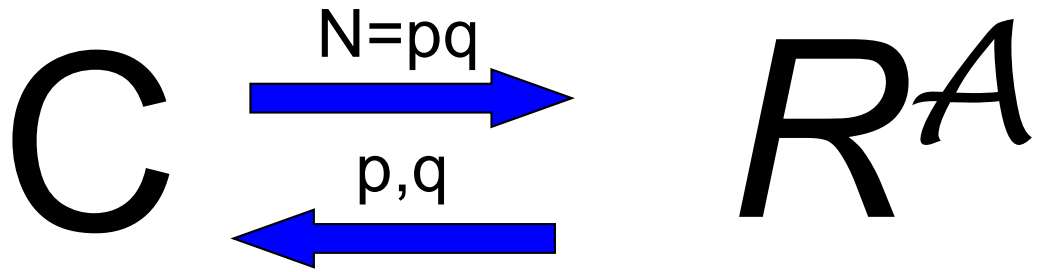


1. Design a **particular attacker A** that breaks $\mathbf{n}^{\varepsilon}$-leakage resilience
2. Show how to **emulate** attacker in poly-time.

Today:

- Restrict attention to **OWP**

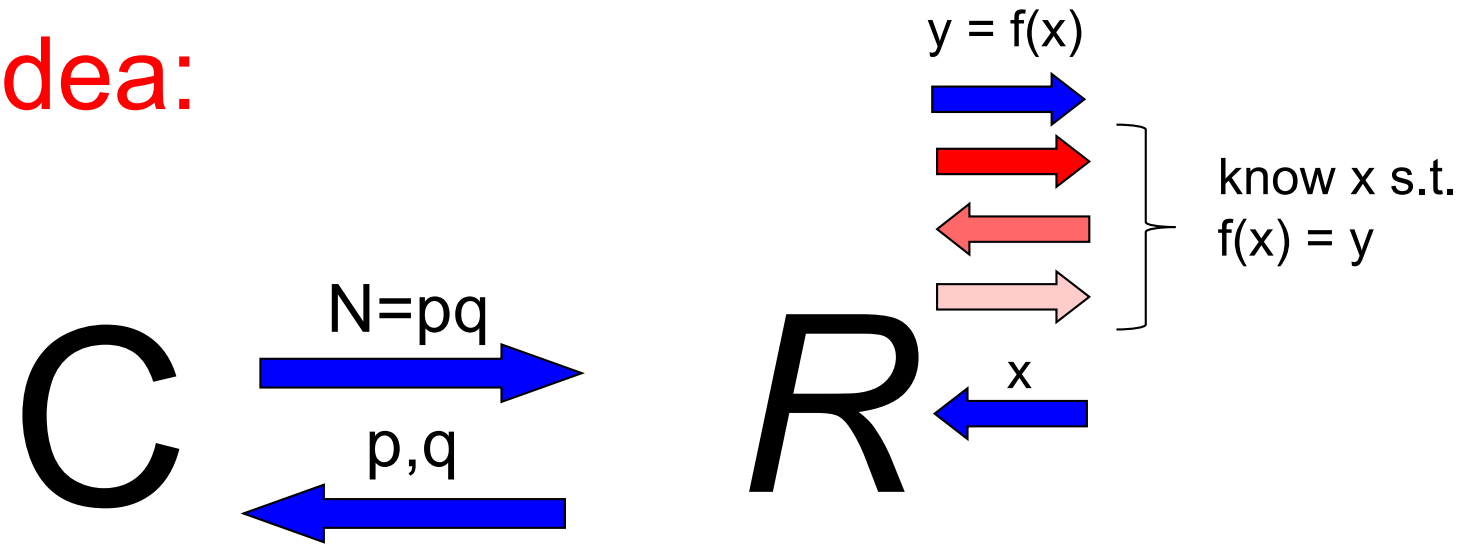
Proof Idea:



Consider **attacker A**, that given y = f(x):

- ask to hear a **succinct argument of knowledge** of the statement "exists x s.t. f(x) = y" (exists based on CRH)
- If the argument is accepting, it recovers any x s.t. f(x) = y, and returns it.

Proof Idea:



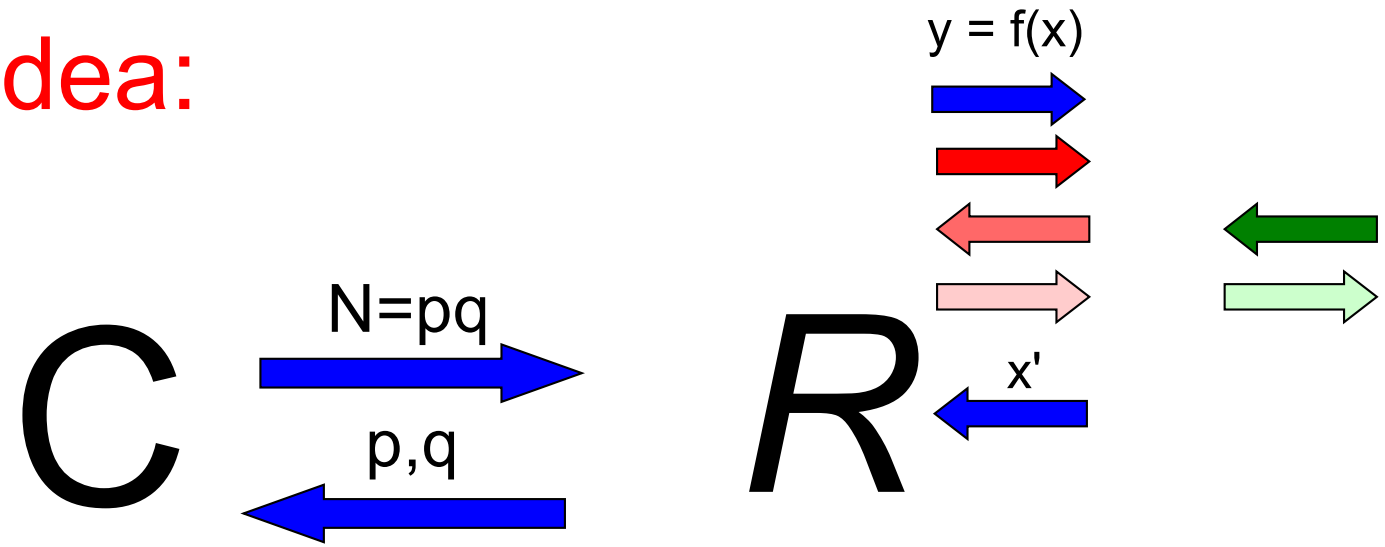
Consider **attacker A**, that given y = f(x):

- ask to hear a **succinct argument of knowledge** of the statement "exists x s.t. f(x) = y" (exists based on CRH)
- If the argument is accepting, it recovers any x s.t. f(x) = y, and returns it.

Remarks:

- Similar method used in [NVZ'14,OPV15] in the context of leakage-resilient ZK and secure computation.

Proof Idea:



Consider **attacker A**, that given y:

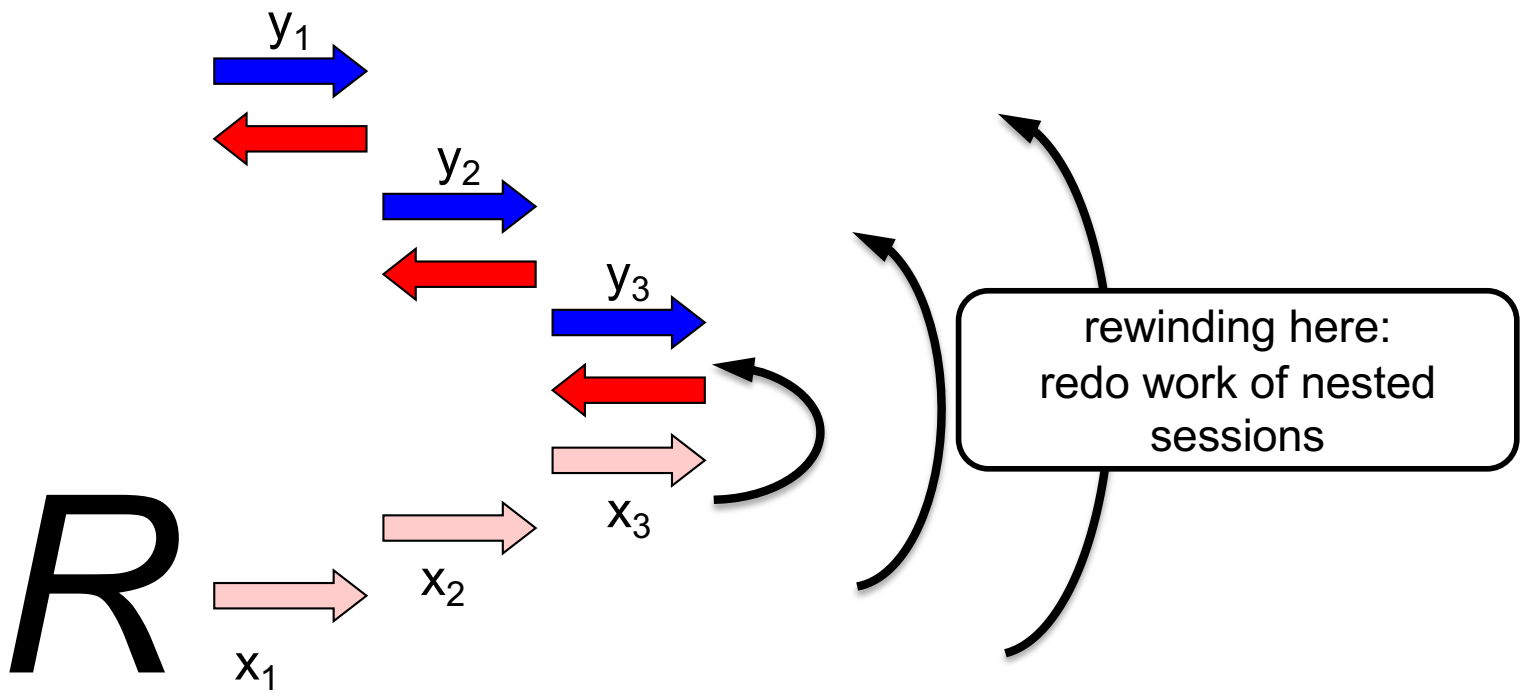
- ask to hear a **succinct argument of knowledge** of the statement "exists x s.t. f(x) = y" (exists based on CRH)
- If the argument is accepting, it recovers any x s.t. f(x) = y, and returns it.

Emulate A:

- **Rewind the AOK** to extract out some x' s.t. f(x') = y
- Return x'

Since f is a OWP, x' must be equal to x returned by A
(fails f is a general OWF, or even if there are more than 1 pre-image)

General Reductions: Problem



Problem: R might nest its oracle calls.

"naïve extraction" requires exponential time
(c.f., Concurrent ZK [DNS'99])

Solution: **Rely on techniques developed in [P11]**

1. use a special form of AOK ("special-soundness"); will no longer be succinct, but will still have a **laconic prover**
2. require R to provide **many**, **$O(n^\epsilon)$**, sequential proofs, then we can find (recursively) find one proof where nesting depth is "small".

Theorem

- Assume the existence of CRH.
- Let f be a OWP
- Let (C,t) be a **O(1)-round intractability assumption**

If there exists a PPT black-box reduction R for basing **$O(n^{\epsilon})$-leakage resilient one-wayness of f** on the hardness of (C,t), then there exists a PPT attacker B that breaks (C,t)

Additional ideas need to extend to OWF with bounded number of pre-images

Main Theorem

- Assume the existence of CRH.
- Let f be a function with **$2^{n^{\wedge}\varepsilon}$-bounded pre-image set sizes.**
- Let (C,t) be a **O(1)-round intractability assumption**

If there exists a PPT black-box reduction R for basing **$O(n^{\varepsilon})$-leakage resilient one-wayness of f** on the hardness of (C,t), then there exists a PPT attacker B that breaks (C,t)

Thank You