Malleable SNARKs and Their Applications

Suvradip Chakraborty¹ Dennis Hofheinz²

Christoph Striecks⁴ Daniele Venturi⁵

Roman Langrehr²

Jesper Buus Nielsen³

¹Visa Research

²FTH Zurich

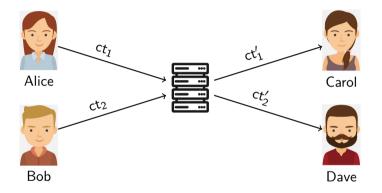
³Aarhus University

⁴AIT Austrian Institute of Technology

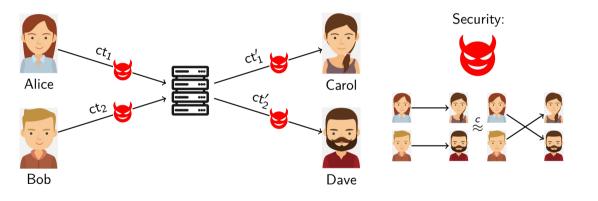
⁵Sapienza University of Rome

2025-05-08

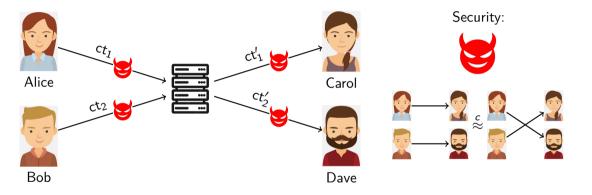
Motivation: Mix-Nets (over-simplified)



Motivation: Mix-Nets (over-simplified)

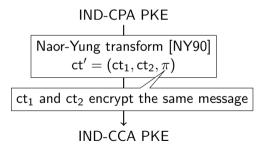


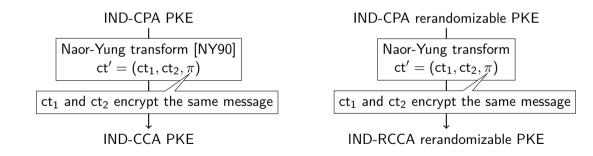
Motivation: Mix-Nets (over-simplified)

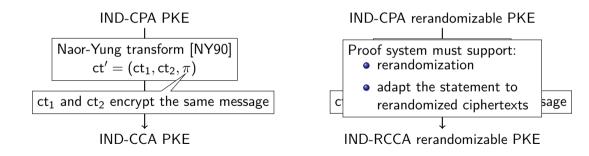


Solution: **e** rerandomizes ciphertexts (+ anonymity of the encryption scheme)

Dec oracle returns \diamond if ciphertext decrypts to m_0^{\star} or m_1^{\star}







- \bullet NP relation ${\cal R}$
- \bullet Allowed transformations ${\cal T}$

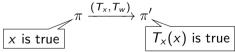
$$\forall (x,w) \in \mathcal{R}, (T_x, T_w) \in \mathcal{T} \rightarrow (T_x(x), T_w(w)) \in \mathcal{R}$$

- \bullet NP relation ${\cal R}$
- \bullet Allowed transformations ${\cal T}$

$$orall (x,w) \in \mathcal{R}, (\mathit{T}_x,\mathit{T}_w) \in \mathcal{T}
ightarrow (\mathit{T}_x(x),\mathit{T}_w(w)) \in \mathcal{R}$$

A malleable NIZK for ${\mathcal R}$ and ${\mathcal T}$ is

- \bullet a NIZK for ${\cal R}$
- with the following extra feature



- $\bullet \ \mathsf{NP} \ \mathsf{relation} \ \mathcal{R}$
- \bullet Allowed transformations ${\cal T}$

$$orall (x,w) \in \mathcal{R}, (\mathit{T}_x,\mathit{T}_w) \in \mathcal{T}
ightarrow (\mathit{T}_x(x),\mathit{T}_w(w)) \in \mathcal{R}$$

A malleable NIZK for ${\mathcal R}$ and ${\mathcal T}$ is

- \bullet a NIZK for ${\cal R}$
- with the following extra feature

$$x \text{ is true} \qquad \pi \xrightarrow{(T_x, T_w)} \pi' \xrightarrow{T_x(x) \text{ is true}}$$

Derivation-privacy: (π, π') indistinguishable from directly generated proofs for x and $T_x(x)$

- $\bullet \ \mathsf{NP} \ \mathsf{relation} \ \mathcal{R}$
- \bullet Allowed transformations ${\cal T}$

$$orall (x,w) \in \mathcal{R}, (\mathit{T}_x,\mathit{T}_w) \in \mathcal{T}
ightarrow (\mathit{T}_x(x),\mathit{T}_w(w)) \in \mathcal{R}$$

A malleable NIZK for ${\mathcal R}$ and ${\mathcal T}$ is

- \bullet a NIZK for ${\cal R}$
- with the following extra feature

$$\begin{array}{c} \pi \xrightarrow{(T_x,T_w)} \pi' \\ \hline T_x(x) \text{ is true} \end{array}$$

Derivation-privacy: (π, π') indistinguishable from directly generated proofs for x and $T_x(x)$ Simulation-soundness: Takes into account that simulated statements can be modified with T Groth-Sahai proofs [CKLM12]:

- \checkmark Very efficient
- Only for a specific R (pairing-product equations)
- $\pmb{\mathsf{X}}$ Only for a specific $\mathcal T$
- × Not post-quantum secure

Groth-Sahai proofs [CKLM12]:

- \checkmark Very efficient
- Only for a specific R (pairing-product equations)
- $\pmb{\times}$ Only for a specific ${\cal T}$
- × Not post-quantum secure

Generically from SNARKs [CKLM13]:

- \checkmark Any NP-relation ${\cal R}$
- ✓ Any set of valid transformations T
- Post-quantum instantiation possible
- X Number of transformations is bounded
- × Not derivation private

Groth-Sahai proofs [CKLM12]:

- \checkmark Very efficient
- Only for a specific R (pairing-product equations)
- $\pmb{\mathsf{X}}$ Only for a specific $\mathcal T$
- × Not post-quantum secure

Generically from SNARKs [CKLM13]:

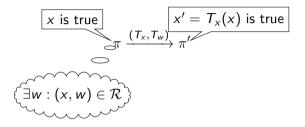
- \checkmark Any NP-relation ${\cal R}$
- ✓ Any set of valid transformations *T*
- Post-quantum instantiation possible
- X Number of transformations is bounded
- × Not derivation private

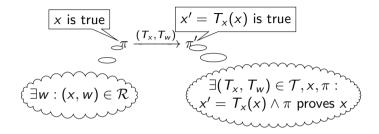
Generically from SNARKs [our work]:

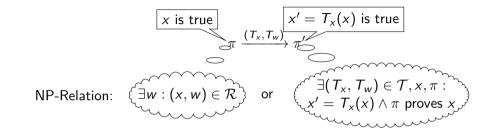
- \checkmark Any NP-relation ${\cal R}$
- $\checkmark \mbox{ Any set of valid} \\ \mbox{ transformations } \mathcal{T}$
- Post-quantum instantiation possible
- ✓ Unbounded transformations
- ✓ Derivation private

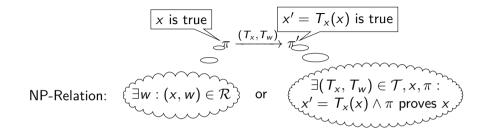
x is true

$$\pi \xrightarrow{(T_x, T_w)} \pi'$$
 $x' = T_x(x)$ is true









Recursive usage of SNARKs

- succinctness avoids blow-up of proof size/verification time
- ✓zero-knowledge ✓derivation private (by zero-knowledge of the SNARK)
- also used for incrementally verifiable computation (IVC), proof carrying data (PCD), blockchains (to compress proofs)

Soundness (SNARK variant): $\forall A$ outputting (x, π) s.t π proves x $\exists \mathcal{E}$ outputting w s.t. $(x, w) \in \mathcal{R}$.

```
Soundness (SNARK variant): \forall \mathcal{A} \text{ outputting } (x, \pi) \text{ s.t } \pi \text{ proves } x \\ \exists \mathcal{E} \text{ outputting } w \text{ s.t. } (x, w) \in \mathcal{R}.
```

In our case: \mathcal{E} can output

• w s.t. $(x, w) \in \mathcal{R}$

Soundness (SNARK variant): $\forall A$ outputting (x, π) s.t π proves $x \exists \mathcal{E}$ outputting w s.t. $(x, w) \in \mathcal{R}$.

In our case: ${\mathcal E}$ can output

• w s.t. $(x, w) \in \mathcal{R}$

•
$$(T_x, T_w) \in \mathcal{T}, \hat{x}, \hat{\pi} \text{ s.t. } x = T_x(\hat{x}) \text{ and } \underline{\hat{\pi} \text{ proves } \hat{x}}$$



Soundness (SNARK variant): $\forall A$ outputting (x, π) s.t π proves $x \exists \mathcal{E}$ outputting w s.t. $(x, w) \in \mathcal{R}$.

In our case: ${\mathcal E}$ can output

• w s.t. $(x, w) \in \mathcal{R}$

•
$$(T_x, T_w) \in \mathcal{T}, \hat{x}, \hat{\pi} \text{ s.t. } x = T_x(\hat{x}) \text{ and } \hat{\pi} \text{ proves } \hat{x}$$

Recursive extraction



Soundness (SNARK variant): $\forall \mathcal{A} \text{ outputting } (x, \pi) \text{ s.t } \pi \text{ proves } x \\ \exists \mathcal{E} \text{ outputting } w \text{ s.t. } (x, w) \in \mathcal{R}.$ In our case: \mathcal{E} can output • $w \text{ s.t. } (x, w) \in \mathcal{R}$ • $(T_x, T_w) \in \mathcal{T}, \hat{x}, \hat{\pi} \text{ s.t. } x = T_x(\hat{x}) \text{ and } \frac{\hat{\pi} \text{ proves } \hat{x}}{\hat{\mu}}$ Recursive extraction



Issue 1: Runtime of the extractor can explode

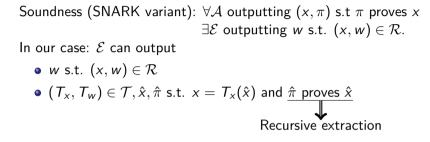
Soundness (SNARK variant): $\forall \mathcal{A} \text{ outputting } (x, \pi) \text{ s.t } \pi \text{ proves } x \\ \exists \mathcal{E} \text{ outputting } w \text{ s.t. } (x, w) \in \mathcal{R}.$ In our case: \mathcal{E} can output • $w \text{ s.t. } (x, w) \in \mathcal{R}$ • $(T_x, T_w) \in \mathcal{T}, \hat{x}, \hat{\pi} \text{ s.t. } x = T_x(\hat{x}) \text{ and } \frac{\hat{\pi} \text{ proves } \hat{x}}{p}$

Issue 1: Runtime of the extractor can explode

Solution: We assume a fast extractor:

 $\mathsf{Time}_\mathcal{E} \leq \mathsf{Time}_\mathcal{A} + \mathsf{poly}(\lambda)$

Recursive extraction



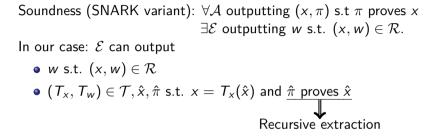


Issue 1: Runtime of the extractor can explode

Issue 2: Recursion might never stop

Solution: We assume a fast extractor:

 $\mathsf{Time}_{\mathcal{E}} \leq \mathsf{Time}_{\mathcal{A}} + \mathsf{poly}(\lambda)$





Issue 1: Runtime of the extractor can explode

Issue 2: Recursion might never stop

Solution: We assume a fast extractor:

Main technical challenge

 $\mathsf{Time}_{\mathcal{E}} \leq \mathsf{Time}_{\mathcal{A}} + \mathsf{poly}(\lambda)$

• Fixed bound *B* on recursion depth

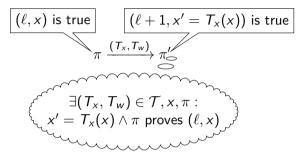
- Fixed bound *B* on recursion depth
- Add a counter ℓ to SNARK statements

- Fixed bound *B* on recursion depth
- Add a counter ℓ to SNARK statements
 - $\ell = 1$ for non-recursive proofs

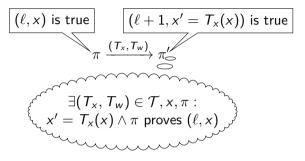
- Fixed bound *B* on recursion depth
- Add a counter ℓ to SNARK statements
 - $\ell = 1$ for non-recursive proofs
 - ℓ is incremented by 1 in each recursion

- Fixed bound *B* on recursion depth
- Add a counter ℓ to SNARK statements
 - $\ell = 1$ for non-recursive proofs
 - ℓ is incremented by 1 in each recursion
 - Recursion is only allowed for $\ell < B$

- Fixed bound *B* on recursion depth
- Add a counter ℓ to SNARK statements
 - $\ell = 1$ for non-recursive proofs
 - ℓ is incremented by 1 in each recursion
 - Recursion is only allowed for $\ell < B$



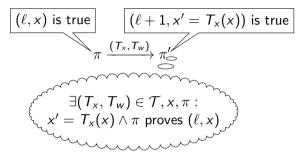
- Fixed bound *B* on recursion depth
- Add a counter ℓ to SNARK statements
 - $\ell = 1$ for non-recursive proofs
 - ℓ is incremented by 1 in each recursion
 - Recursion is only allowed for $\ell < B$



✓ Soundness

Warm-up solution: Counters [CKLM13]

- Fixed bound *B* on recursion depth
- Add a counter ℓ to SNARK statements
 - $\ell = 1$ for non-recursive proofs
 - ℓ is incremented by 1 in each recursion
 - Recursion is only allowed for $\ell < B$

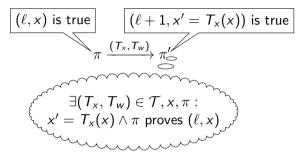


Soundness

X Limited derivation privacy (counters leak # recursions)

Warm-up solution: Counters [CKLM13]

- Fixed bound *B* on recursion depth
- Add a counter ℓ to SNARK statements
 - $\ell = 1$ for non-recursive proofs
 - ℓ is incremented by 1 in each recursion
 - Recursion is only allowed for $\ell < B$



Soundness

- X Limited derivation privacy (counters leak # recursions)
- X Only bounded number of recursions

Non-solution: One-way permutation

 $\mathsf{OWP}\ f:X\to X$

Counter ℓ is replaced by $\xi \in X$

- Initially: $\xi \leftarrow X$
- Recursion $\xi' = f(\xi)$

Idea: Extractor would have to break one-wayness if

recursion for extraction > # recursion to generate the proof (honestly)

Non-solution: One-way permutation

 $\mathsf{OWP}\ f:X\to X$

Counter ℓ is replaced by $\xi \in X$

- Initially: $\xi \leftarrow X$
- Recursion $\xi' = f(\xi)$

Idea: Extractor would have to break one-wayness if

recursion for extraction > # recursion to generate the proof (honestly)

Issue: Soundness has to hold for adversarially generated proofs E.g. adversary can choose initial $\xi \in X$ from a different distribution.

Non-solution: One-way permutation

 $\mathsf{OWP}\ f:X\to X$

Counter ℓ is replaced by $\xi \in X$

- Initially: $\xi \leftarrow X$
- Recursion $\xi' = f(\xi)$

Idea: Extractor would have to break one-wayness if

recursion for extraction > # recursion to generate the proof (honestly)

Issue: Soundness has to hold for adversarially generated proofs E.g. adversary can choose initial $\overline{\xi \in X}$ from a different distribution.

We need a variant of OWF that is secure if the adversary chooses the value *x* to invert.

Adversarial one-way functions (AOWFs)

```
Stateful 2-stage adversary (\mathcal{A}_1, \mathcal{A}_2)
```

```
1<sup>st</sup> attempt: \begin{array}{c} \mathcal{A}_1 \text{ outputs } y \\ \mathcal{A}_2 \text{ outputs } x \end{array} (\mathcal{A}_1, \mathcal{A}_2) wins if f(x) = y
```

Adversarial one-way functions (AOWFs)

Stateful 2-stage adversary $(\mathcal{A}_1, \mathcal{A}_2)$

1st attempt: $\begin{array}{c} \mathcal{A}_1 \text{ outputs } y \\ \mathcal{A}_2 \text{ outputs } x \end{array}$ $(\mathcal{A}_1, \mathcal{A}_2) \text{ wins if } f(x) = y$

Unsatisfiable! A_1 samples $x \leftarrow X$ and outputs y := f(x). A_2 outputs x.

Stateful 2-stage adversary $(\mathcal{A}_1, \mathcal{A}_2)$

1st attempt: $\begin{array}{c} \mathcal{A}_1 \text{ outputs } y \\ \mathcal{A}_2 \text{ outputs } x \end{array} (\mathcal{A}_1, \mathcal{A}_2) \text{ wins if } f(x) = y$

Unsatisfiable! A_1 samples $x \leftarrow X$ and outputs y := f(x). A_2 outputs x.

Make the problem harder for the adversary:

2nd attempt:
$$\begin{array}{c} \mathcal{A}_1 \text{ outputs } y \\ \mathcal{A}_2 \text{ outputs } x \end{array}$$
 $(\mathcal{A}_1, \mathcal{A}_2) \text{ wins if } f^n(x) = y$

Stateful 2-stage adversary $(\mathcal{A}_1, \mathcal{A}_2)$

1st attempt: $\begin{array}{c} \mathcal{A}_1 \text{ outputs } y \\ \mathcal{A}_2 \text{ outputs } x \end{array} (\mathcal{A}_1, \mathcal{A}_2) \text{ wins if } f(x) = y$

Unsatisfiable! A_1 samples $x \leftarrow X$ and outputs y := f(x). A_2 outputs x.

Make the problem harder for the adversary:

2nd attempt:
$$\begin{array}{c} \mathcal{A}_1 \text{ outputs } y \\ \mathcal{A}_2 \text{ outputs } x \end{array}$$
 $(\mathcal{A}_1, \mathcal{A}_2) \text{ wins if } f^n(x) = y$

lf

• *f* is only <u>sequentially</u> computable

Stateful 2-stage adversary $(\mathcal{A}_1, \mathcal{A}_2)$

1st attempt: $\begin{array}{c} \mathcal{A}_1 \text{ outputs } y \\ \mathcal{A}_2 \text{ outputs } x \end{array} (\mathcal{A}_1, \mathcal{A}_2) \text{ wins if } f(x) = y$

Unsatisfiable! A_1 samples $x \leftarrow X$ and outputs y := f(x). A_2 outputs x.

Make the problem harder for the adversary:

2nd attempt:
$$\begin{array}{c} \mathcal{A}_1 \text{ outputs } y \\ \mathcal{A}_2 \text{ outputs } x \end{array}$$
 $(\mathcal{A}_1, \mathcal{A}_2) \text{ wins if } f^n(x) = y$

lf

- *f* is only <u>sequentially</u> computable
- *n* depends on the runtime of A_1 (Time_{A_1} < $n \cdot \text{Time}_{f(\cdot)}$)

there is no trivial attack!

• Justification: RO is an AOWFs

- Justification: RO is an AOWFs
- ✓ Soundness

- Justification: RO is an AOWFs
- ✓ Soundness
- ✓ Unbounded # recursions

- Justification: RO is an AOWFs
- ✓ Soundness
- ✓ Unbounded # recursions
- X Derivation privacy

- Justification: RO is an AOWFs
- ✓ Soundness
- ✓ Unbounded # recursions
- X Derivation privacy
- ✓ Derivation privacy if AOWF is unlinkable (requires randomization)

- Justification: RO is an AOWFs
- ✓ Soundness
- ✓ Unbounded # recursions
- X Derivation privacy
- ✓ Derivation privacy if AOWF is <u>unlinkable</u> (requires randomization)
 - For our candidate this can be achieved by inputting additional random bits in the hash function

Applications

- Reverse firewalls (RF)
 - First RF for NIZKs/SNARKs

- Reverse firewalls (RF)
 - $\bullet~\mbox{First}~\mbox{RF}~\mbox{for}~\mbox{NIZKs}/\mbox{SNARKs}$
- IND-RCCA secure rerandomizable PKE (adaptation of [CKLM12])
 - First post-quantum scheme

- Reverse firewalls (RF)
 - $\bullet~\mbox{First}~\mbox{RF}~\mbox{for}~\mbox{NIZKs}/\mbox{SNARKs}$
- IND-RCCA secure rerandomizable PKE (adaptation of [CKLM12])
 - First post-quantum scheme
- IND-RCCA secure updatable encryption (adaptation of [KLR19])
 - First scheme with full security
 - First post-quantum scheme

- Reverse firewalls (RF)
 - $\bullet~\mbox{First}~\mbox{RF}~\mbox{for}~\mbox{NIZKs}/\mbox{SNARKs}$
- IND-RCCA secure rerandomizable PKE (adaptation of [CKLM12])
 - First post-quantum scheme
- IND-RCCA secure updatable encryption (adaptation of [KLR19])
 - First scheme with full security
 - First post-quantum scheme
- Targeted non-malleable (TNM-CCA1) homomorphic encryption (HE) (non-generic) (adaptation of [BSW12])
 - Needs higher arity transformations
 - First unlinkable scheme
 - First scheme with unbounded homomorphic operations

- Reverse firewalls (RF)
 - First RF for NIZKs/SNARKs
- IND-RCCA secure rerandomizable PKE (adaptation of [CKLM12])
 - First post-quantum scheme
- IND-RCCA secure updatable encryption (adaptation of Requires assumption specific
 - First scheme with full security
 - First post-quantum scheme
- Targeted non-malleable (TNM-CCA1) homomorphic encryption (HE) (non-generic) (adaptation of [BSW12])
 - Needs higher arity transformations
 - First unlinkable scheme
 - First scheme with unbounded homomorphic operations

to the underlying HE scheme

Our contributions

• Malleable NIZK's from recursive SNARKS

Our contributions

- Malleable NIZK's from recursive SNARKS
 - Counter-based approach (similar to [CKLM13])

Our contributions

- Malleable NIZK's from recursive SNARKS
 - Counter-based approach (similar to [CKLM13])
 - AOWF-based approach is more flexible

Our contributions

- Malleable NIZK's from recursive SNARKS
 - Counter-based approach (similar to [CKLM13])
 - AOWF-based approach is more flexible
- Many applications of malleable SNARKS
 - Reverse firewalls
 - Variants of Naor-Yung

Our contributions

- Malleable NIZK's from recursive SNARKS
 - Counter-based approach (similar to [CKLM13])
 - AOWF-based approach is more flexible
- Many applications of malleable SNARKS
 - Reverse firewalls
 - Variants of Naor-Yung

Open problems

- More AOWF candidates
 - Relation to time-lock puzzles/verfiable delay functions/proofs of sequential work

Our contributions

- Malleable NIZK's from recursive SNARKS
 - Counter-based approach (similar to [CKLM13])
 - AOWF-based approach is more flexible
- Many applications of malleable SNARKS
 - Reverse firewalls
 - Variants of Naor-Yung

Open problems

- More AOWF candidates
 - Relation to time-lock puzzles/verfiable delay functions/proofs of sequential work
- Good candidates for SNARKs with fast extraction

Dan Boneh, Gil Segev, and Brent Waters.

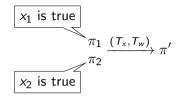
Targeted malleability: homomorphic encryption for restricted computations. In Shafi Goldwasser, editor, ITCS 2012: 3rd Innovations in Theoretical Computer Science, pages 350–366, Cambridge, MA, USA, January 8–10, 2012. Association for Computing Machinery. doi:10.1145/2090236.2090264.

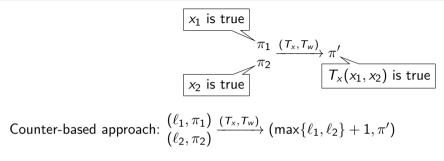
 Melissa Chase, Markulf Kohlweiss, Anna Lysyanskaya, and Sarah Meiklejohn.
 Malleable proof systems and applications.
 In David Pointcheval and Thomas Johansson, editors, <u>Advances in Cryptology –</u> <u>EUROCRYPT 2012</u>, volume 7237 of <u>Lecture Notes in Computer Science</u>, pages 281–300, Cambridge, UK, April 15–19, 2012. Springer, Berlin, Heidelberg, Germany. doi:10.1007/978-3-642-29011-4_18. Melissa Chase, Markulf Kohlweiss, Anna Lysyanskaya, and Sarah Meiklejohn.
 Succinct malleable NIZKs and an application to compact shuffles.
 In Amit Sahai, editor, <u>TCC 2013</u>: 10th Theory of Cryptography Conference, volume 7785 of <u>Lecture Notes in Computer Science</u>, pages 100–119, Tokyo, Japan, March 3–6, 2013.
 Springer, Berlin, Heidelberg, Germany.
 doi:10.1007/978-3-642-36594-2_6.

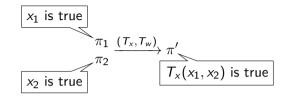
 Michael Klooß, Anja Lehmann, and Andy Rupp.
 (R)CCA secure updatable encryption with integrity protection.
 In Yuval Ishai and Vincent Rijmen, editors, <u>Advances in Cryptology – EUROCRYPT 2019</u>, <u>Part I</u>, volume 11476 of <u>Lecture Notes in Computer Science</u>, pages 68–99, Darmstadt, Germany, May 19–23, 2019. Springer, Cham, Switzerland. doi:10.1007/978-3-030-17653-2_3.

🗟 Moni Naor and Moti Yung.

Public-key cryptosystems provably secure against chosen ciphertext attacks. In <u>22nd Annual ACM Symposium on Theory of Computing</u>, pages 427–437, Baltimore, MD, USA, May 14–16, 1990. ACM Press. doi:10.1145/100216.100273. Alice, Bob, and other faces, Server: freepik.com Matryoshka doll: holz-leute.de

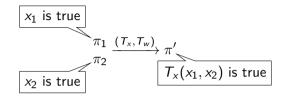






Counter-based approach: $\begin{pmatrix} \ell_1, \pi_1 \end{pmatrix} \xrightarrow{(\mathcal{T}_x, \mathcal{T}_w)} (\max\{\ell_1, \ell_2\} + 1, \pi')$

X Only for constant-depth



Counter-based approach:
$$(\ell_1, \pi_1) \xrightarrow{(\mathcal{T}_x, \mathcal{T}_w)} (\max\{\ell_1, \ell_2\} + 1, \pi')$$

X Only for constant-depth

AOWF-based approach:

- ✓ Unbounded depth
- Needs higher arity variant of AOWFs (works for hash functions)
- Statements and proofs must be input to the AOWF
- Extractor must cache extracted SNARKs

Roman Langrehr (ETH Zurich)