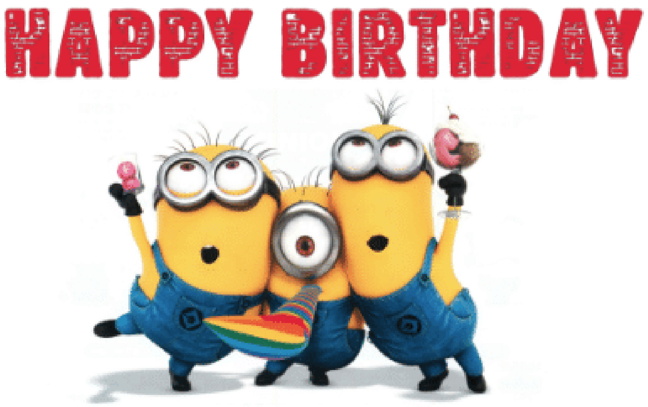
# Linking Stam's Bounds With Generalized Truncation

Bart Mennink
Radboud University (The Netherlands)

CT-RSA 2019
March 7, 2019

# Birthday Paradox




For a random selection of $23$ people, with a probability at least $50\%$ two of them share the same birthday

# Birthday Paradox

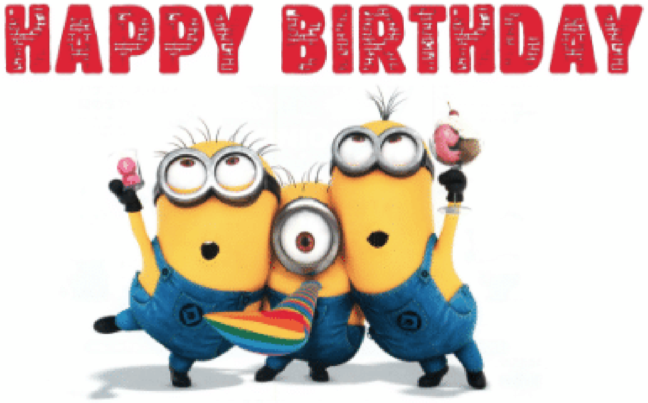

For a random selection of $23$ people, with a probability at least $50\%$ two of them share the same birthday

## General Birthday Paradox

- Consider space $\mathcal{S} = \{0,1\}^n$
- Randomly draw $q$ elements from $\mathcal{S}$
- Expected number of collisions:

$$\mathbf{Ex}\left[\text{collisions}\right] = \binom{q}{2}/2^n$$

# Birthday Paradox




For a random selection of $23$ people, with a probability at least $50\%$ two of them share the same birthday

## General Birthday Paradox

- Consider space $\mathcal{S} = \{0,1\}^n$
- Randomly draw $q$ elements from $\mathcal{S}$
- Expected number of collisions:

$$\mathbf{Ex}\left[\text{collisions}\right] = \binom{q}{2} / 2^n$$

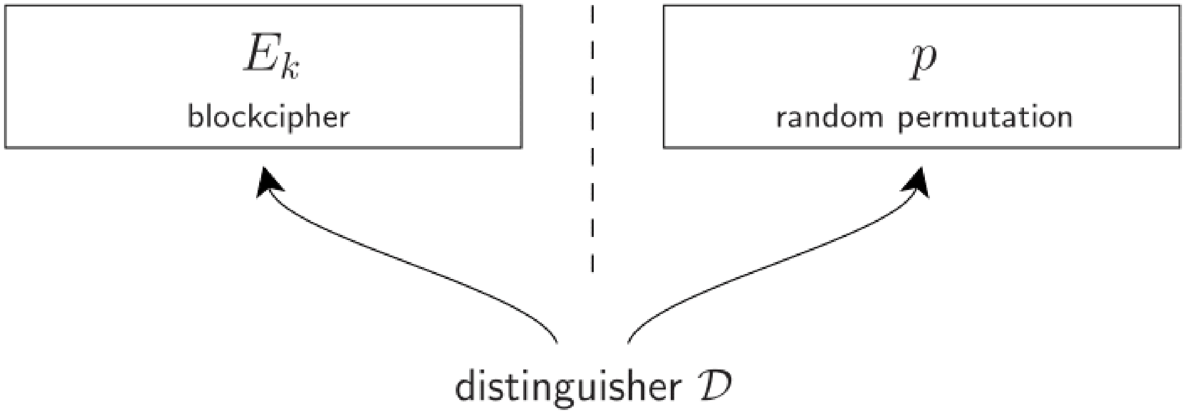
- Important phenomenon in cryptography

# Pseudorandom Permutation

| $E_k$ | $p$ |
|---|---|
| blockcipher | random permutation |

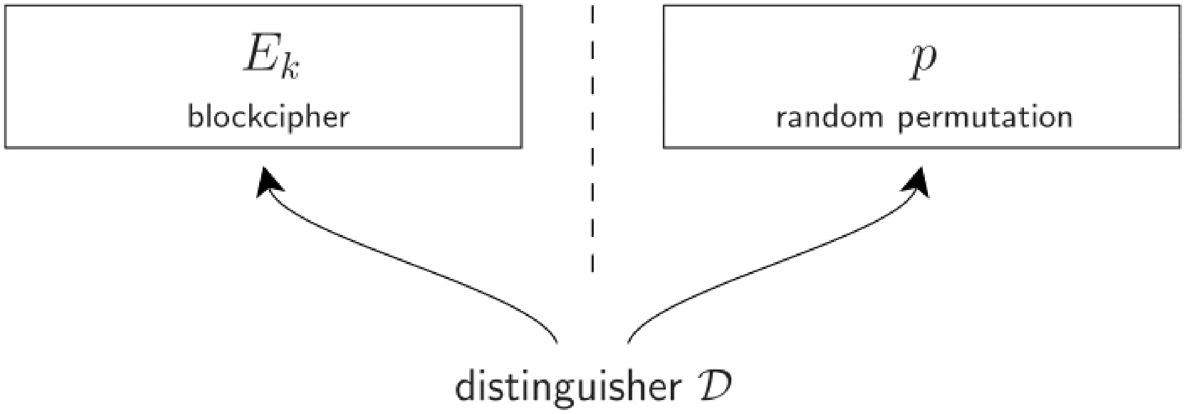
- Two oracles: $E_k$ (for secret random key $k$) and $p$

# Pseudorandom Permutation



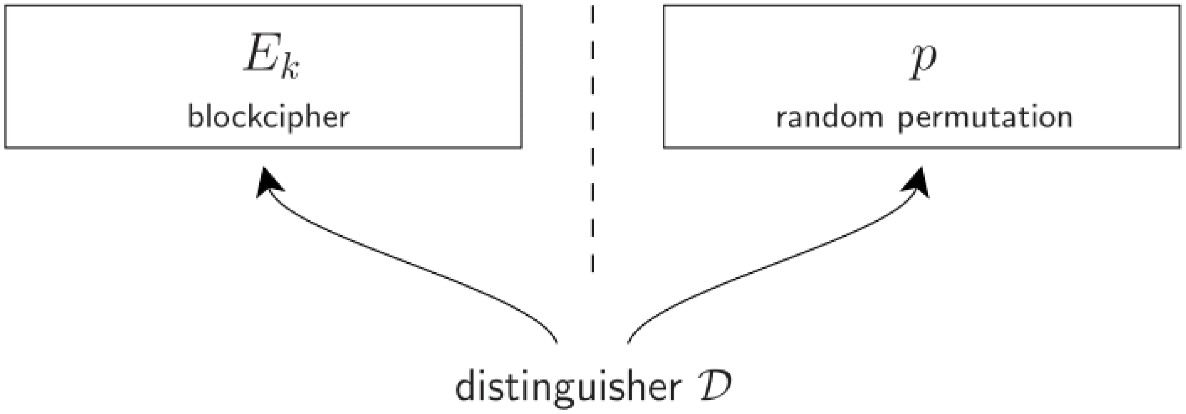

- Two oracles: $E_k$ (for secret random key $k$) and $p$
- Distinguisher $\mathcal{D}$ has query access to either $E_k$ or $p$

# Pseudorandom Permutation




- Two oracles: $E_k$ (for secret random key $k$) and $p$
- Distinguisher $\mathcal{D}$ has query access to either $E_k$ or $p$
- $\mathcal{D}$ tries to determine which oracle it communicates with

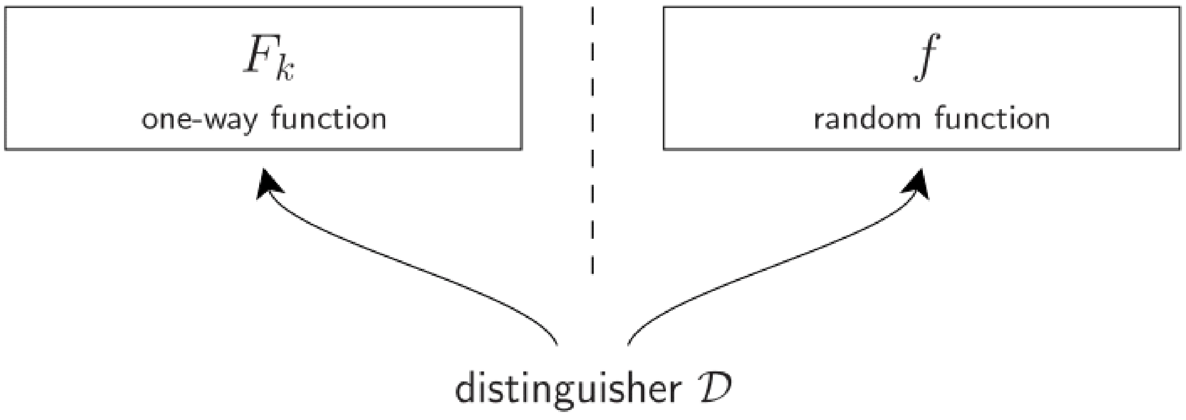
# Pseudorandom Permutation




- Two oracles: $E_k$ (for secret random key $k$) and $p$
- Distinguisher $\mathcal{D}$ has query access to either $E_k$ or $p$
- $\mathcal{D}$ tries to determine which oracle it communicates with

$$\mathbf{Adv}_E^{\mathrm{prp}}(\mathcal{D}) = \left|\mathbf{Pr}\left(\mathcal{D}^{E_k} = 1\right) - \mathbf{Pr}\left(\mathcal{D}^{p} = 1\right)\right|$$

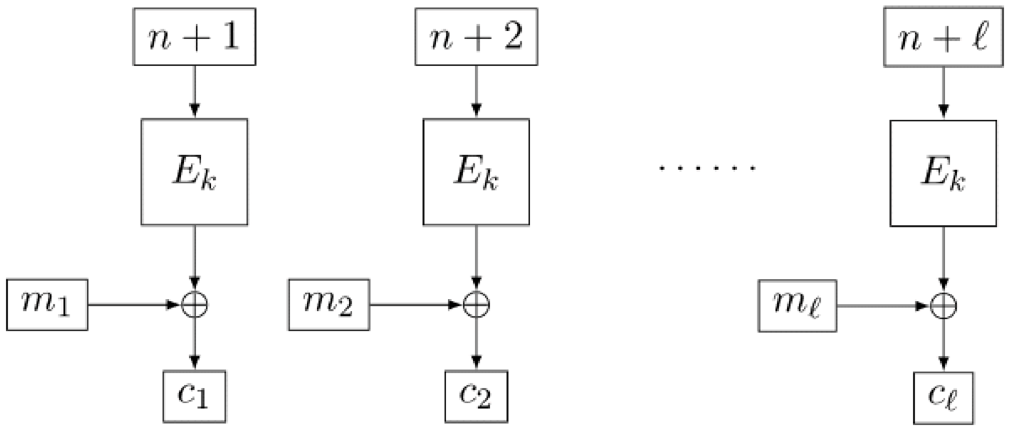
# Pseudorandom Function



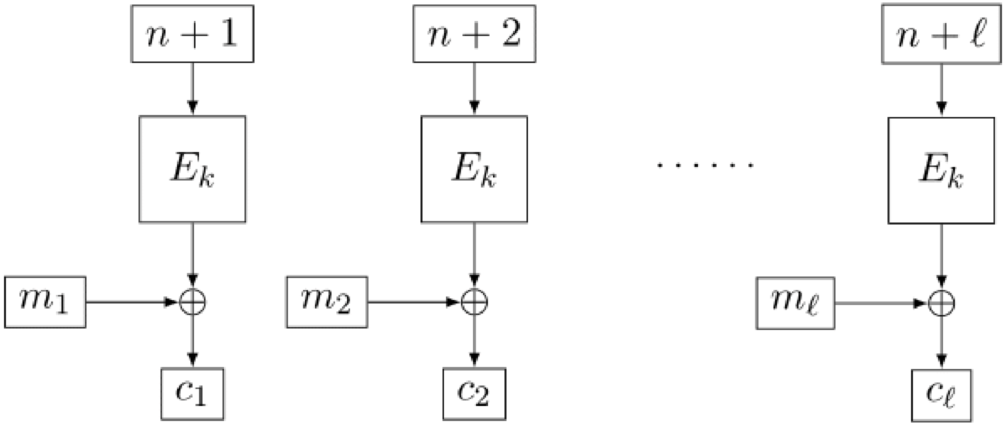

- Two oracles: $F_k$ (for secret random key $k$) and $f$
- Distinguisher $\mathcal{D}$ has query access to either $F_k$ or $f$
- $\mathcal{D}$ tries to determine which oracle it communicates with

$$\mathbf{Adv}_F^{\mathrm{prf}}(\mathcal{D}) = \left|\mathbf{Pr}\left(\mathcal{D}^{F_k} = 1\right) - \mathbf{Pr}\left(\mathcal{D}^{f} = 1\right)\right|$$

# Counter Mode Based on Pseudorandom Permutation



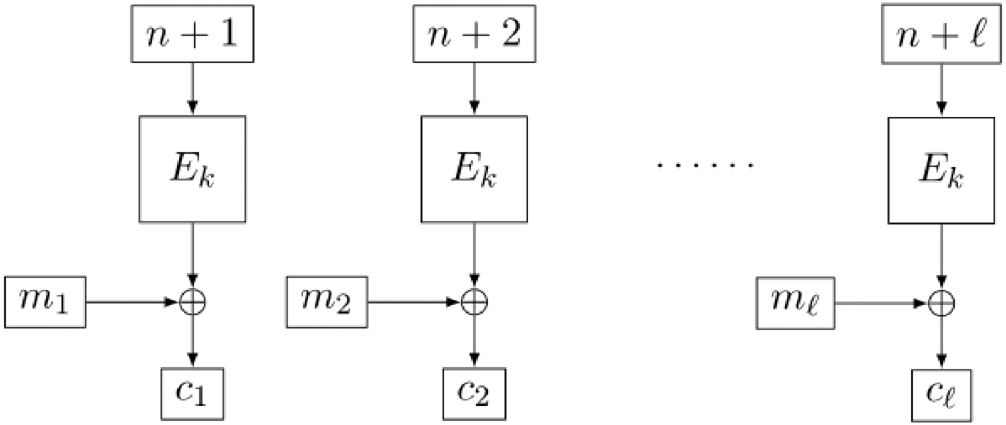
# Counter Mode Based on Pseudorandom Permutation




- Security bound:

$$\mathbf{Adv}^{\mathrm{cpa}}_{\mathsf{CTR}[E]}(\sigma) \le \mathbf{Adv}^{\mathrm{prp}}_{E}(\sigma) + \binom{\sigma}{2}/2^n$$

# Counter Mode Based on Pseudorandom Permutation

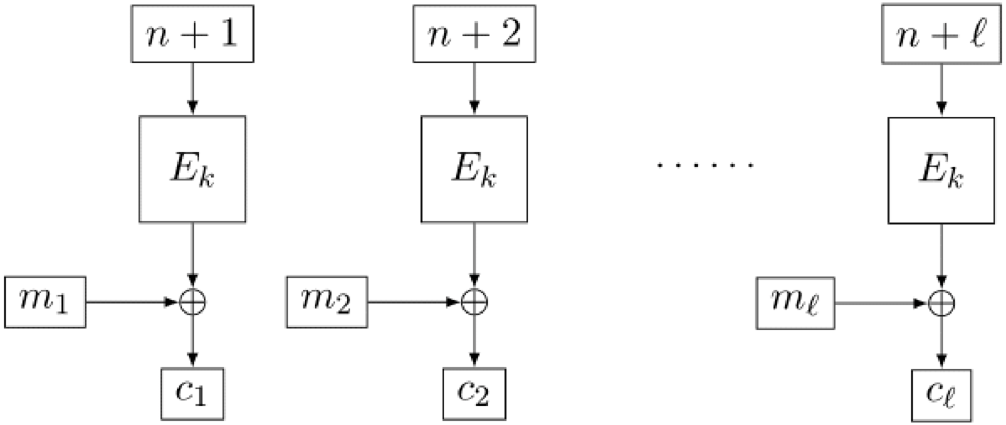



- Security bound:

$$\mathbf{Adv}^{\mathrm{cpa}}_{\mathsf{CTR}[E]}(\sigma) \leq \mathbf{Adv}^{\mathrm{prp}}_{E}(\sigma) + \binom{\sigma}{2}/2^n$$

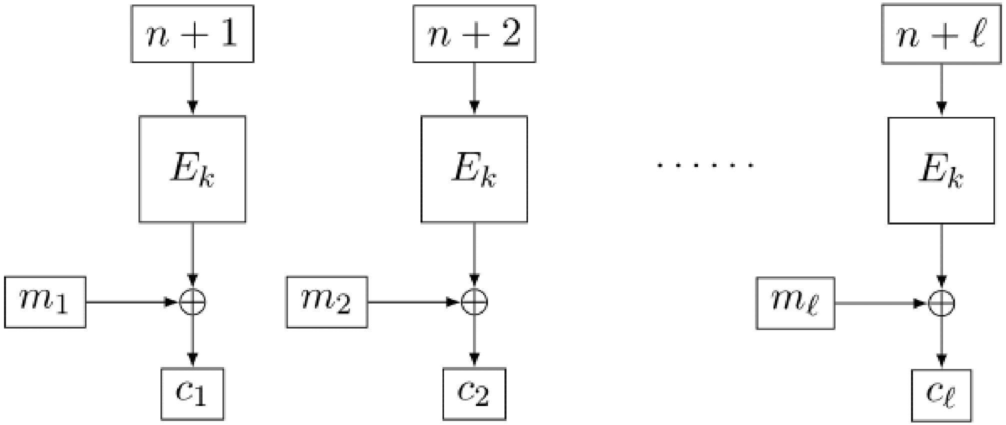
- $\mathsf{CTR}[E]$ is secure as long as:
  - $E_k$ is a secure PRP
  - Number of encrypted blocks $\sigma \ll 2^{n/2}$

# Counter Mode Based on Pseudorandom Permutation




- $m_i \oplus c_i$ is distinct for all $\sigma$ blocks
- Unlikely to happen for random string

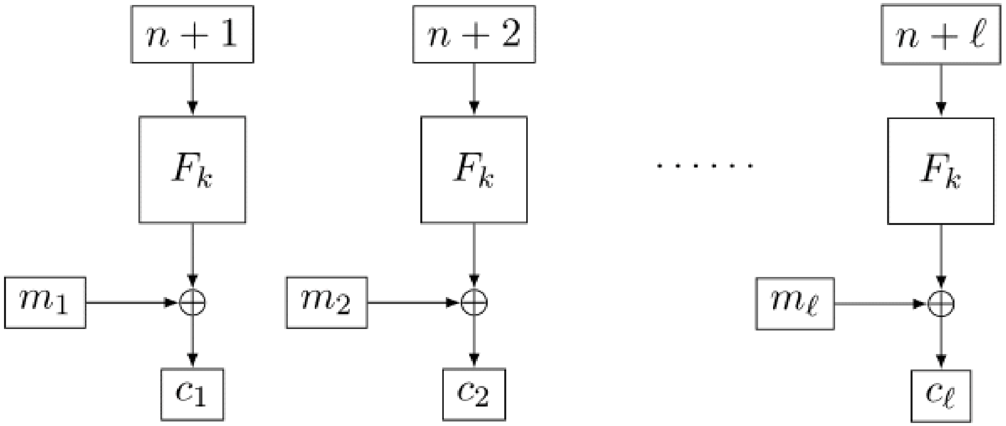
# Counter Mode Based on Pseudorandom Permutation



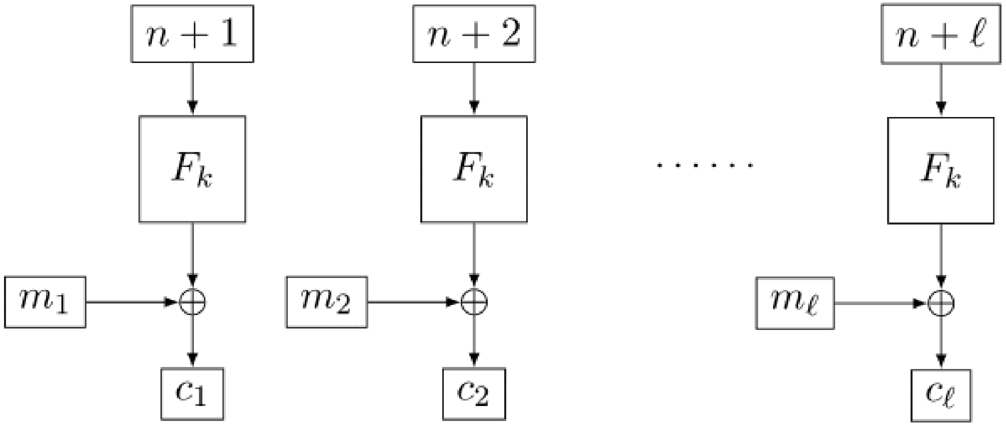

- $m_i \oplus c_i$ is distinct for all $\sigma$ blocks
- Unlikely to happen for random string
- Distinguishing attack in $\sigma \approx 2^{n/2}$ blocks:

$$\binom{\sigma}{2}/2^n \lesssim \mathbf{Adv}^{\mathrm{cpa}}_{\mathsf{CTR}[E]}(\sigma)$$

# Counter Mode Based on Pseudorandom Function



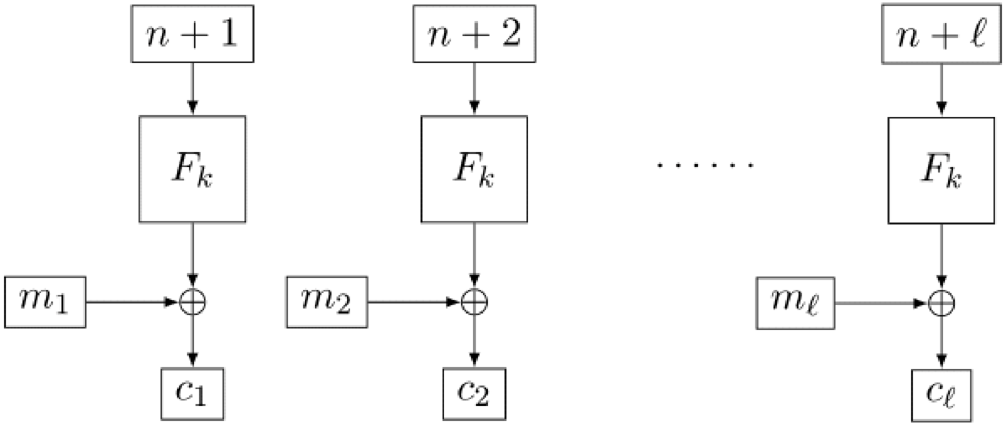
# Counter Mode Based on Pseudorandom Function




- Security bound:

$$\mathbf{Adv}^{\text{cpa}}_{\mathsf{CTR}[F]}(\sigma) \leq \mathbf{Adv}^{\text{prf}}_{F}(\sigma)$$

# Counter Mode Based on Pseudorandom Function




- Security bound:

$$\mathbf{Adv}^{\mathrm{cpa}}_{\mathsf{CTR}[F]}(\sigma) \leq \mathbf{Adv}^{\mathrm{prf}}_{F}(\sigma)$$

- $\mathsf{CTR}[F]$ is secure as long as $F_k$ is a secure PRF
- Birthday bound security loss disappeared

# Sweet32 Attack

On the Practical (In-)Security of 64-bit Block Ciphers: Collision Attacks on HTTP over TLS and OpenVPN

Bhargavan, Leurent, ACM CCS 2016


- TLS supported Triple-DES
- OpenVPN used Blowfish
- Both Blowfish and Triple-DES have 64-bit state
- Practical birthday-bound attack on encryption mode


$E_k$
conversion
$F_k$

# Various PRP-PRF Conversion Functions

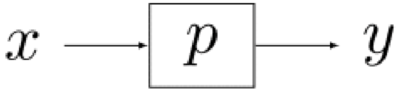
## Naive Switch

$$x \longrightarrow \boxed{p} \longrightarrow y$$

$2^{n/2}$ secure

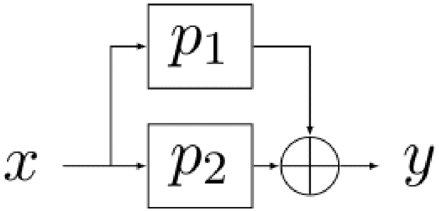
# Various PRP-PRF Conversion Functions

## Naive Switch




$2^{n/2}$ secure

## Xor of Permutations [BKR98]

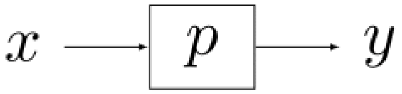



$2^n$ secure

[BI99,Luc00,Pat08,DHT17]

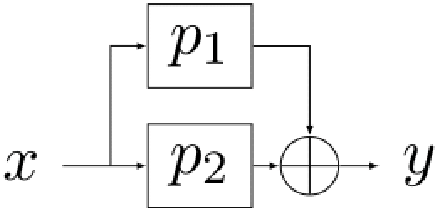
# Various PRP-PRF Conversion Functions

### Naive Switch




$2^{n/2}$ secure

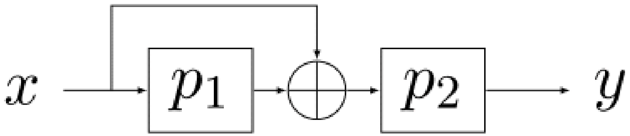
### Xor of Permutations [BKR98]




$2^n$ secure

[BI99,Luc00,Pat08,DHT17]

### EDM [CS16]

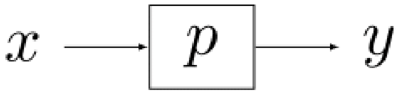



$2^n/n$ secure

[DHT17,MN17]

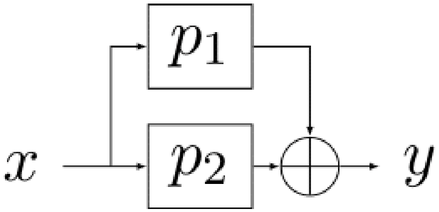
# Various PRP-PRF Conversion Functions

## Naive Switch




$2^{n/2}$ secure

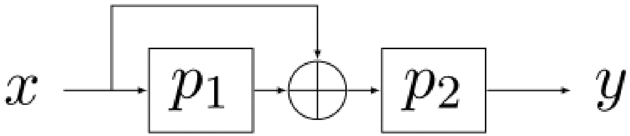
## Xor of Permutations [BKR98]




$2^n$ secure

[BI99,Luc00,Pat08,DHT17]

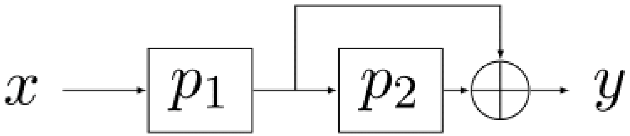
## EDM [CS16]




$2^n/n$ secure

[DHT17,MN17]

## EDMD [MN17]

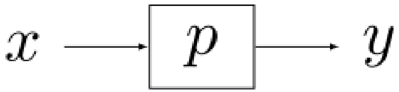



$2^n$ secure

[MN17]

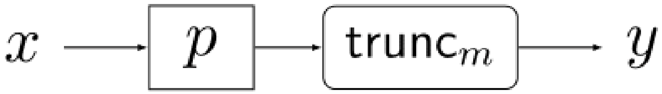
# Various PRP-PRF Conversion Functions

## Naive Switch




$2^{n/2}$ secure

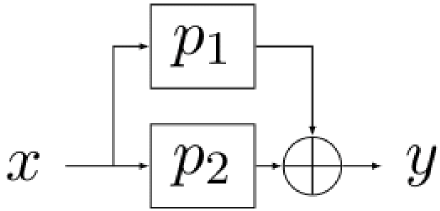
## Truncation [HWKS98]




$2^{n-m/2}$ secure

[Sta78,BI99,GG15,BN18]

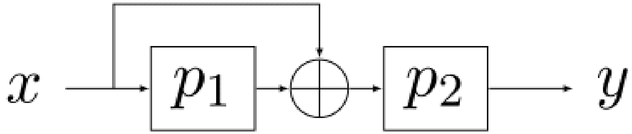
## Xor of Permutations [BKR98]




$2^n$ secure

[BI99,Luc00,Pat08,DHT17]

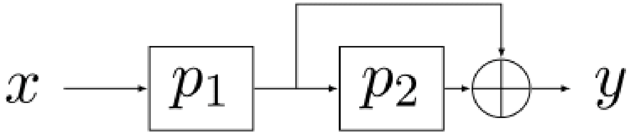
## EDM [CS16]




$2^n/n$ secure

[DHT17,MN17]

## EDMD [MN17]

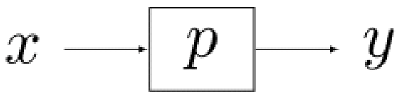



$2^n$ secure

[MN17]

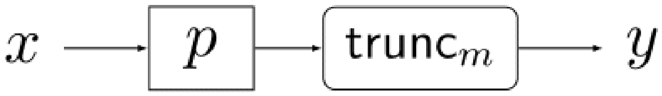
# Various PRP-PRF Conversion Functions

### Naive Switch




$2^{n/2}$ secure

### Truncation [HWKS98]

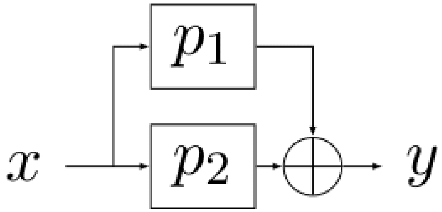



$2^{n-m/2}$ secure

[Sta78,BI99,GG15,BN18]

This work

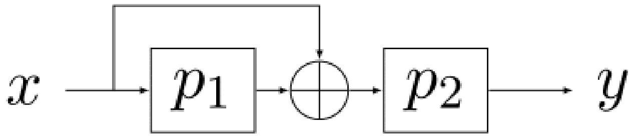
### Xor of Permutations [BKR98]




$2^n$ secure

[BI99,Luc00,Pat08,DHT17]

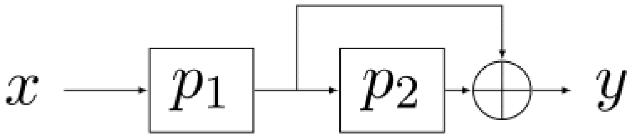
### EDM [CS16]




$2^n/n$ secure

[DHT17,MN17]

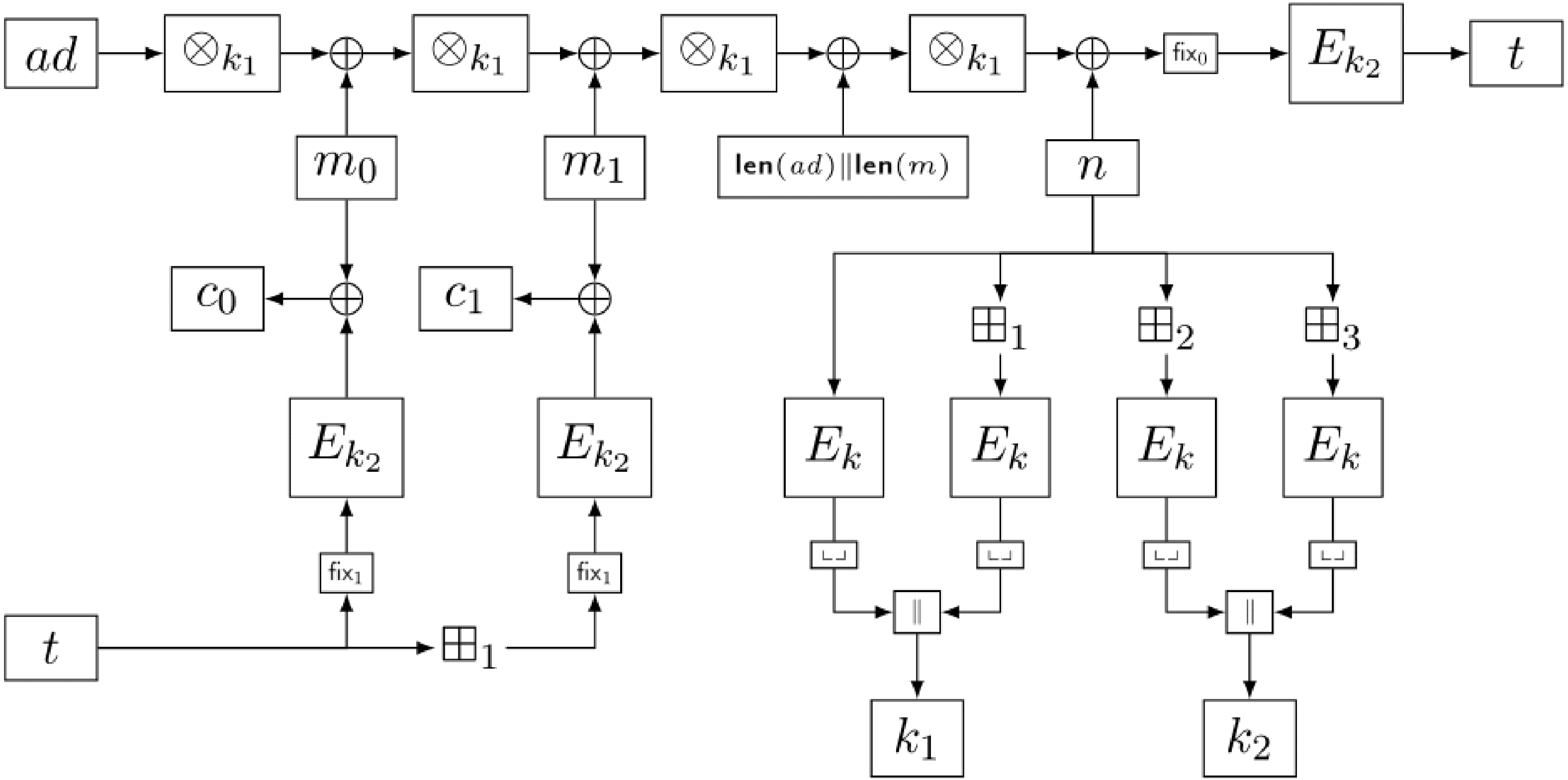
### EDMD [MN17]



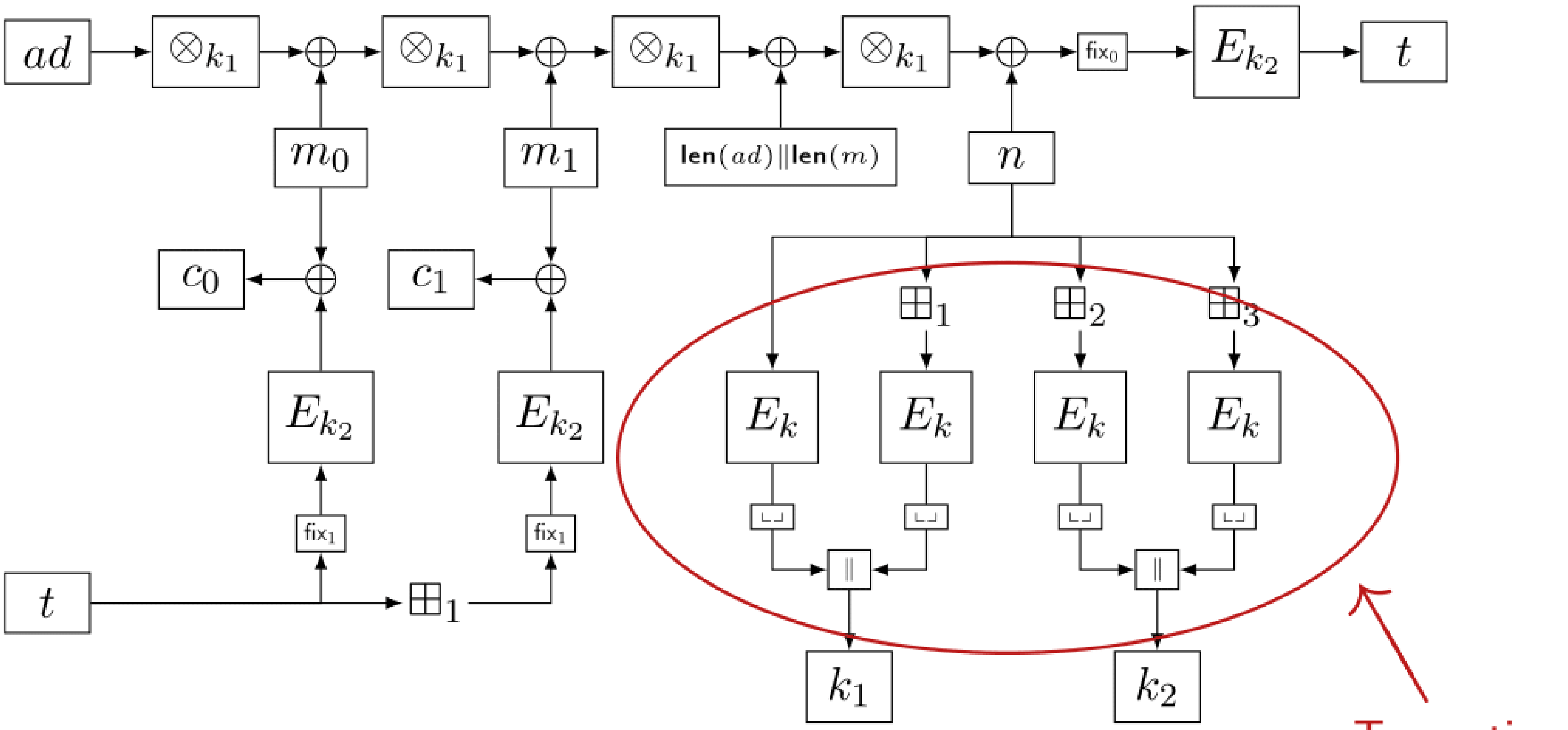

$2^n$ secure

[MN17]

# Truncation in GCM-SIV



# Truncation in GCM-SIV



# History on Truncation

1998 Hall et al.

security for most $n, m$

# History on Truncation

1998 — Hall et al.

security for most $n, m$

1999 — Bellare and Impagliazzo

covered more parameters

# History on Truncation

1998 Hall et al.

security for most $n, m$

1999 Bellare and Impagliazzo

covered more parameters

2015 Gilboa and Gueron

covered all parameters

# History on Truncation

- 1998 — Hall et al.
  security for most $n, m$
- 1999 — Bellare and Impagliazzo
  covered more parameters
- 2015 — Gilboa and Gueron
  covered all parameters
- 2016 — Gilboa and Gueron
  proof of tightness

# History on Truncation

- 1978 Stam →
- 1998 Hall et al.
  security for most $n, m$
- 1999 Bellare and Impagliazzo
  covered more parameters
- 2015 Gilboa and Gueron
  covered all parameters
- 2016 Gilboa and Gueron
  proof of tightness

**Distance between sampling with and without replacement**

by A. J. Stam*

**Summary** Two random samples of size $n$ are taken from a set containing $N$ objects of $H$ types, first with and then without replacement. Let $d$ be the absolute ($L_1$-)distance and $I$ the Kullback-Leibler information distance between the distributions of the sample compositions without and with replacement. Sample composition is meant with respect to types; it does not matter whether order of sampling is included or not. A bound on $I$ and $d$ is derived, that depends only on $n$, $N$, $H$. The bound on $I$ is not higher than $2I$. For fixed $H$ we have $d \to 0$, $I \to 0$ as $N \to \infty$ if and only if $n/N \to 0$. Let $W_r$ be the epoch at which for the $r$-th time an object of type 1 appears. Bounds on the distances between the joint distributions of $W_1, \ldots, W_r$ without and with replacement are given.

stronger bound than [HWKS98,BI99,GG15]

# History on Truncation

- 1978 Stam →
- 1998 Hall et al.
  security for most $n, m$
- 1999 Bellare and Impagliazzo
  covered more parameters
- 2015 Gilboa and Gueron
  covered all parameters
- 2016 Gilboa and Gueron
  proof of tightness
- 2018 Bhattacharya and Nandi
  reconstruction of Stam's analysis in $\chi^2$

**Distance between sampling with and without replacement**

by A. J. STAM*

**Summary** Two random samples of size $n$ are taken from a set containing $N$ objects of $H$ types, first with and then without replacement. Let $d$ be the absolute ($L_1$-)distance and $I$ the KULLBACK-LEIBLER information distance between the distributions of the sample compositions without and with replacement. Sample composition is meant with respect to types; it does not matter whether order of sampling is included or not. A bound on $I$ and $d$ is derived, that depends only on $n$, $N$, $H$. The bound on $I$ is not higher than $2I$. For fixed $H$ we have $d \to 0$, $I \to 0$ as $N \to \infty$ if and only if $n/N \to 0$. Let $W_r$ be the epoch at which for the $r$-th time an object of type 1 appears. Bounds on the distances between the joint distributions of $W_1, \ldots, W_r$ without and with replacement are given.

stronger bound than [HWKS98,BI99,GG15]

# History on Truncation

- 1978 — Stam → 
- 1998 — Hall et al.
  security for most $n, m$
- 1999 — Bellare and Impagliazzo
  covered more parameters
- 2015 — Gilboa and Gueron
  covered all parameters
- 2016 — Gilboa and Gueron
  proof of tightness
- 2018 — Bhattacharya and Nandi
  reconstruction of Stam's analysis in $\chi^2$

**Distance between sampling with and without replacement**

by A. J. STAM*

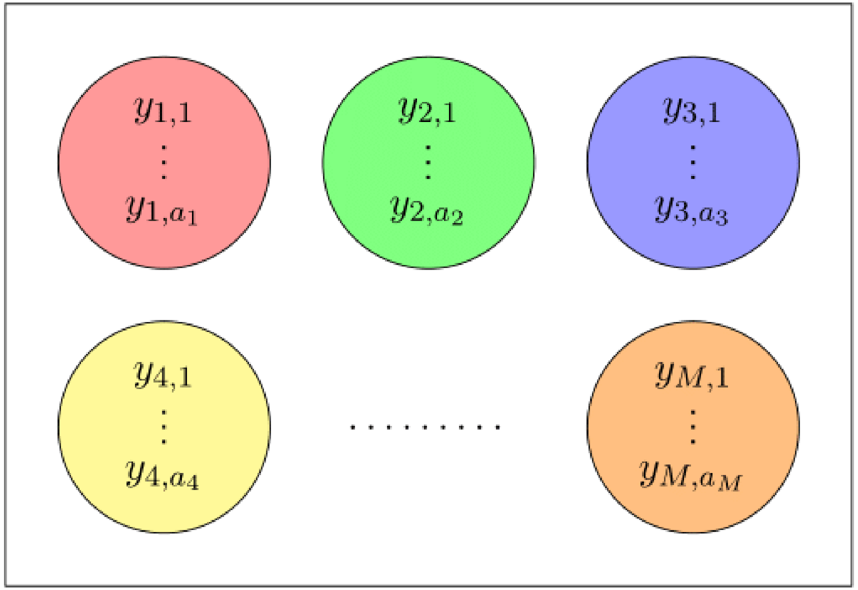
**Summary** Two random samples of size $n$ are taken from a set containing $N$ objects of $H$ types, first with and then without replacement. Let $d$ be the absolute ($L_1$-)distance and $I$ the KULLBACK-LEIBLER information distance between the distributions of the sample compositions without and with replacement. Sample composition is meant with respect to types; it does not matter whether order of sampling is included or not. A bound on $I$ and $d$ is derived, that depends only on $n$, $N$, $H$. The bound on $I$ is not higher than $2I$. For fixed $H$ we have $d \to 0$, $I \to 0$ as $N \to \infty$ if and only if $n/N \to 0$. Let $W_r$ be the epoch at which for the $r$-th time an object of type 1 appears. Bounds on the distances between the joint distributions of $W_1, \ldots, W_r$ without and with replacement are given.

stronger bound than [HWKS98,BI99,GG15]

Stam's bounds are more general [Sta78,Sta86]

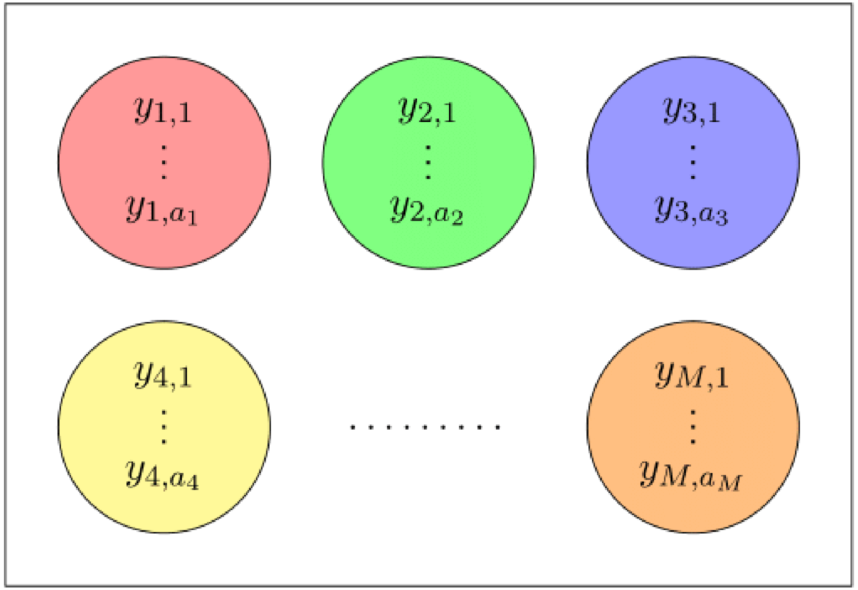
# Stam's Bound [Sta78]

- Partition $N$ elements into $M$ colors:



# Stam's Bound [Sta78]

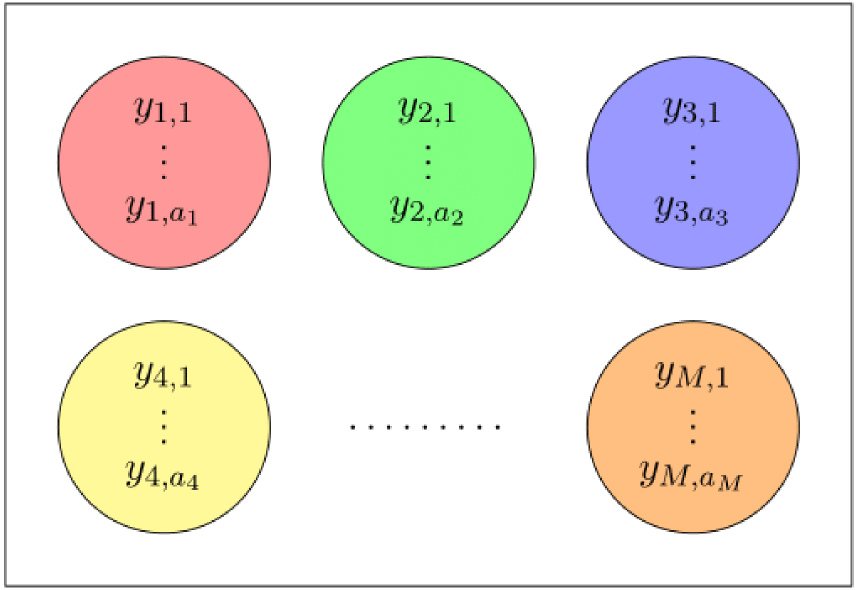
- Partition $N$ elements into $M$ colors:




- $X$: sample $q$ elements without replacement
- $Y$: sample $q$ elements with replacement

# Stam's Bound [Sta78]

- Partition $N$ elements into $M$ colors:

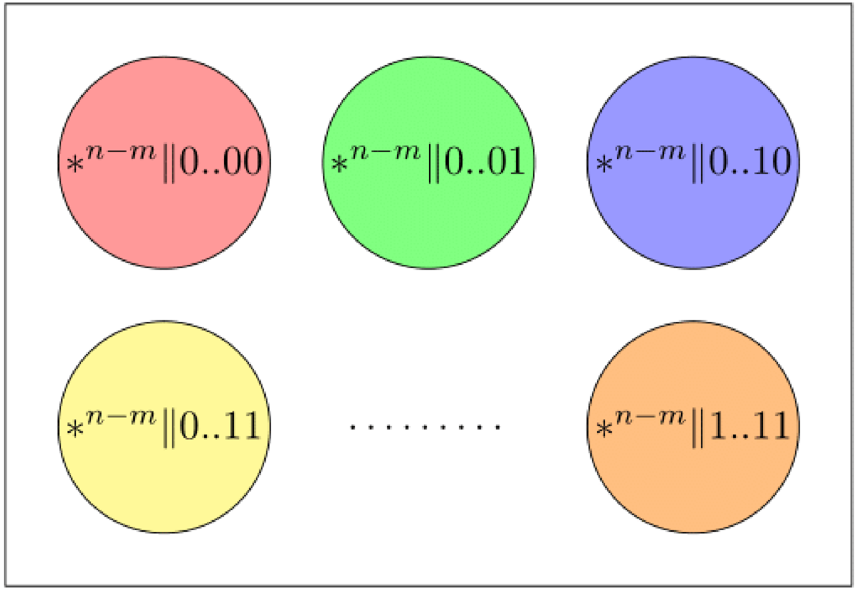



- $X$: sample $q$ elements without replacement
- $Y$: sample $q$ elements with replacement
- Stam [Sta78] (simplified): $\Delta(X,Y) \leq \frac{Mq^2}{N^2}$

# Applying Stam's Bounds to Plain Truncation

$$x \xrightarrow[n]{} \boxed{p} \xrightarrow[n]{} \boxed{\text{trunc}} \xrightarrow[m]{} y$$

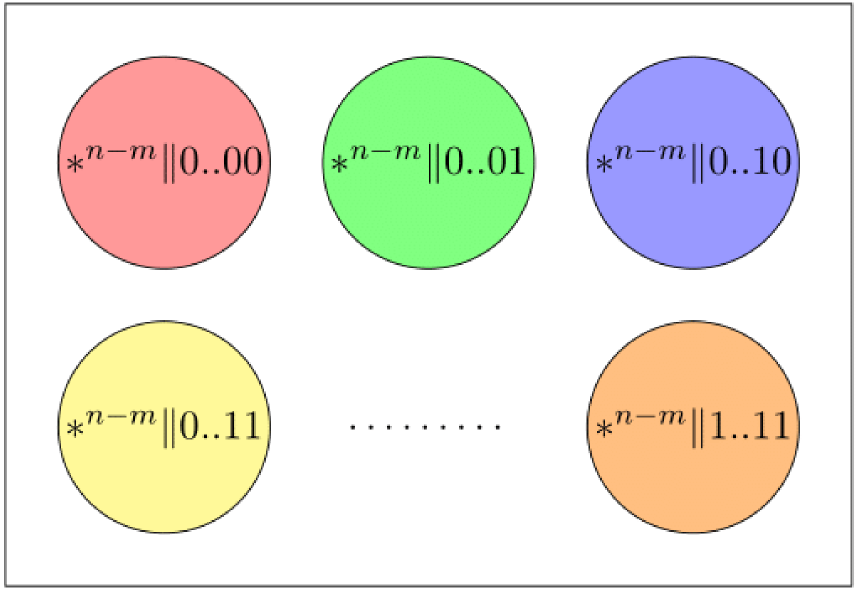
- Partition $2^n$ elements into $2^m$ colors:



# Applying Stam's Bounds to Plain Truncation

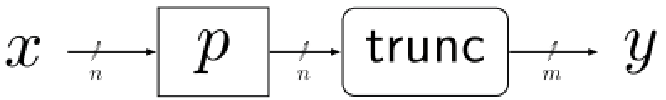
$x \xrightarrow[n]{} \boxed{p} \xrightarrow[n]{} \boxed{\text{trunc}} \xrightarrow[m]{} y$

- Partition $2^n$ elements into $2^m$ colors:

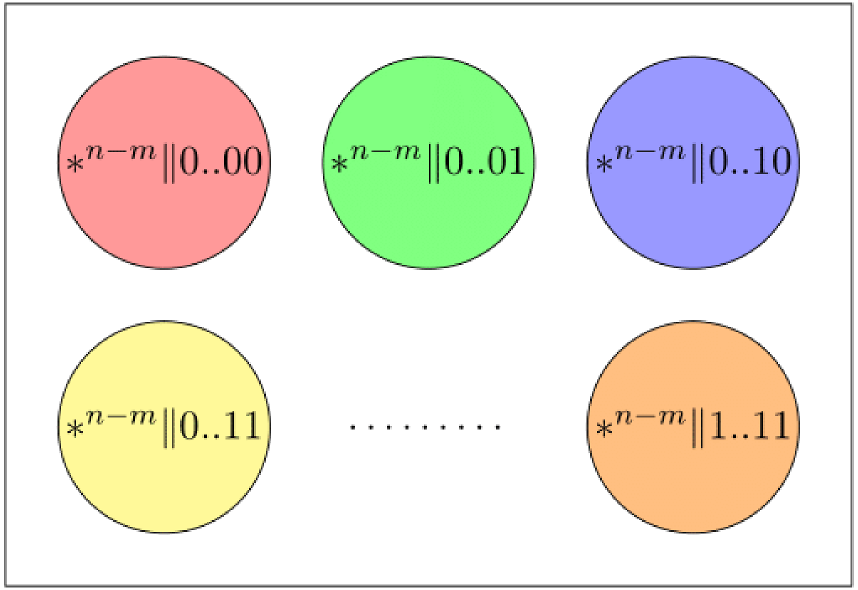



- Truncated permutation $\equiv$ sampling without replacement

# Applying Stam's Bounds to Plain Truncation

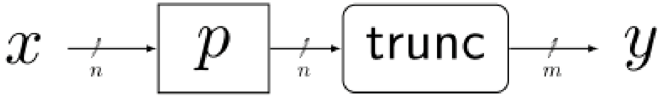



- Partition $2^n$ elements into $2^m$ colors:

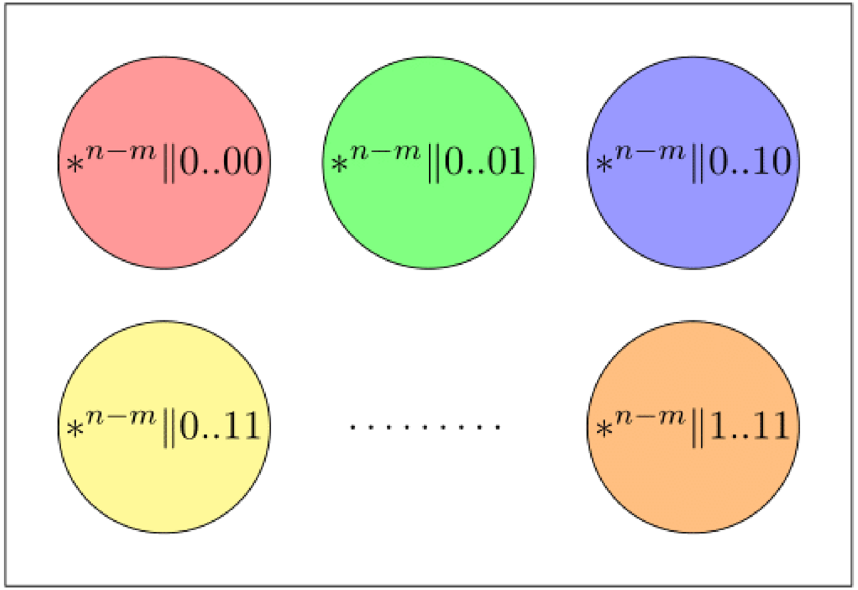



- Truncated permutation $\equiv$ sampling without replacement
- Random function $\equiv$ sampling with replacement

# Applying Stam's Bounds to Plain Truncation




- Partition $2^n$ elements into $2^m$ colors:

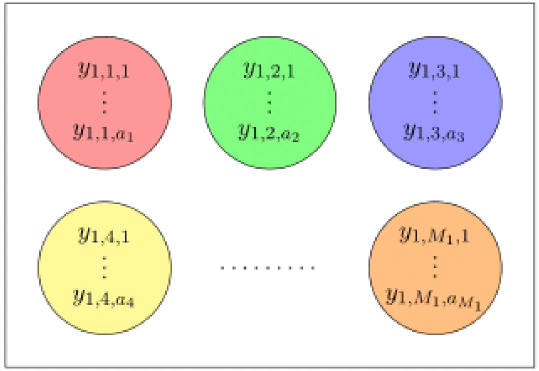



- Truncated permutation $\equiv$ sampling without replacement
- Random function $\equiv$ sampling with replacement
- Truncated permutation is $2^{n-m/2}$ PRF secure

# Generalized Stam's Bound [Sta86]

- Possibly different partitions for the $q$ drawings:

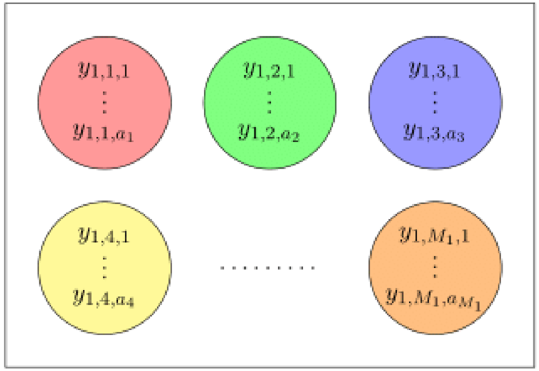
drawing $1$



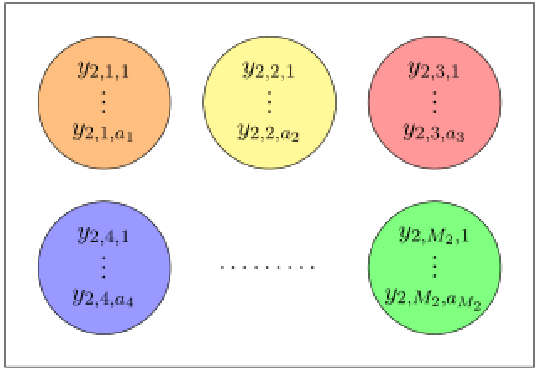
# Generalized Stam's Bound [Sta86]

- Possibly different partitions for the $q$ drawings:

drawing 1

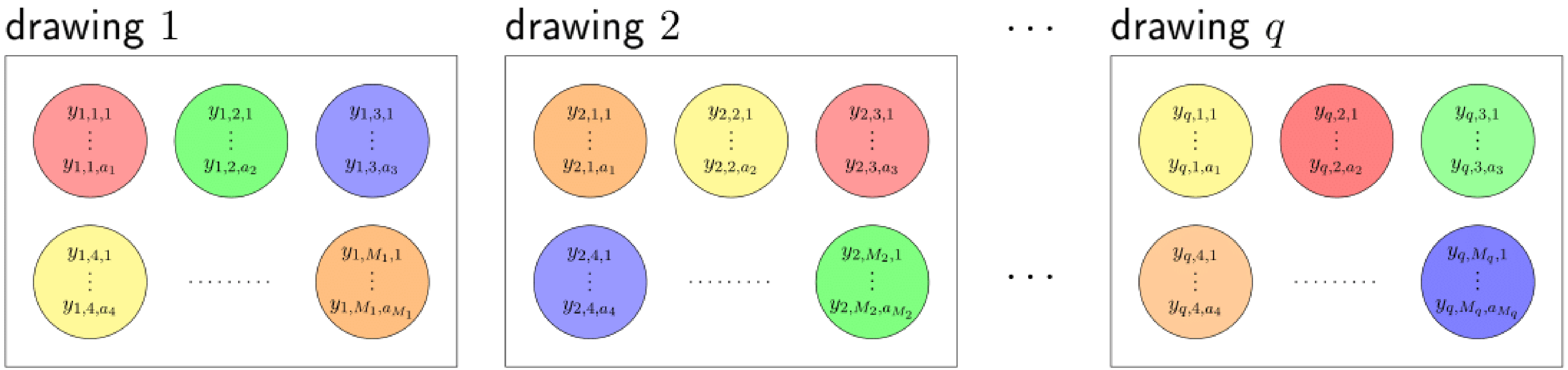



drawing 2



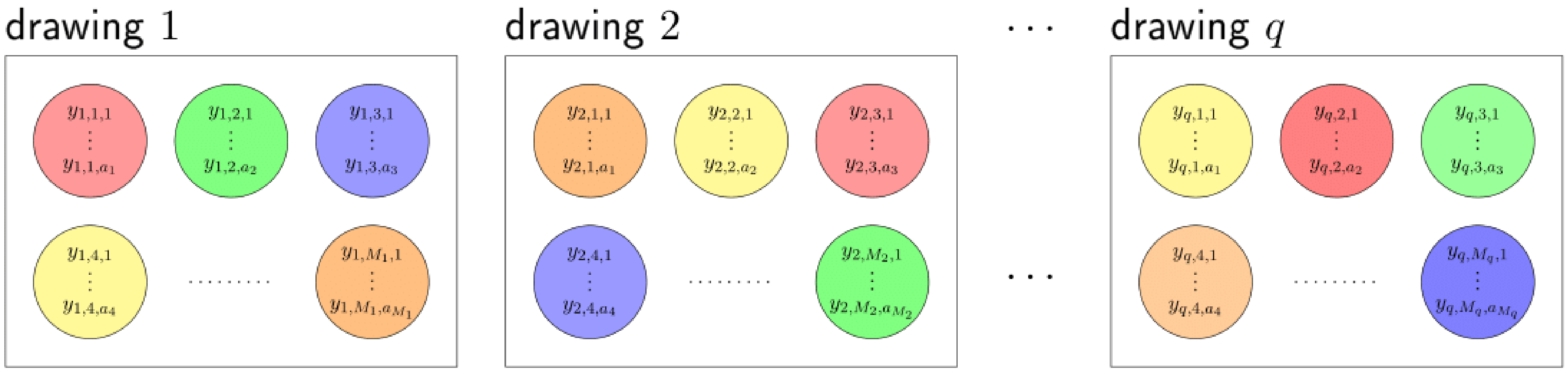
# Generalized Stam's Bound [Sta86]

- Possibly different partitions for the $q$ drawings:



# Generalized Stam's Bound [Sta86]

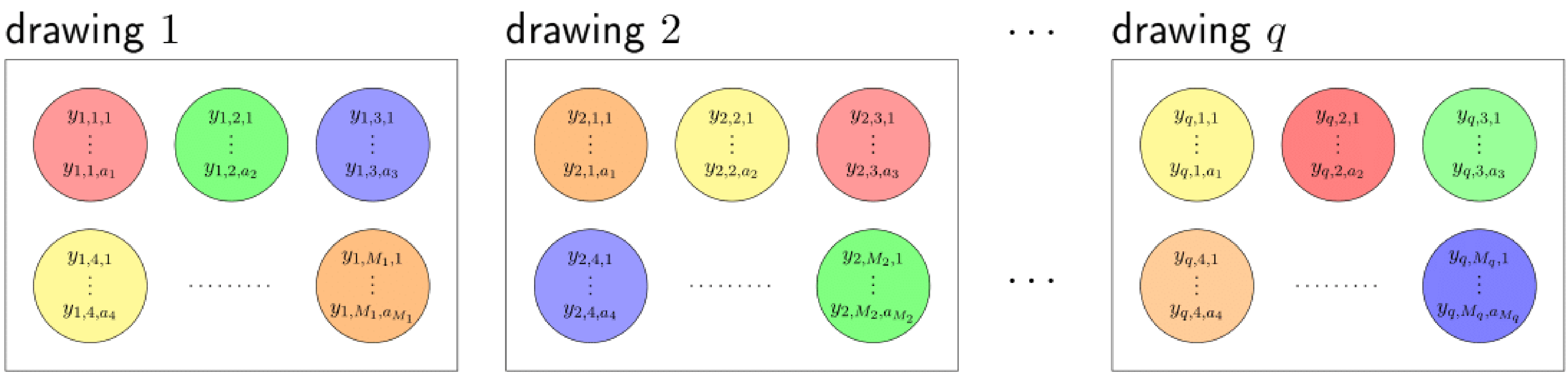
- Possibly different partitions for the $q$ drawings:




- $X$: sample $q$ elements without replacement
- $Y$: sample $q$ elements with replacement

# Generalized Stam's Bound [Sta86]

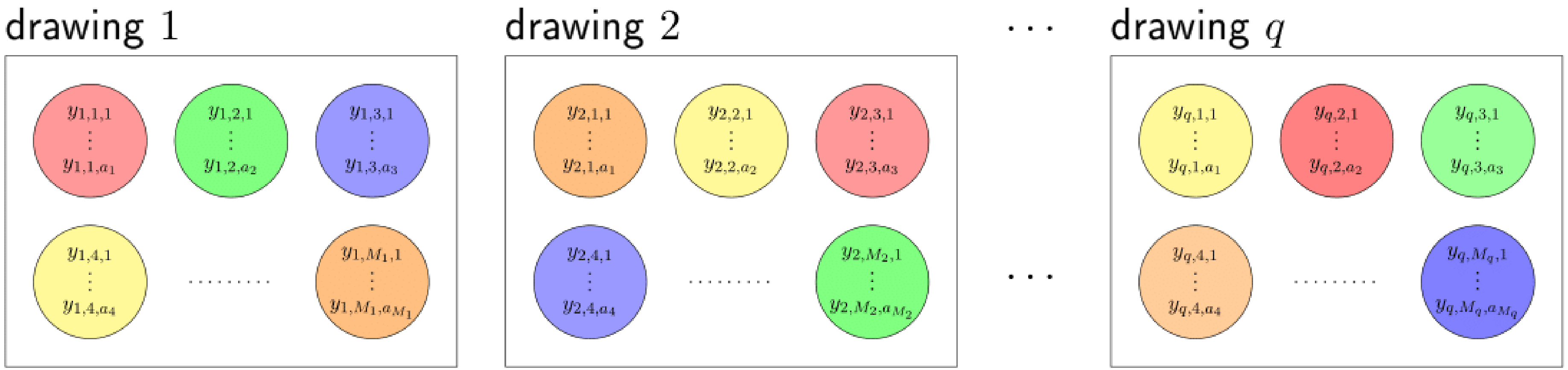
- Possibly different partitions for the $q$ drawings:




- $X$: sample $q$ elements without replacement
- $Y$: sample $q$ elements with replacement
- Stam [Sta86] (simplified): $\Delta(X, Y) \leq \sum_{i=1}^{q} \frac{M_i i}{N^2}$

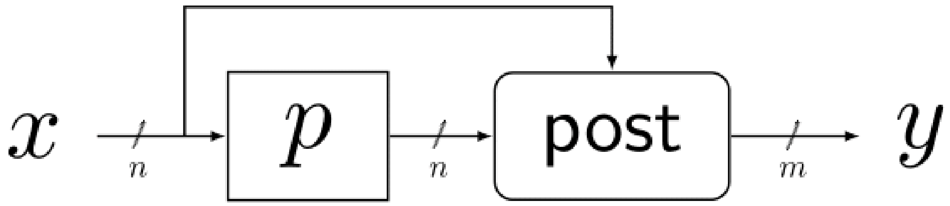
# Generalized Stam's Bound [Sta86]

- Possibly different partitions for the $q$ drawings:



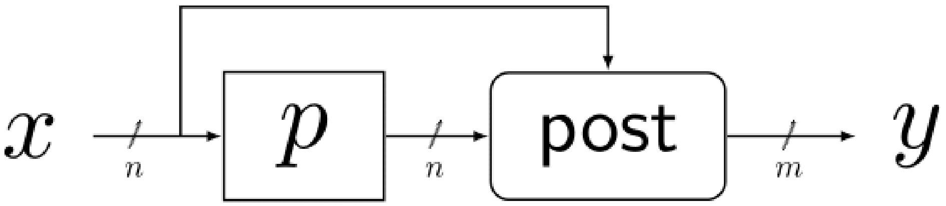

- $X$: sample $q$ elements without replacement
- $Y$: sample $q$ elements with replacement
- Stam [Sta86] (simplified): $\Delta(X,Y) \leq \sum_{i=1}^{q} \frac{M_i i}{N^2}$
- If $M_1 = \cdots = M_q = M$, then $\Delta(X,Y) \leq \frac{Mq^2}{N^2}$

# Generalized Truncation




- We translate Stam's bounds to generalized truncation
- Understand and transform proof techniques to PRF security

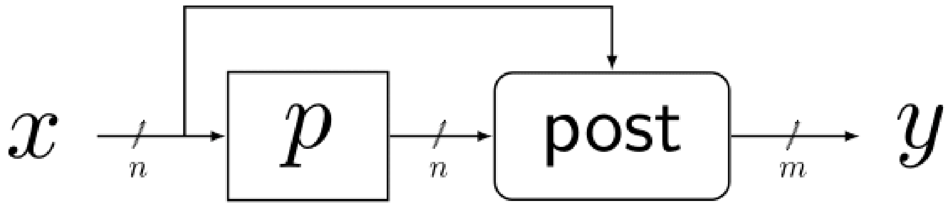
# Generalized Truncation




- We translate Stam's bounds to generalized truncation
- Understand and transform proof techniques to PRF security

| condition on post | PRF security | note |
|---|---|---|
| plain truncation | $2^{n-m/2}$ | known result, based on [Sta78] |

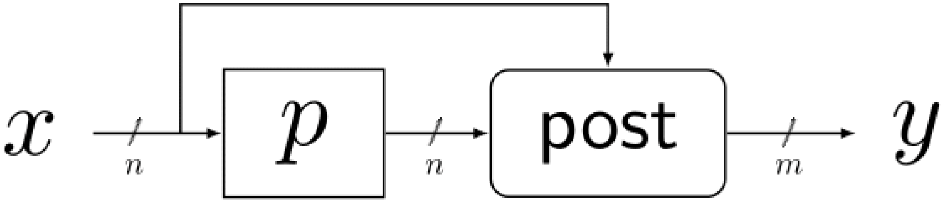
# Generalized Truncation




- We translate Stam's bounds to generalized truncation
- Understand and transform proof techniques to PRF security

| condition on post | PRF security | note |
|---|---|---|
| plain truncation | $2^{n-m/2}$ | known result, based on [Sta78] |
| balanced and $x$-independent | $2^{n-m/2}$ | equivalent, also based on [Sta78] |

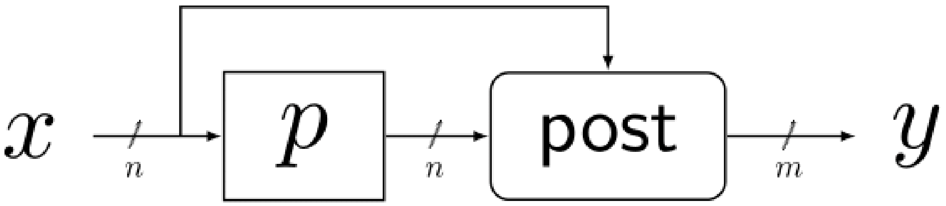
# Generalized Truncation




- We translate Stam's bounds to generalized truncation
- Understand and transform proof techniques to PRF security

| condition on post | PRF security | note |
|---|---|---|
| plain truncation | $2^{n-m/2}$ | known result, based on [Sta78] |
| balanced and $x$-independent | $2^{n-m/2}$ | equivalent, also based on [Sta78] |
| balanced | $2^{n-m/2}$ | same bound, but based on [Sta86] |

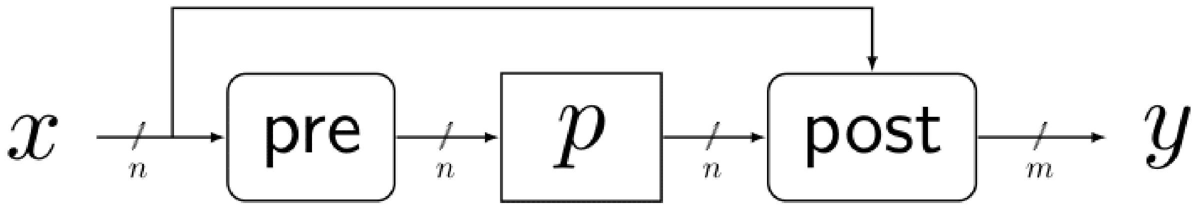
# Generalized Truncation



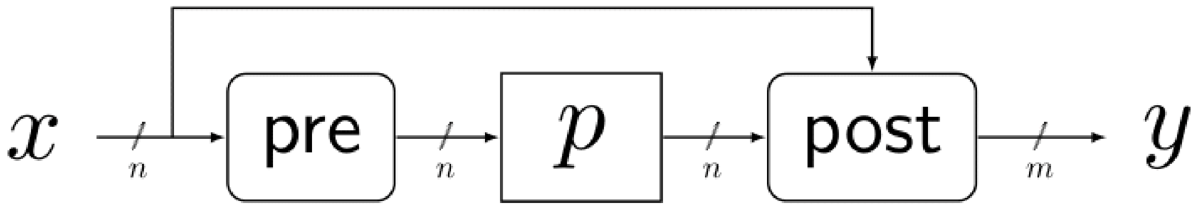

- We translate Stam's bounds to generalized truncation
- Understand and transform proof techniques to PRF security

| condition on post | PRF security | note |
|---|---|---|
| plain truncation | $2^{n-m/2}$ | known result, based on [Sta78] |
| balanced and $x$-independent | $2^{n-m/2}$ | equivalent, also based on [Sta78] |
| balanced | $2^{n-m/2}$ | same bound, but based on [Sta86] |
| arbitrary | $\min\{2^{n-m/2}, 2^{2n-2m}/\gamma^2\}$ | $\gamma$ quantifies unbalancedness: $\lvert \mathsf{post}^{-1}[x](y) - 2^{n-m} \rvert \leq \gamma$, extra proof step needed |

# What About Pre-Processing?

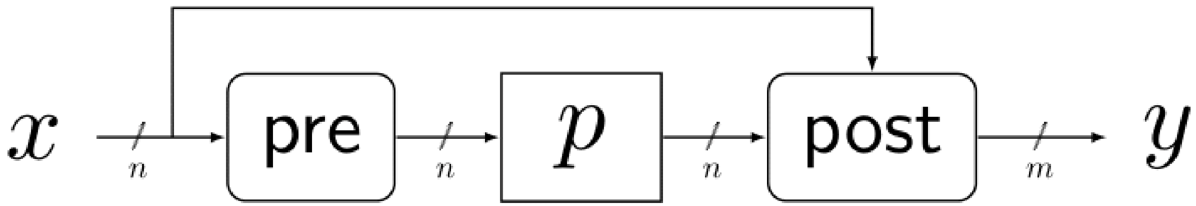


# What About Pre-Processing?



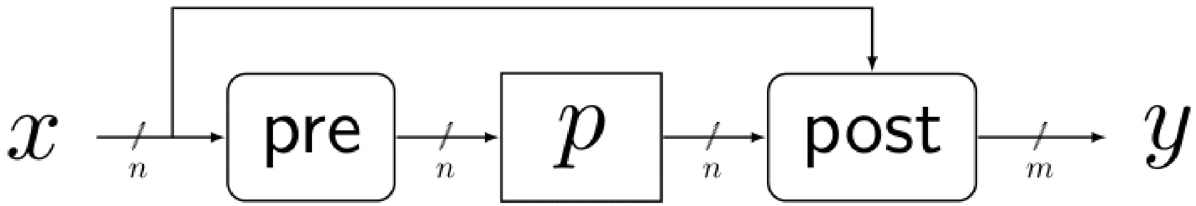

- Balanced pre?
  $\rightarrow$ pre could be absorbed into $p$

# What About Pre-Processing?




- Balanced pre?
  $\rightarrow$ pre could be absorbed into $p$
- Unbalanced pre?
  $\rightarrow$ pre allows for collisions
  $\rightarrow$ output bias notable for $m \approx n$

# What About Pre-Processing?




- Balanced pre?
  $\rightarrow$ pre could be absorbed into $p$
- Unbalanced pre?
  $\rightarrow$ pre allows for collisions
  $\rightarrow$ output bias notable for $m \approx n$
- Pre-processing may only degrade security

# Conclusion

**Truncation**

- Recently popularized by GCM-SIV (but usage disputed [IS17,BHT18])
- Security already covered by ancient result
- Generalized truncation: detailed security treatment

**Application**

- Simple truncation is best option
- Advantage over XoP?

**Thank you for your attention!**