Finding Random Collisions for Random Degree-2 Functions

Xinyu Mao (University of Waterloo)

@TCC 2026

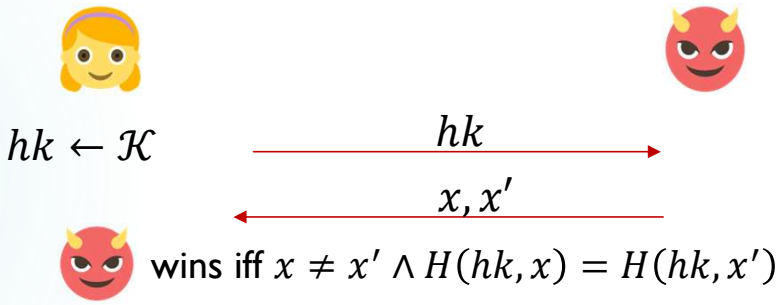
Collision-Resistant Hash Function (CRH)

The first input is viewed as key

CRH

- A keyed hash function $H: \mathcal{K} \times \{0,1\}^m \to \{0,1\}^\ell$
- Shrinking: $\ell < m$.
- **Collision resistance:** $\forall$ PPT A

$$\Pr_{hk \leftarrow \mathcal{K}}[A(hk) = (x, x') \text{ s.t. } x \neq x' \wedge H(hk, x) = H(hk, x')] \text{ is negligible in } \lambda.$$



Applications: digital signature, vector commitments, succinct arguments…

Random Degree-2 Functions

Let p be a prime. The hash function family $\mathcal{Q}_{p,M,N}$ consists of functions

$$h\colon \mathbb{F}_p^N \to \mathbb{F}_p^M, \qquad h(x)_i = \gamma_i + \ell_i^\top x + \sum_{1 \le a \le b \le N} q_{i,a,b} x_a x_b$$

where $\gamma_i \in \mathbb{F}_p$, $\ell_i \in \mathbb{F}_p^N$, $q_{i,a,b} \in \mathbb{F}_p$ for all i, a, b.

- Hash functions are widely used, and we want efficient, low-complexity ones.
- Low-degree functions are friendly to applications like MPC, FHE, etc.
  - Degree-one functions are linear affine functions, which are too simple…

Random Degree-2 Functions Are Not Collision-Resistant [Ding&Yang 08]

$$h\colon \mathbb{F}_p^N \to \mathbb{F}_p^M, \qquad h(x)_i = \gamma_i + \ell_i^\top x + \sum_{1 \le a \le b \le N} q_{i,a,b} x_a x_b$$

Equivalently, $h(x)_i = \gamma_i + \ell_i^\top x + x^\top Q_i x$ where $Q_i = \begin{bmatrix} q_{i,1,1}, & 0 & \cdots & 0 \\ \vdots & & \ddots & \vdots \\ & & & 0 \\ q_{i,1,N}, & & \cdots & q_{i,N,N} \end{bmatrix}$

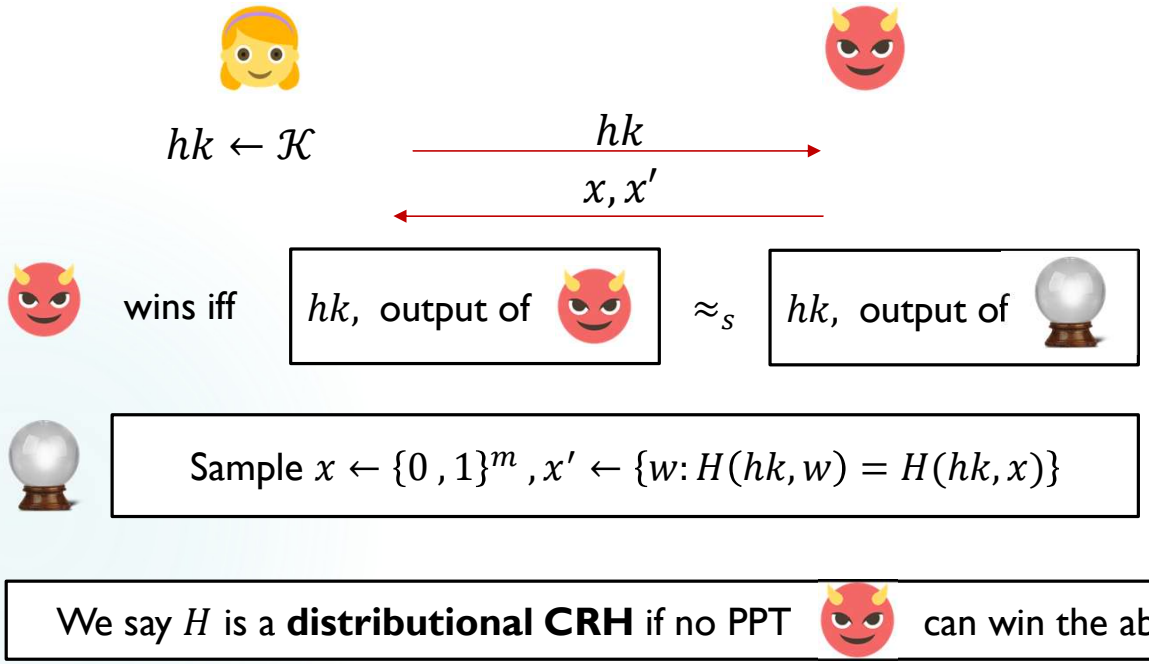
Differential attack

1. Choose a non-zero difference vector $\Delta \in \mathbb{F}_p^N$.
2. Solve the **linear system** $h(x) = h(x + \Delta)$ to get a solution x^*.
3. Output $(x^*, x^* + \Delta)$.

$$h(x)_i = h(x+\Delta)_i \quad \Leftrightarrow \quad \Delta^\top \left(Q_i + Q_i^\top\right) \cdot x + \Delta^\top Q_i \Delta + \ell_i^\top \Delta = 0$$

Distributional Collision Resistance

$H: \mathcal{K} \times \{0,1\}^m \to \{0,1\}^\ell \quad (\ell < m)$



We say H is a **distributional CRH** if no PPT can win the above game.

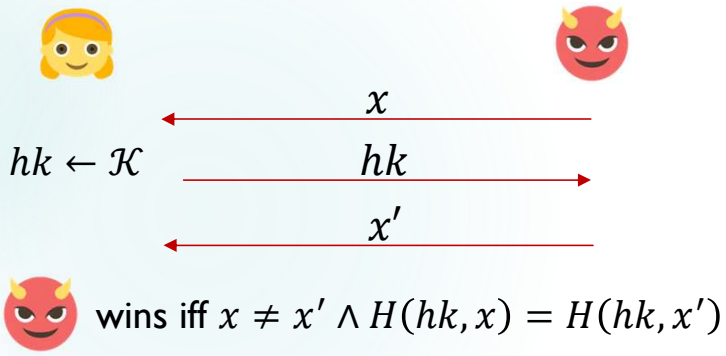
Application: dCRH implies constant-round statistically-hiding commitment scheme

Other Relaxations of CRH

$$H: \mathcal{K} \times \{0,1\}^m \to \{0,1\}^\ell \quad (\ell < m)$$

Target-collision resistance
Also known as **universal one-way hash function (UOWHF)**

Application: digital signature

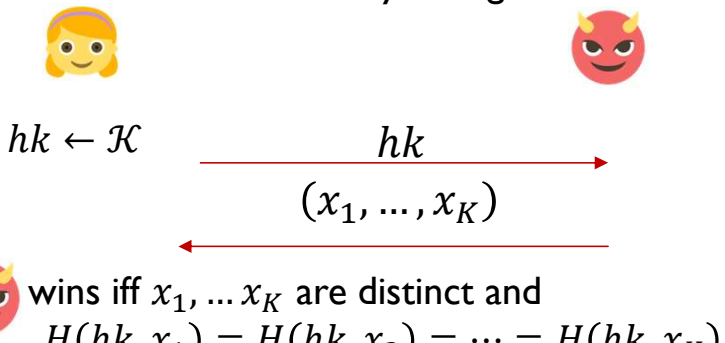


wins iff $x \neq x' \wedge H(hk, x) = H(hk, x')$

***K*-multi-collision resistance (*K*-MCR)**

Applications: 4-message zero-knowledge argument

constant-round statistically hiding commitment schemes



wins iff $x_1, \ldots x_K$ are distinct and
$H(hk, x_1) = H(hk, x_2) = \cdots = H(hk, x_K)$

2-MCRH = standard CRH

We require that all PPT wins with negligible probability.

What Was Known for Random Degree Two Functions

- Zhandry: The differential attack can be generalized to find many collisions.
  - Choose $\Delta_1, \ldots, \Delta_k$ and solve the linear system $h(x) = h(x + \Delta_1) = \cdots = h(x + \Delta_k)$.
  - $\mathcal{Q}_{p,N,M}$ is not multi-collision-resistant.
- AHI+17: Assuming the intractability of solving a random system of quadratic equations over $\mathbb{F}_p$, $\mathcal{Q}_{p,N,M}$ is a UOWHF.
- **Question**: Is $\mathcal{Q}_{p,N,M}$ distributionally collision-resistant (under reasonable assumptions)?

New: Random Degree-2 Functions Are Not **Distributionally** Collision-Resistant

Differential attack $D(h)$

1. Pick difference vector $\Delta \in \mathbb{F}_p^N \setminus \{0\}$ **uniformly at random**.
2. Pick a solution x^* to the linear system $h(x) = h(x + \Delta)$ **uniformly at random**.
3. Output $(x^*, x^* + \Delta)$.

Theorem. For a random $h \leftarrow \mathcal{Q}_{p,N,M}$, the statistical distance between $(h, D(h))$ and the ideal collision distribution is $O(p^{-(N-M)/2})$.

Corollary. If p^{N-M} is superpolynomial, then $\mathcal{Q}_{p,N,M}$ is not a dCRH.

Proof Outline

For $h \leftarrow \mathcal{Q}_{p,N,M}$,

$$\big(h, D(h)\big) \approx_s \big(h, \mathrm{AllCol}(h)\big) \approx_s (h, \mathrm{IdealCol}(h))$$

- $\mathrm{AllCol}(h)$: uniform distribution over $\{(x, x') : h(x) = h(x')\}$.
- $\mathrm{IdealCol}(h)$: the ideal collision distribution in the definition of dCRH.
  - Sample $x \leftarrow \mathbb{F}_p^N, x' \leftarrow \{w : h(w) = h(x)\}$; output (x, x').

$$\big(h, D(h)\big) \approx_s \big(h, \mathrm{AllCol}(h)\big)$$

- Let $\big(H, (X, Y)\big)$ and $\big(H, (X', Y')\big)$ denote the two distributions respectively.
- For a fixed Δ, $h(x) = h(x + \Delta)$ is a linear system $A_{h,\Delta} x = b_\Delta$ where

$$A_{h,\Delta} = \begin{pmatrix} \Delta^\top (Q_1 + Q_1^\top) \\ \vdots \\ \Delta^\top (Q_N + Q_N^\top) \end{pmatrix}, \qquad b_\Delta = \begin{pmatrix} \Delta^\top Q_1 \Delta + \ell_1^\top \Delta \\ \vdots \\ \Delta^\top Q_N \Delta + \ell_N^\top \Delta \end{pmatrix}$$

Lemma. For a random h, $A_{h,\Delta}, b_\Delta$ are uniformly random.

- Call Δ **good** if $A_{h,\Delta}$ is full-rank.
- **Observation.** $\big(H, (X, Y)\big)|_{Y-X \text{ is good}} \equiv \big(H, (X', Y')\big)|_{Y'-X' \text{is good}}$
- Statistical distance is bounded by

$$\Pr[(Y - X) \text{ is not good}] + \Pr[(Y' - X') \text{ is not good}] \leq O(p^{M-N})$$

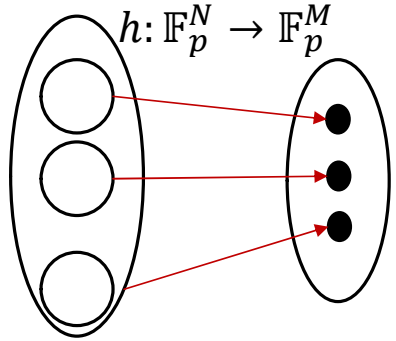
Lemma + direct calculation

$$\left(h, \mathrm{AllCol}(h)\right) \approx_s \left(h, \mathrm{IdealCol}(h)\right)$$

$\forall y, |h^{-1}(y)| = p^{N-M}$

- h is **regular** → two distributions are identical except on diagonal pairs (x, x).
- We define a quantity called **fiber variance** to measure how regular h is.

$$v(h) := \frac{1}{p^M} \sum_{y \in \mathbb{F}_p^M} \left(\frac{|h^{-1}(y)|}{p^{N-M}} - 1 \right)^2$$



- **Lemma 1.** For every fixed h, $SD\left(\mathrm{AllCol}(h), \mathrm{IdealCol}(h)\right) \le \sqrt{v(h)} + p^{M-N}$.
- **Lemma 2.** $\mathcal{Q}_{p,N,M}$ is pairwise independent → $\mathbb{E}_{h \leftarrow \mathcal{Q}_{p,N,M}}[v(h)] = O(p^{M-N})$.

Differential attack $D(h)$

1. Pick difference vector $\Delta \in \mathbb{F}_p^N \setminus \{0\}$ **uniformly at random**.
2. Pick a solution x^* to the linear system $h(x) = h(x + \Delta)$ **uniformly at random**.
3. Output $(x^*, x^* + \Delta)$.

For $h \leftarrow Q_{p,N,M}$,

$$\big(h, D(h)\big) \approx_s \big(h, \mathrm{AllCol}(h)\big) \approx_s (h, \mathrm{IdealCol}(h))$$

Thanks for Listening ☺