

Computationally Differentially Private Inner Product Protocols Imply Oblivious Transfer

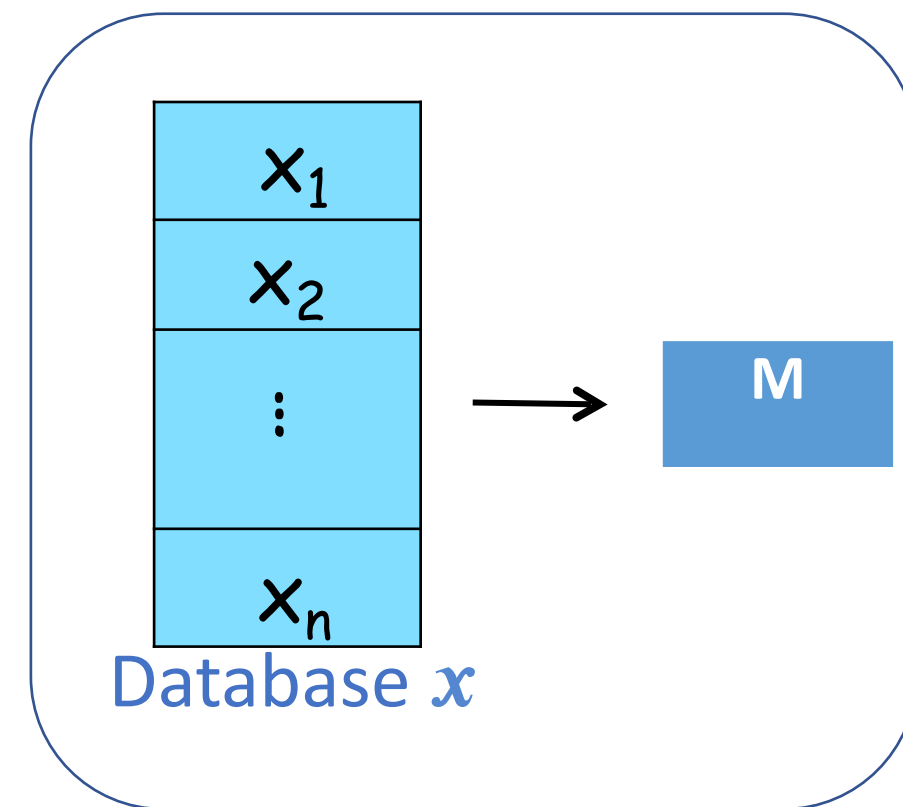
Chao Yan

Joint work with

Iftach Haitner, Noam Mazon, Jad Silbak and Eliad Tsfadia

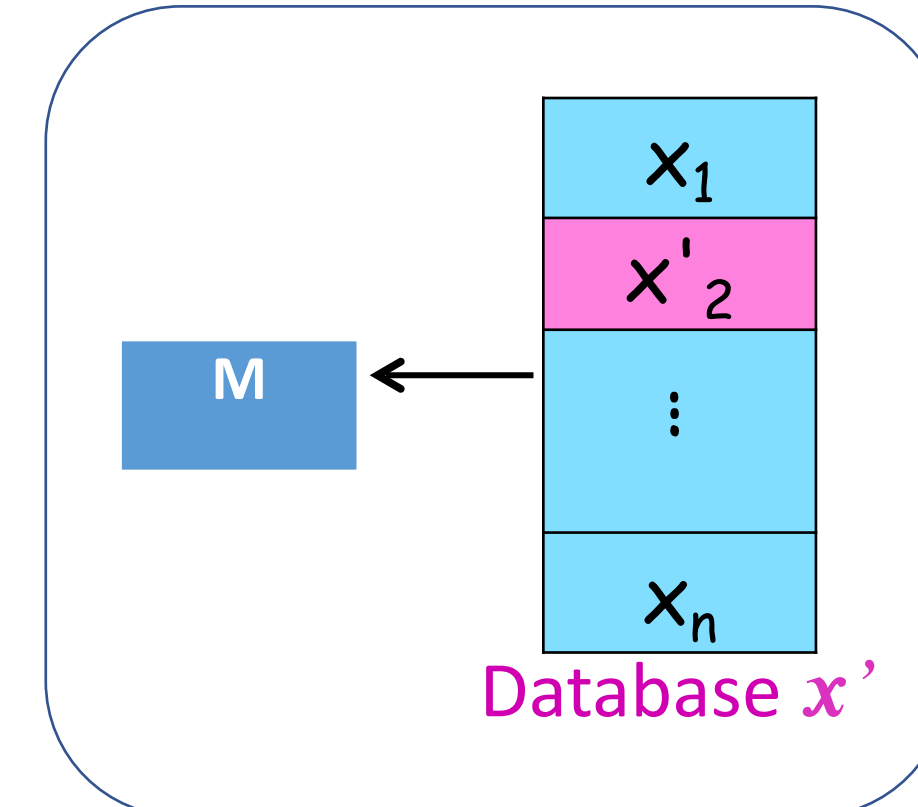
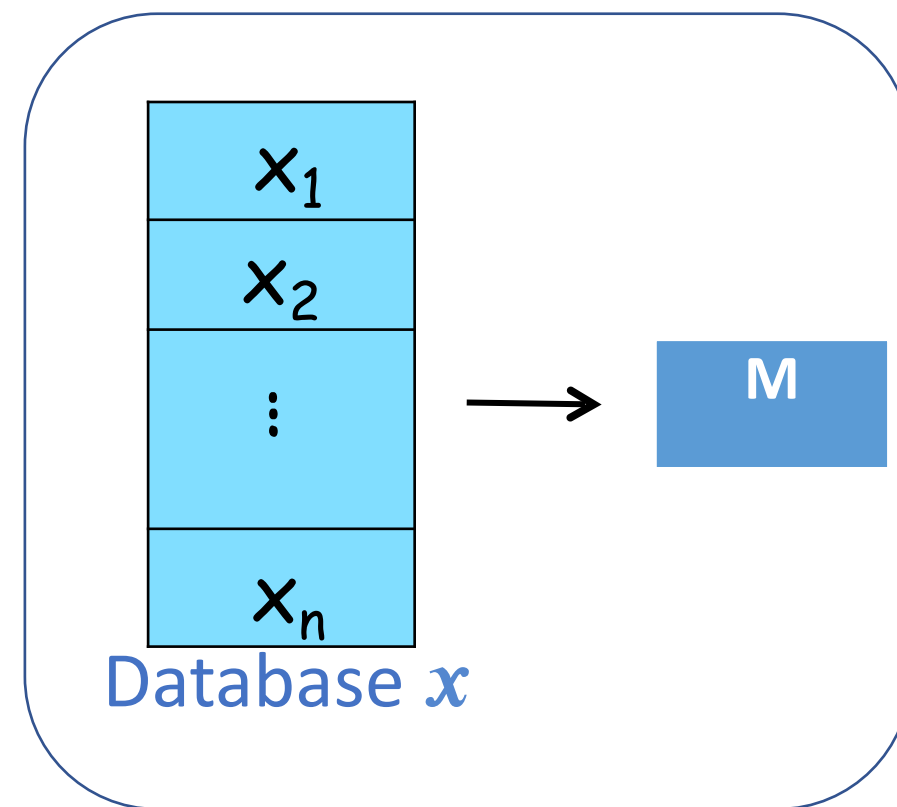
Differential Privacy

Dwork, McSherry, Nissim, Smith 2006



Differential Privacy

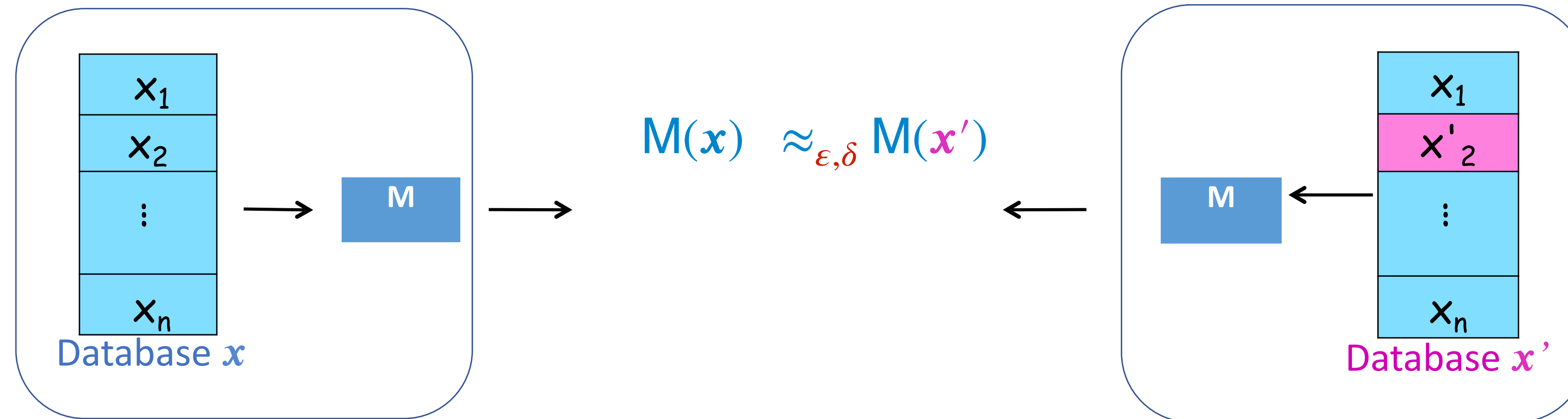
Dwork, McSherry, Nissim, Smith 2006



Differential Privacy

Dwork, McSherry, Nissim, Smith 2006

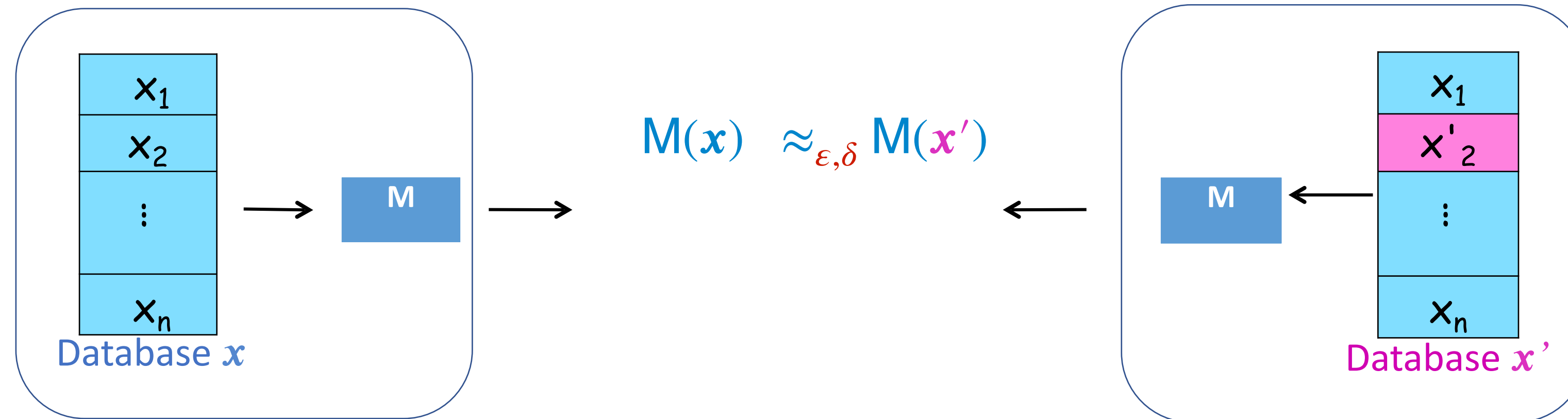
One record does not change the output distribution “too much”



Differential Privacy

Dwork, McSherry, Nissim, Smith 2006

One record does not change the output distribution “too much”



M is (ϵ, δ) -DP

$\forall$ neighboring databases x, x' and $\forall$ (unbounded) distinguisher D :

$$\Pr[D(M(x)) = 1] \leq e^{\epsilon} \cdot \Pr[D(M(x')) = 1] + \delta$$

Centralized DP

Google

$$x = (x_1, \dots, x_n)$$



$$f(x) = \sum_{i=1}^n x_i$$

$$M(x) = \sum_{i=1}^n x_i + \textit{noise}$$

Centralized DP

Google

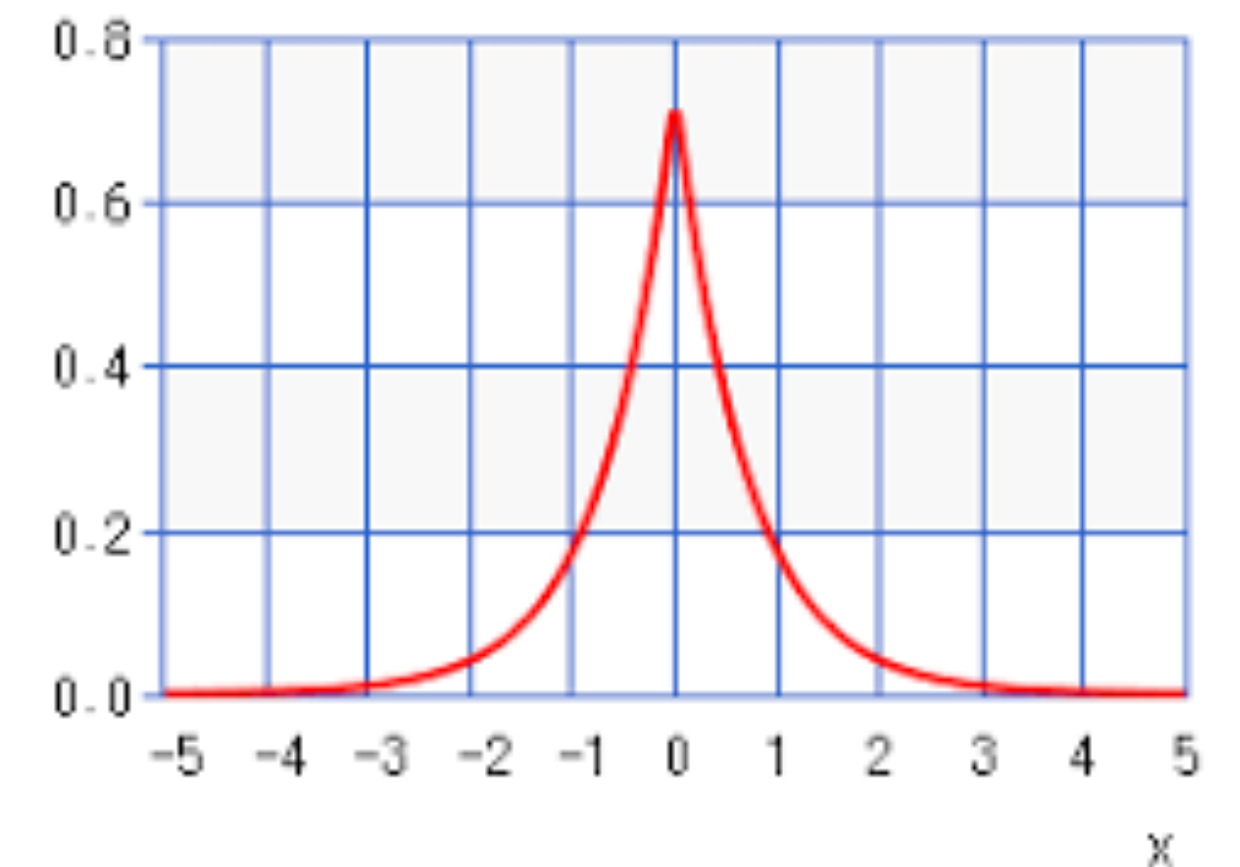
$$x = (x_1, \dots, x_n)$$



$$f(x) = \sum_{i=1}^n x_i$$

$$M(x) = \sum_{i=1}^n x_i + \mathbf{Laplace}(1/\epsilon)$$

w.h.p $\text{noise} \leq O(1)$



Centralized DP



$$x = (x_1, \dots, x_n)$$

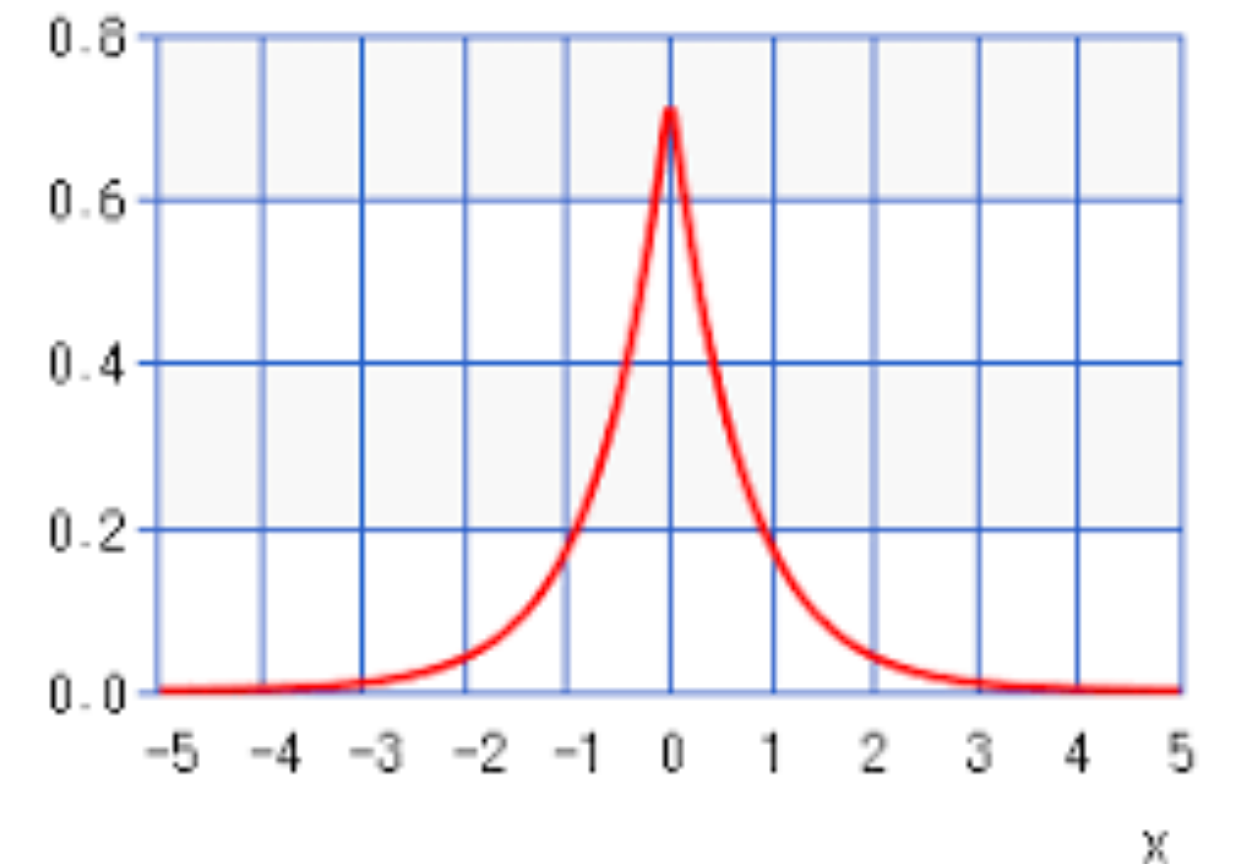


$$f(x) = \sum_{i=1}^n x_i$$

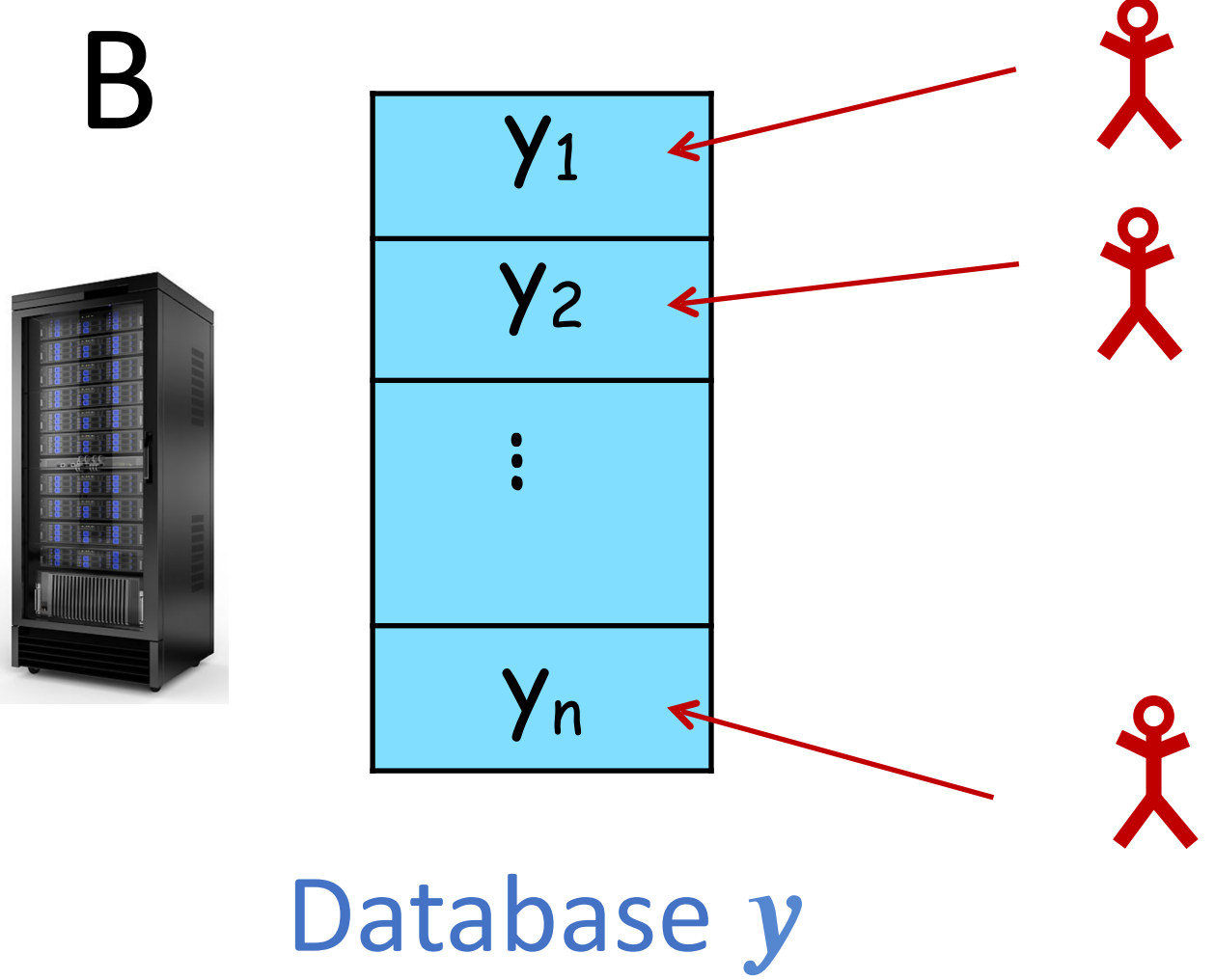
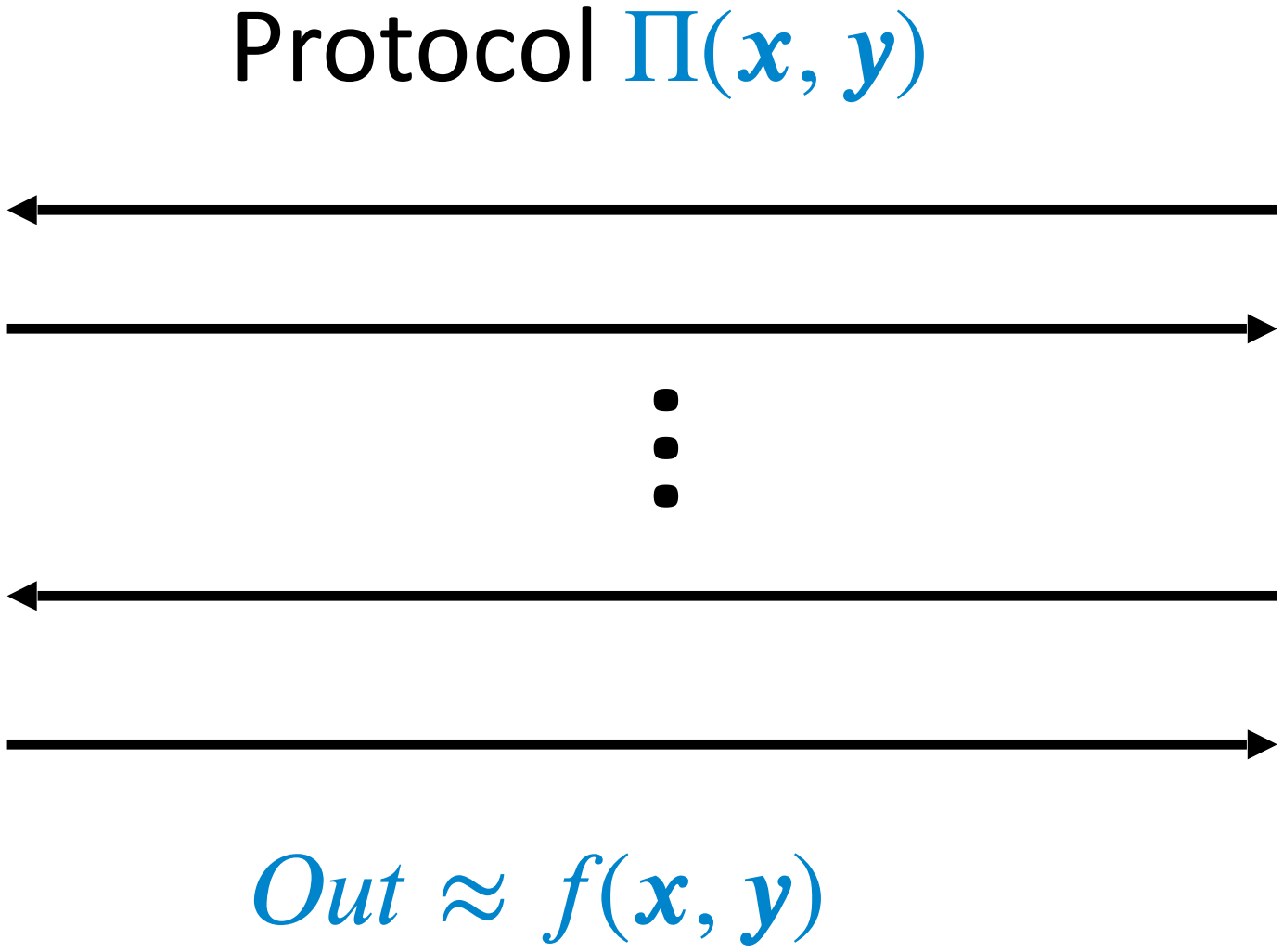
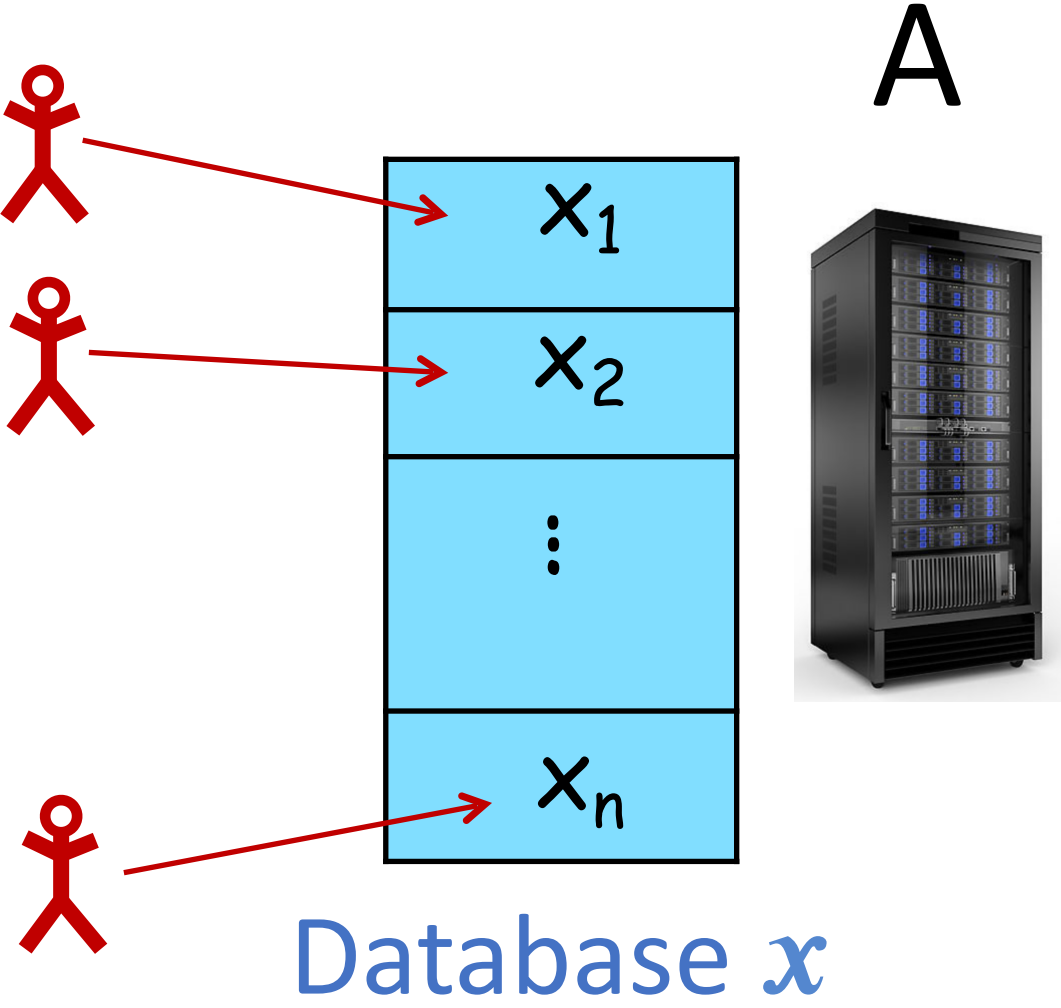
$$M(x) = \sum_{i=1}^n x_i + \mathbf{Laplace}(1/\epsilon)$$

centralized-DP: M can
access the **entire**
database x

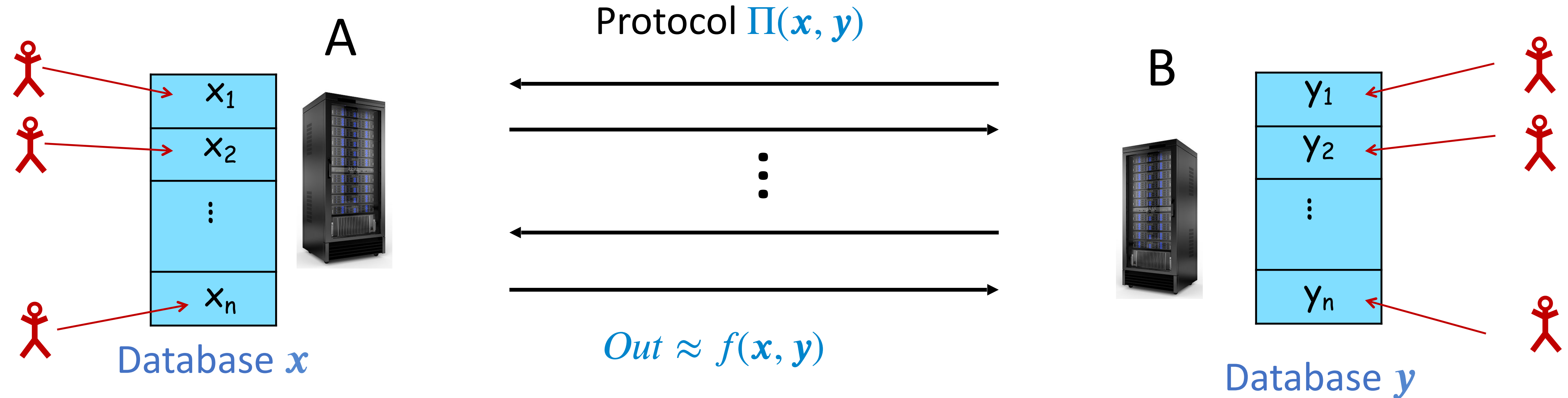
$$\mathbf{w.h.p} \text{ noise} \leq O(1)$$



Two-Party DP



Two-Party DP



Goal: Estimate $f(x, y)$ while preserving (ϵ, δ) -DP:

$$\forall x, \forall \text{ neigh. } y, y': \text{view}_A^\Pi(x, y) \approx_{\epsilon, \delta} \text{view}_A^\Pi(x, y')$$

$\text{view}_A^\Pi(x, y)$ – A's view in $\Pi(x, y)$ (input, coins and transcript).

(and same for B)

Two-Party DP

Google

$$x = (x_1, \dots, x_n)$$

$$out_1 = \sum_i x_i + Noise$$

$$out_2 = \sum_i y_i + Noise$$

$$out = out_1 + out_2$$



$$y = (y_1, \dots, y_n)$$



Two-Party DP

Google

$x = (x_1, \dots, x_n)$

$$out_1 = \sum_i x_i + Noise$$

$$out_2 = \sum_i y_i + Noise$$

$$out = out_1 + out_2$$



$y = (y_1, \dots, y_n)$



w.h.p $noise \leq O(1)$

Inner Product

A

$$x \in \{-1, 1\}^n$$



?

B

$$y \in \{-1, 1\}^n$$



$$\langle x, y \rangle = \sum_{i=1}^n x_i y_i - \text{measures correlation between databases}$$

DP Inner Product



DP Inner Product

Centralized Model

Can achieve **constant** error.

DP Inner Product

Centralized Model

Can achieve **constant** error.

Two-Party Protocol

For uniform inputs:

$$\overset{A}{x} \in \{-1, 1\}^n$$

$$\overset{B}{y} \in \{-1, 1\}^n$$

$$out = 0$$

- With prob. 0.99: $|out - \langle x, y \rangle| \approx \sqrt{n}$
- Can be generalized for every input distribution.

DP Inner Product

Centralized Model

Can achieve **constant** error.

Two-Party Protocol

For uniform inputs:

$$\overset{A}{x} \in \{-1, 1\}^n$$

$$\overset{B}{y} \in \{-1, 1\}^n$$

$$out = 0$$

- With prob. 0.99: $|out - \langle x, y \rangle| \approx \sqrt{n}$
- Can be generalized for every input distribution.

Lower Bound

McGregor, Mironov, Pitassi, Reingold, Talwar and Vadhan 2010

For every DP protocol:

$$|out - \langle x, y \rangle| \approx \sqrt{n}$$

DP Inner Product

Centralized Model

Can achieve **constant** error.

Two-Party Protocol

For uniform inputs:

$$\overset{A}{x} \in \{-1, 1\}^n$$

$$\overset{B}{y} \in \{-1, 1\}^n$$

$$out = 0$$

- With prob. 0.99: $|out - \langle x, y \rangle| \approx \sqrt{n}$
- Can be generalized for every input distribution.

Computational Two-Party DP

Can achieve **constant** error, using
Cryptography

Lower Bound

McGregor, Mironov, Pitassi, Reingold, Talwar and Vadhan 2010

For every DP protocol:

$$|out - \langle x, y \rangle| \approx \sqrt{n}$$

DP Inner Product

Centralized Model

Can achieve **constant** error.

Two-Party Protocol

For uniform inputs:

$$\overset{A}{x} \in \{-1, 1\}^n$$

$$\overset{B}{y} \in \{-1, 1\}^n$$

$$out = 0$$

- With prob. 0.99: $|out - \langle x, y \rangle| \approx \sqrt{n}$
- Can be generalized for every input distribution.

Computational Two-Party DP

Can achieve **constant** error, using
Cryptography

Oblivious Transfer and
Secure MPC

Lower Bound

McGregor, Mironov, Pitassi, Reingold, Talwar and Vadhan 2010

For every DP protocol:

$$|out - \langle x, y \rangle| \approx \sqrt{n}$$

Computational DP

Beimel, Nissim, Omri 2008 Mironov, Pandey, Reingold, Vadhan 2009

M is (ϵ, δ) -DP

$\forall$ neighboring databases x, x' and $\forall$ (unbounded) distinguisher D :

$$\Pr[D(M(x)) = 1] \leq e^\epsilon \cdot \Pr[D(M(x')) = 1] + \delta$$

Computational DP

Beimel, Nissim, Omri 2008 Mironov, Pandey, Reingold, Vadhan 2009

M is (ϵ, δ) -DP

$\forall$ neighboring databases x, x' and $\forall$ (unbounded) distinguisher D :

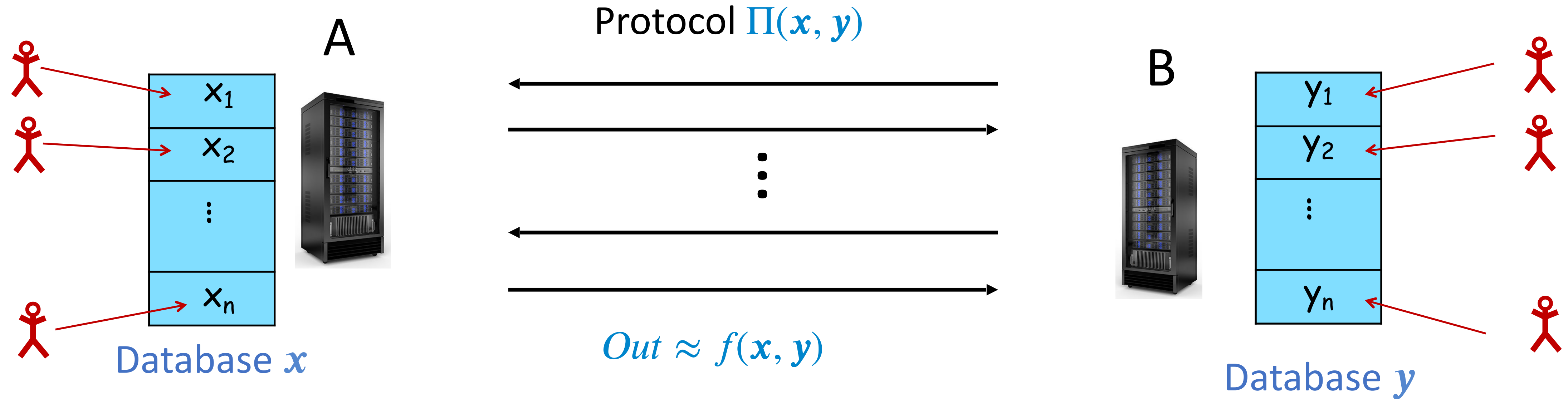
$$\Pr[D(M(x)) = 1] \leq e^\epsilon \cdot \Pr[D(M(x')) = 1] + \delta$$

M is (ϵ, δ) -CDP

$\forall$ neighboring databases x, x' and $\forall$ efficient (PPT) distinguisher D :

$$\Pr[D(M(x)) = 1] \leq e^\epsilon \cdot \Pr[D(M(x')) = 1] + \delta$$

Two-Party CDP



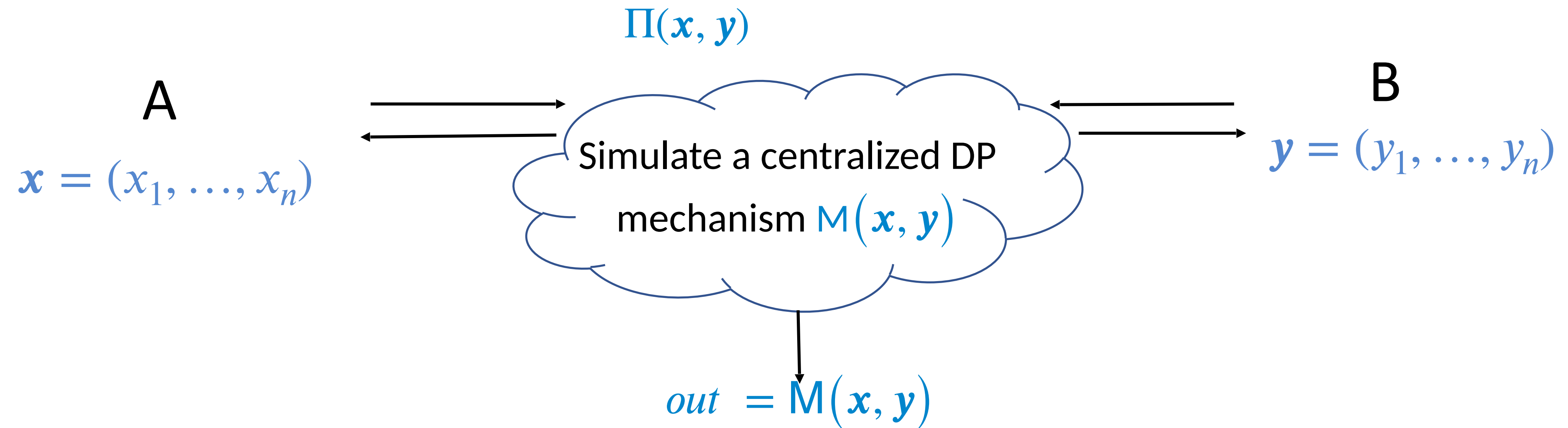
Goal: Estimate $f(x, y)$ while preserving (ϵ, δ) -CDP:

$$\forall x, \forall \text{ neigh. } y, y': \text{view}_A^\Pi(x, y) \approx_{\epsilon, \delta}^c \text{view}_A^\Pi(x, y')$$

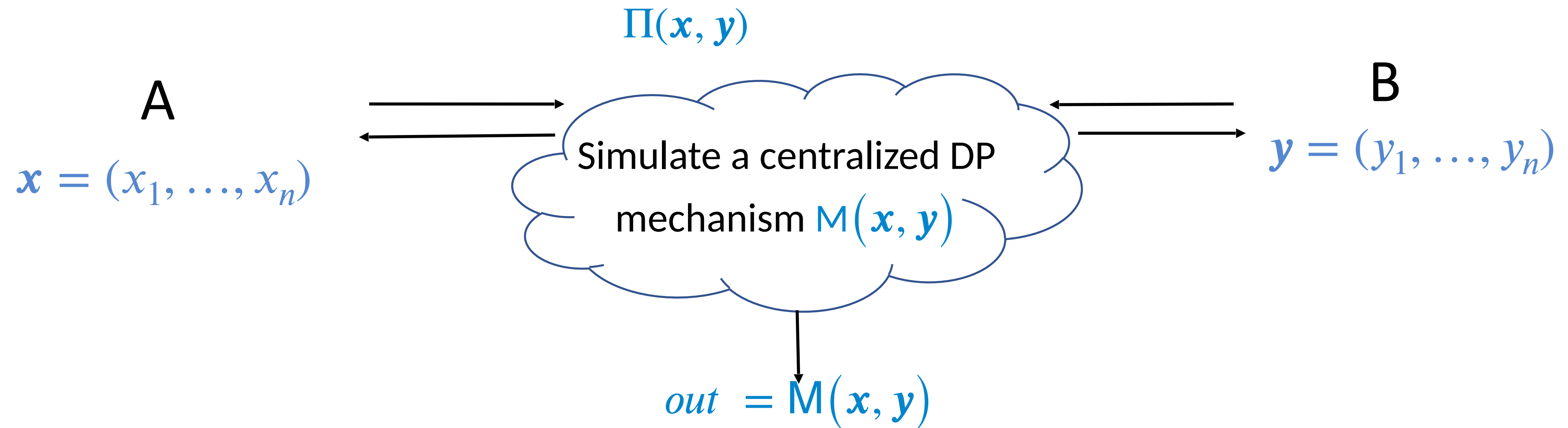
$\text{view}_A^\Pi(x, y)$ – A's view in $\Pi(x, y)$ (input, coins and transcript).

(and same for B)

CDP via Secure Multiparty Computation

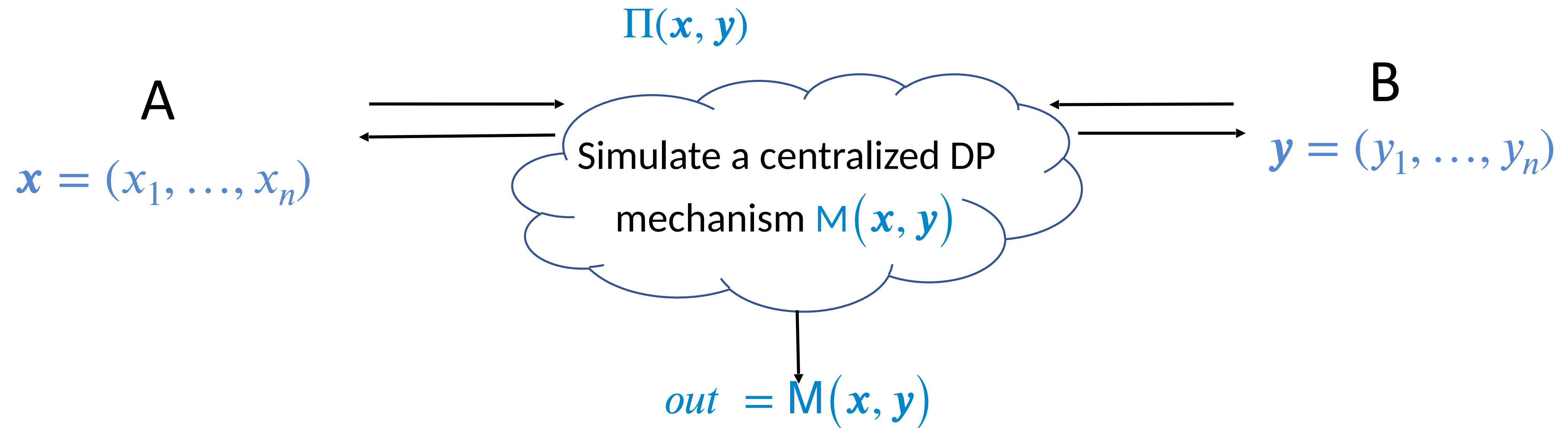


CDP via Secure Multiparty Computation



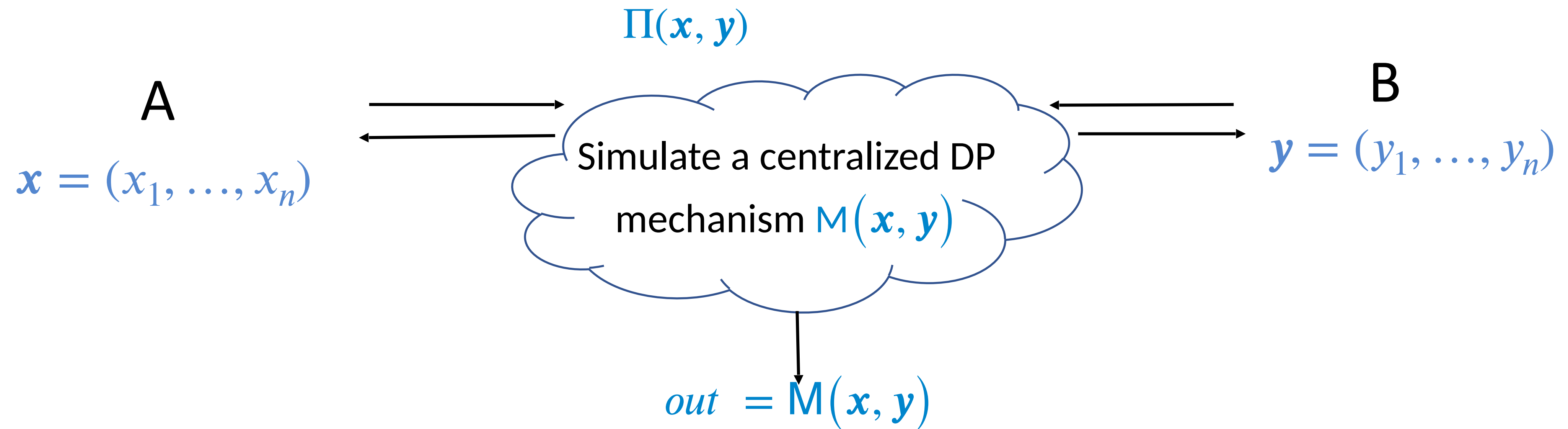
- M is (centralized) (ϵ, δ) -DP $\implies \Pi$ is (ϵ, δ) -CDP.

CDP via Secure Multiparty Computation



- M is (centralized) (ϵ, δ) -DP $\implies \Pi$ is (ϵ, δ) -CDP.
- Secure MPC via *Oblivious Transfer* (OT).

CDP via Secure Multiparty Computation



- M is (centralized) (ϵ, δ) -DP $\implies \Pi$ is (ϵ, δ) -CDP.
- Secure MPC via *Oblivious Transfer* (OT).
- For computing IP, take $M(x, y) = \langle x, y \rangle + Lap(2/\epsilon)$.

The Complexity of Two-Party CDP

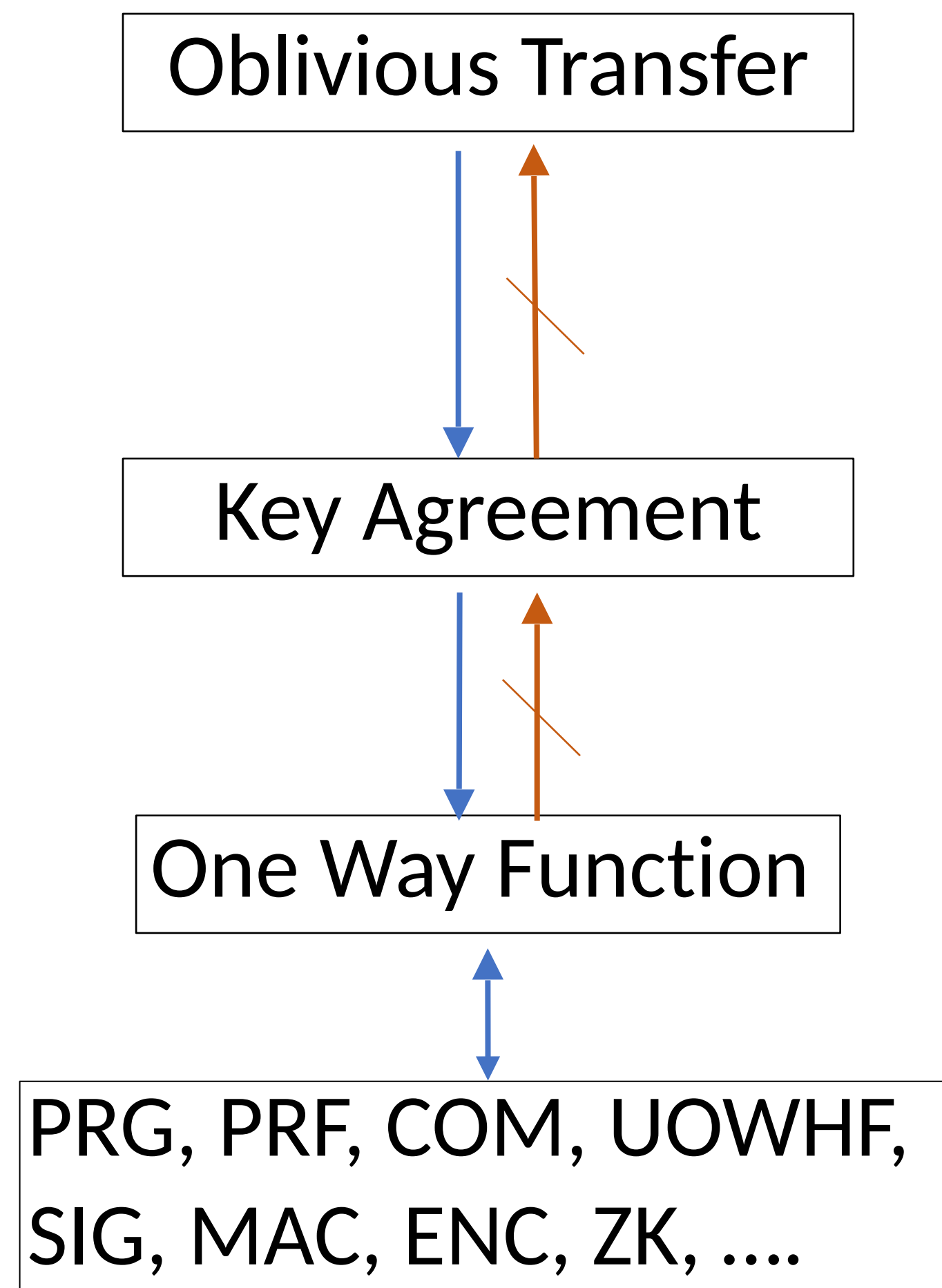
Using OT, we can construct very accurate CDP protocols!

The Complexity of Two-Party CDP

Using OT, we can construct very accurate CDP protocols!

Do we have to use OT?

Complexity Hierarchy

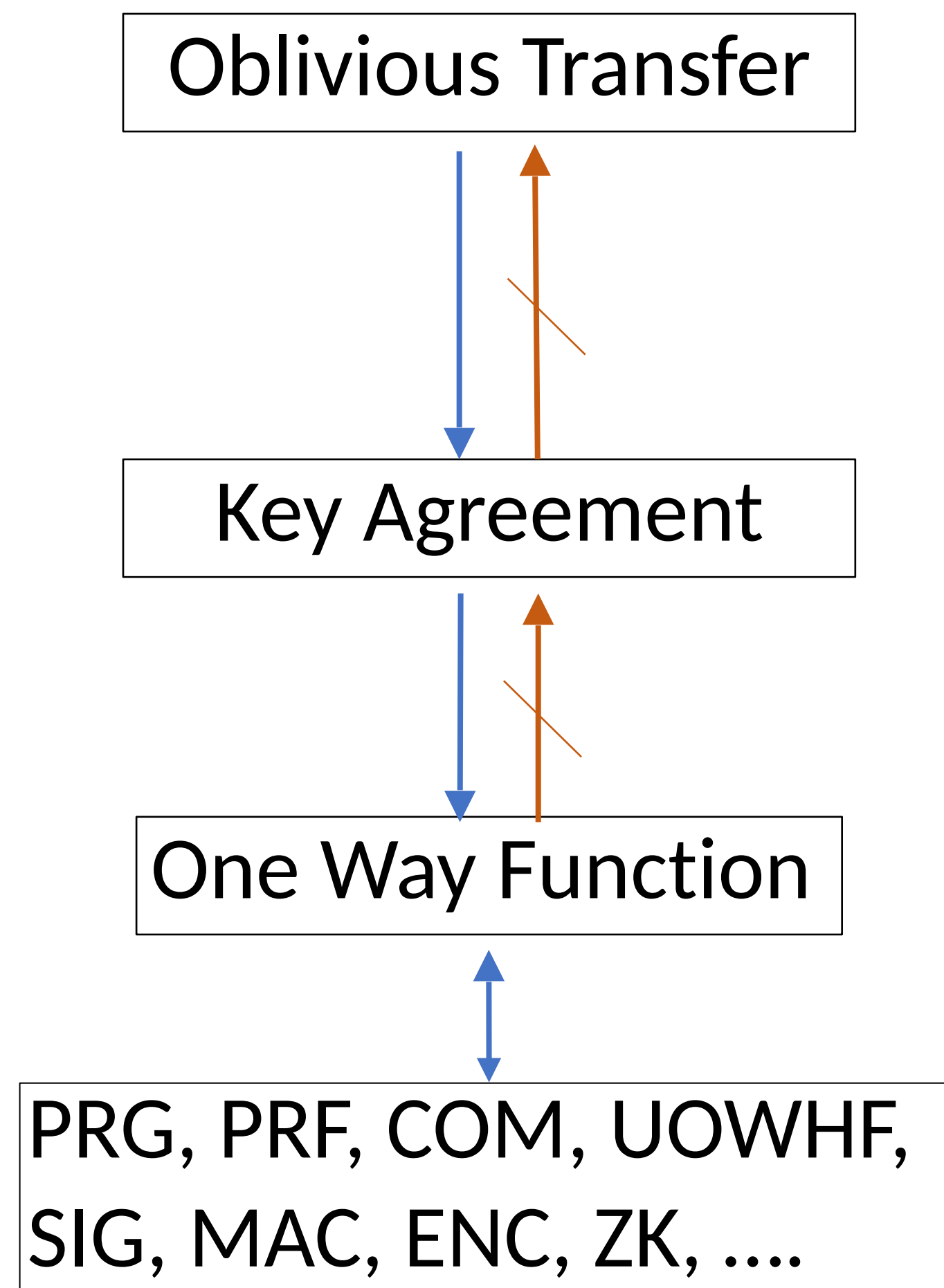


Complexity Hierarchy

“Non-trivial”:

Possible in two-party CDP

Impossible in two-party DP

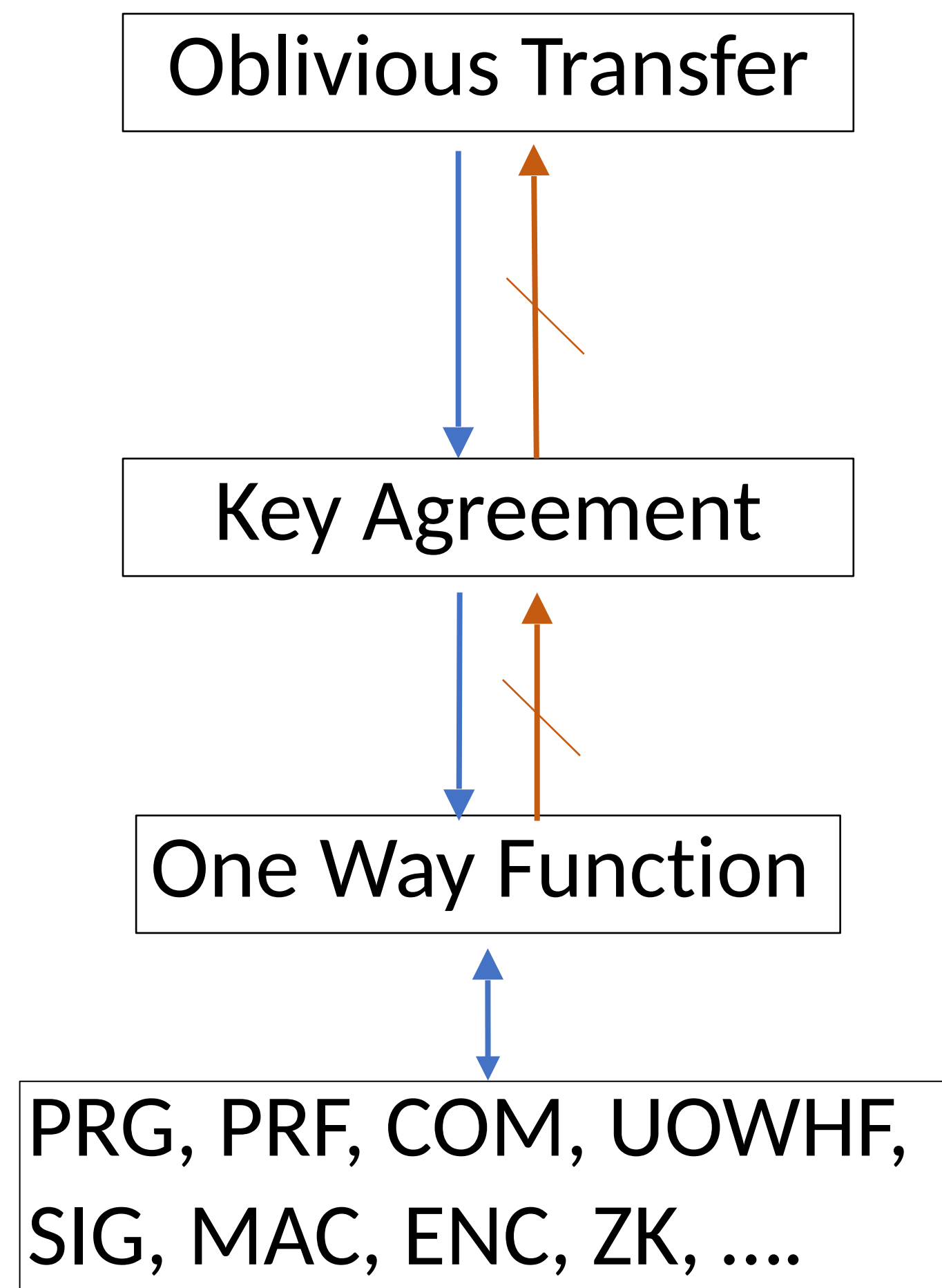


Complexity Hierarchy

“Non-trivial”:

Possible in two-party CDP

Impossible in two-party DP



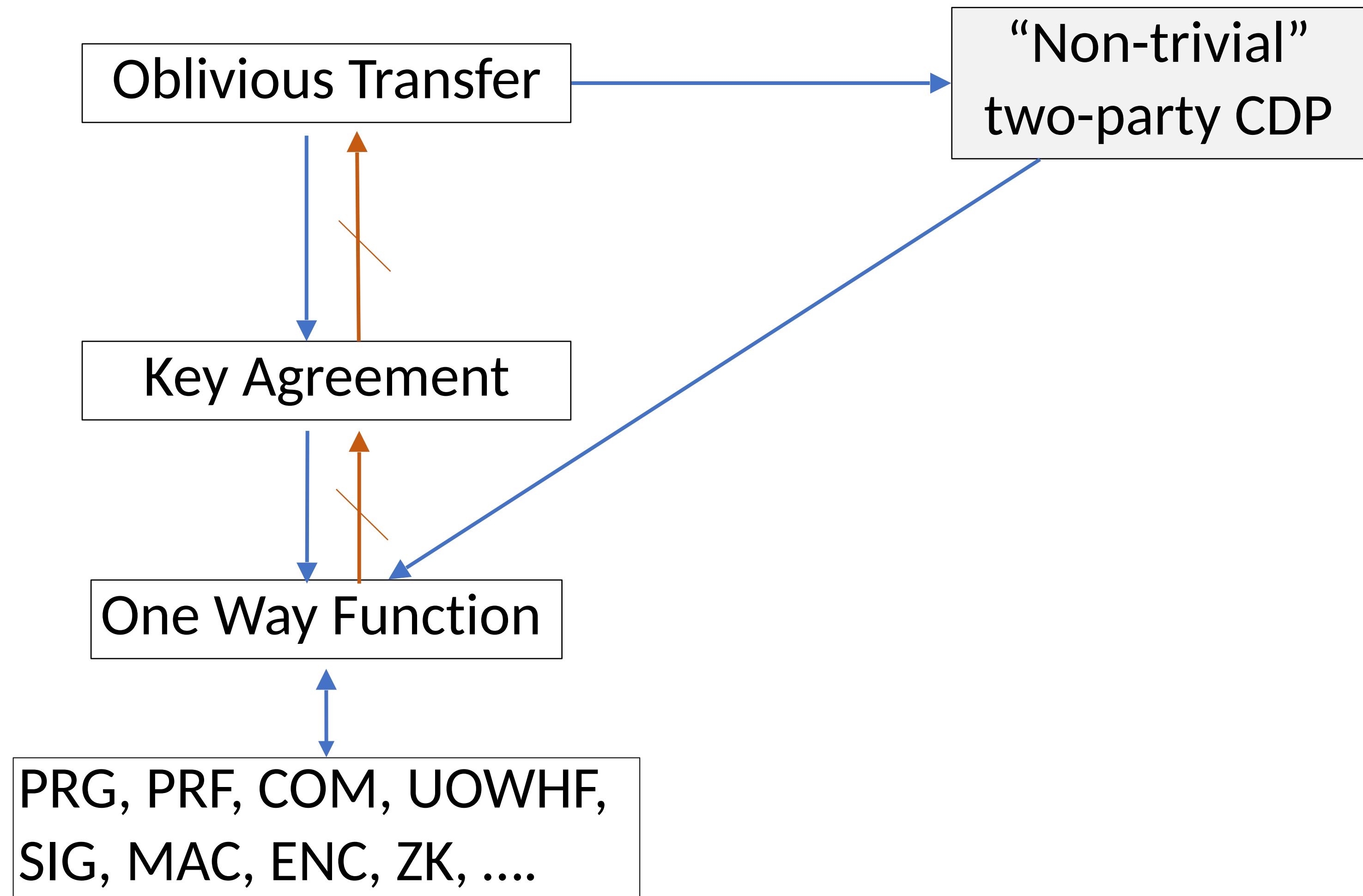
“Non-trivial”
two-party CDP

Complexity Hierarchy

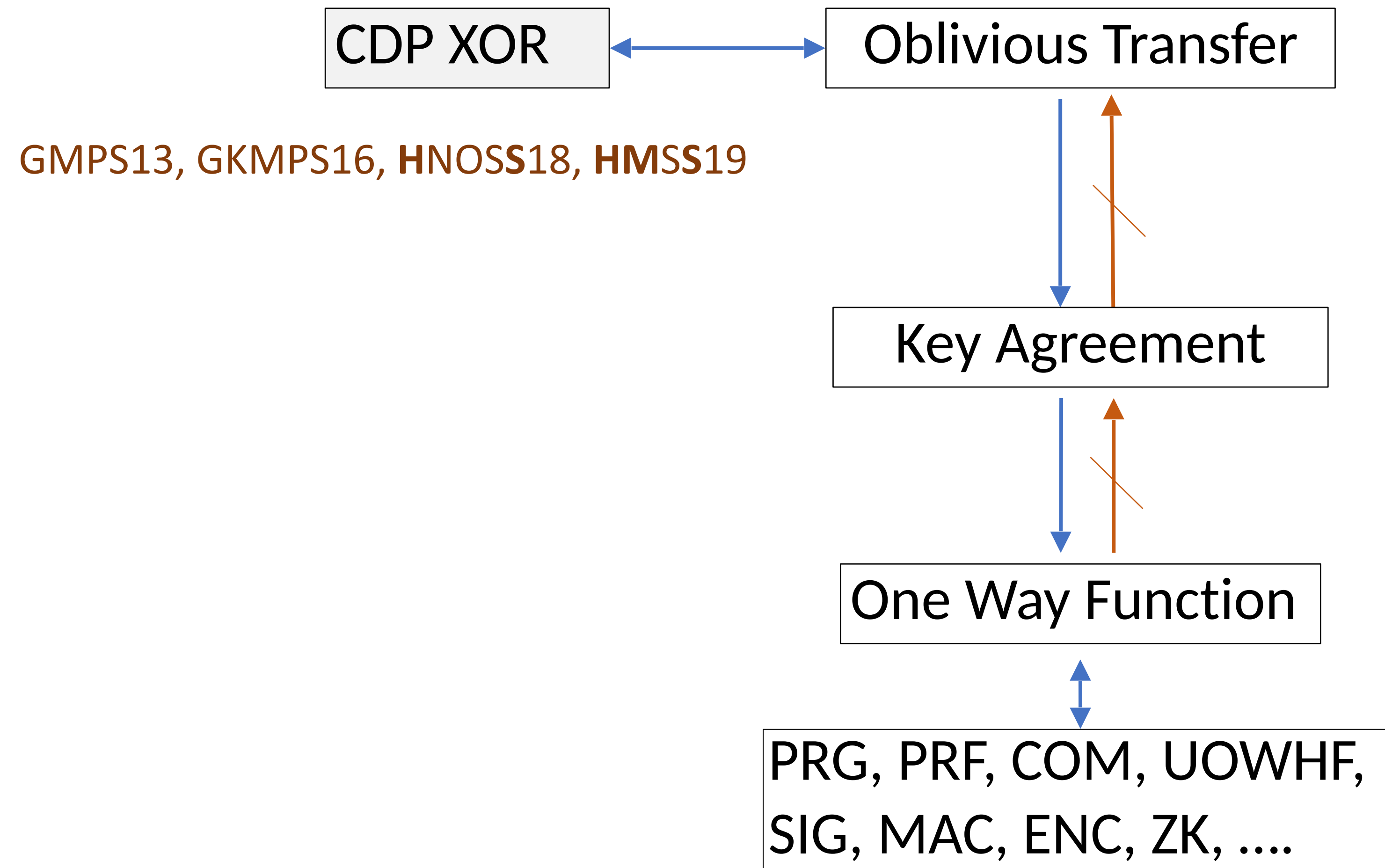
“Non-trivial”:

Possible in two-party CDP

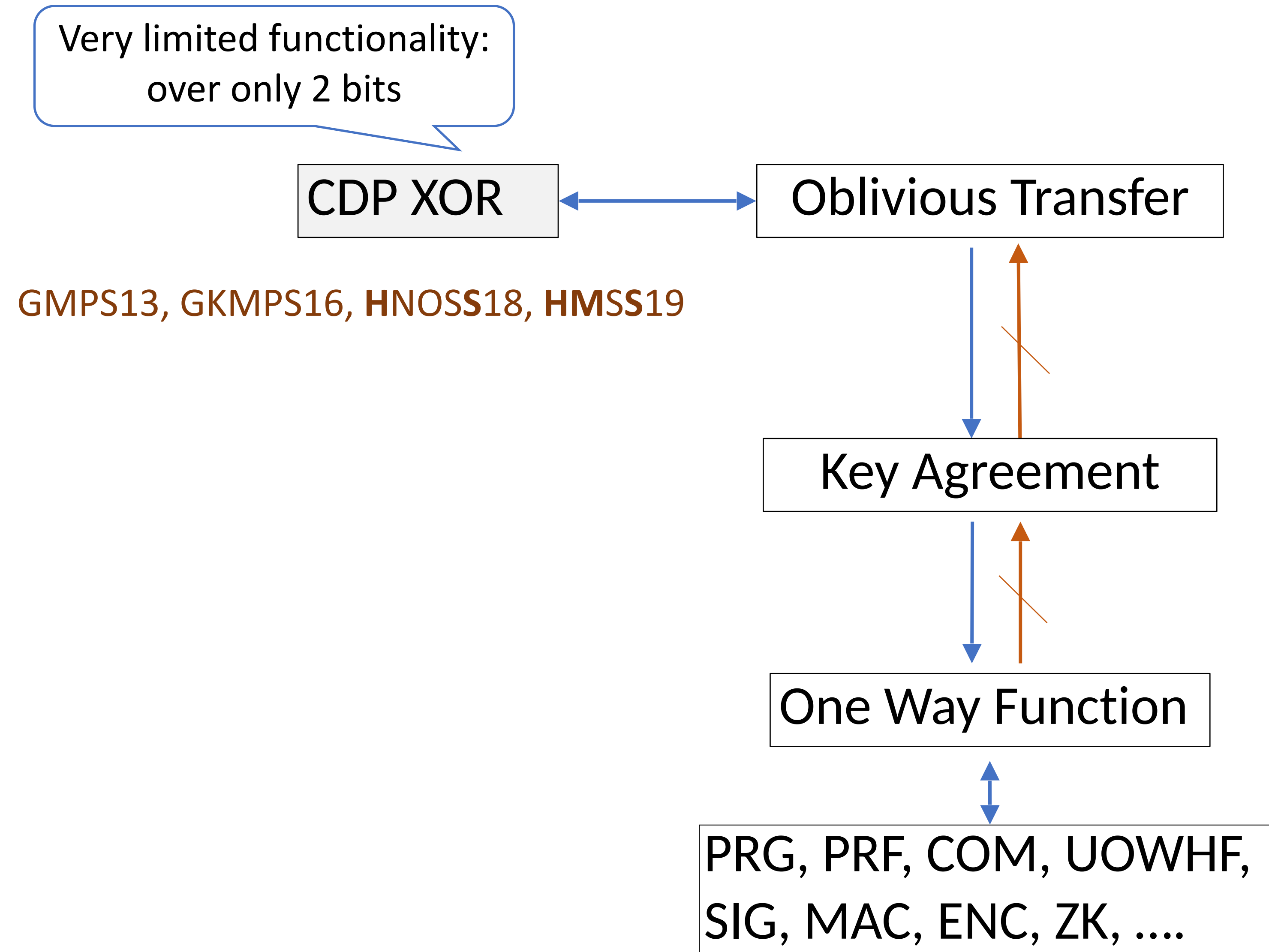
Impossible in two-party DP



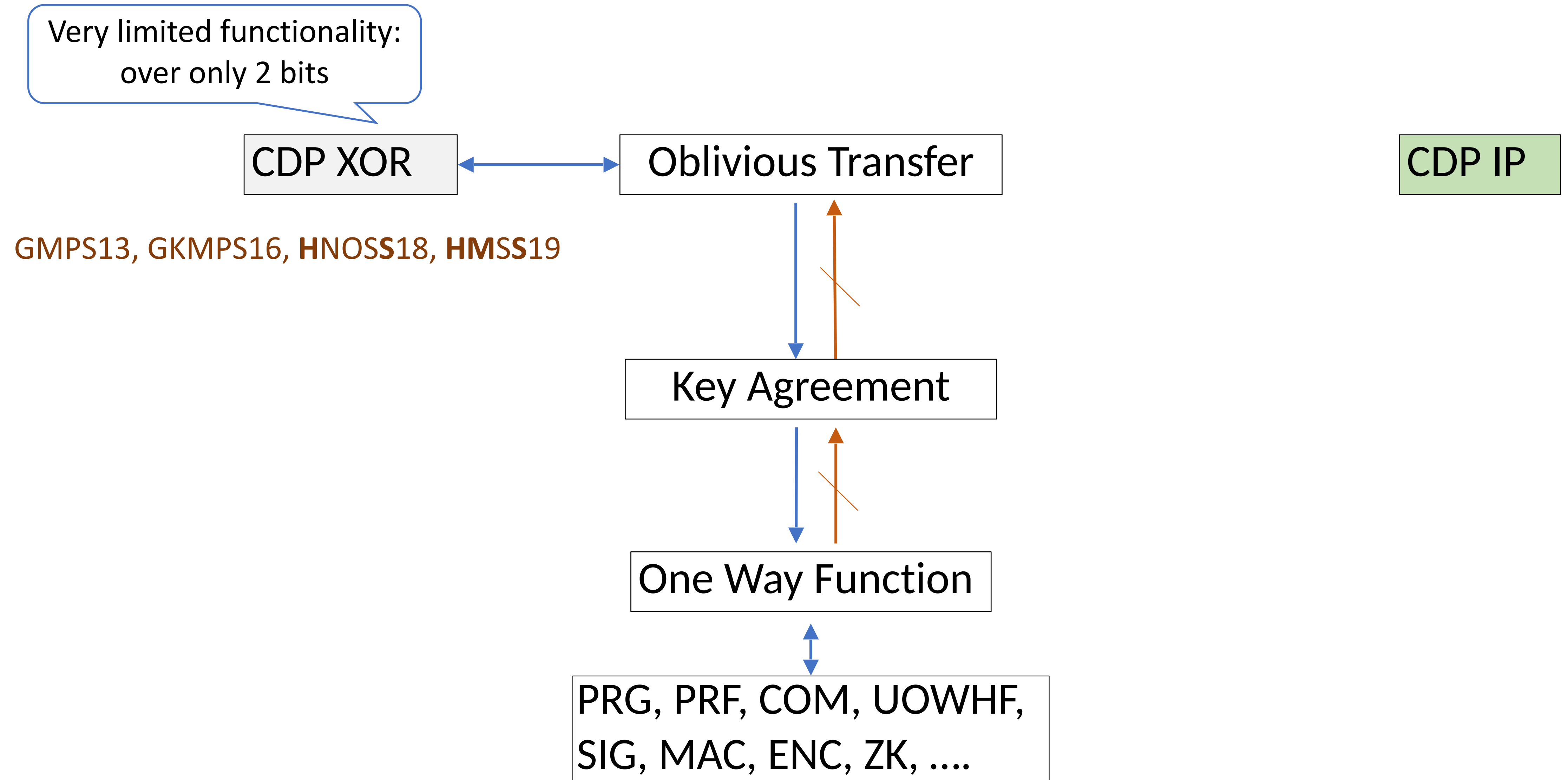
Complexity Hierarchy



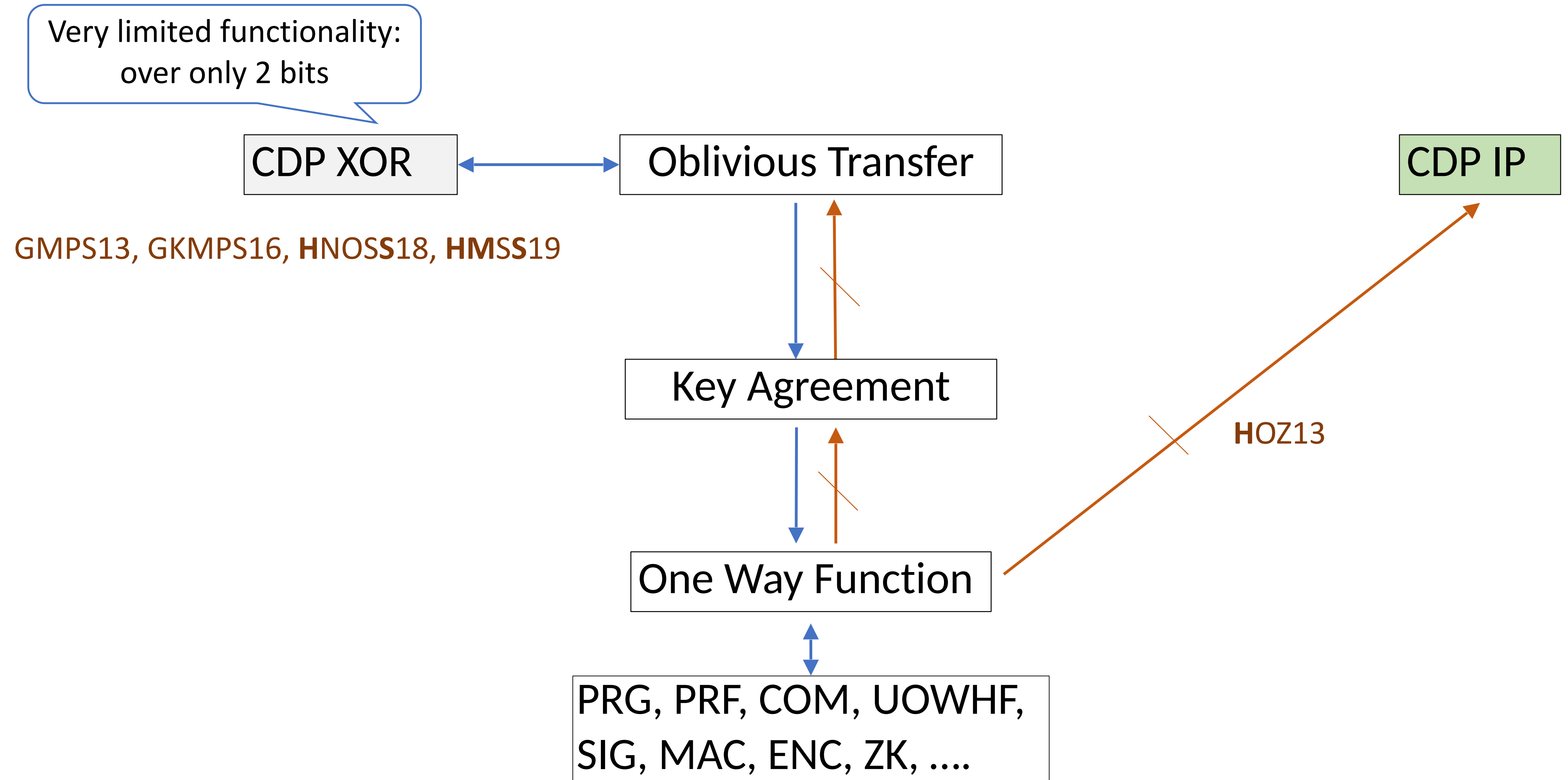
Complexity Hierarchy



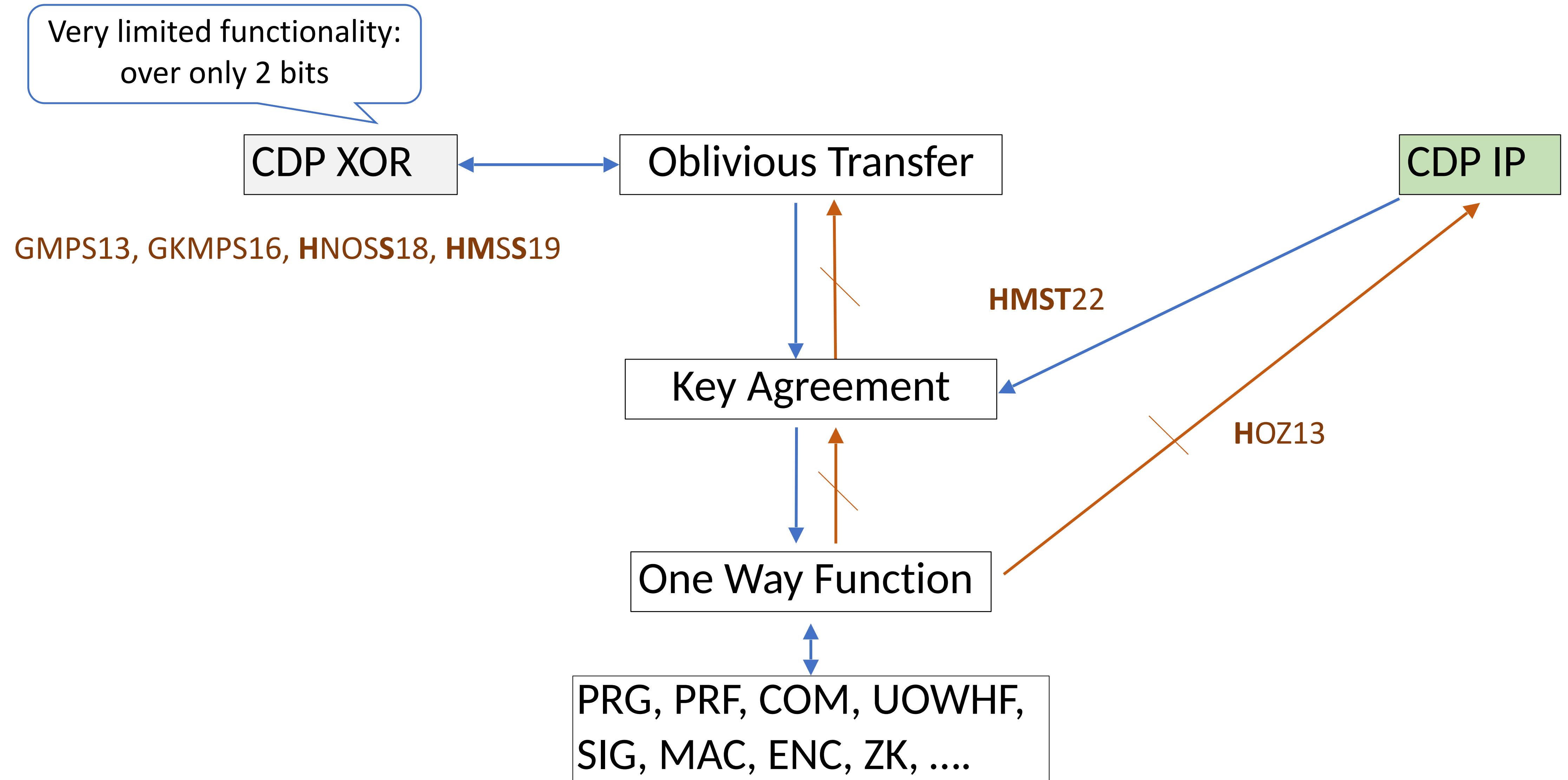
Complexity Hierarchy



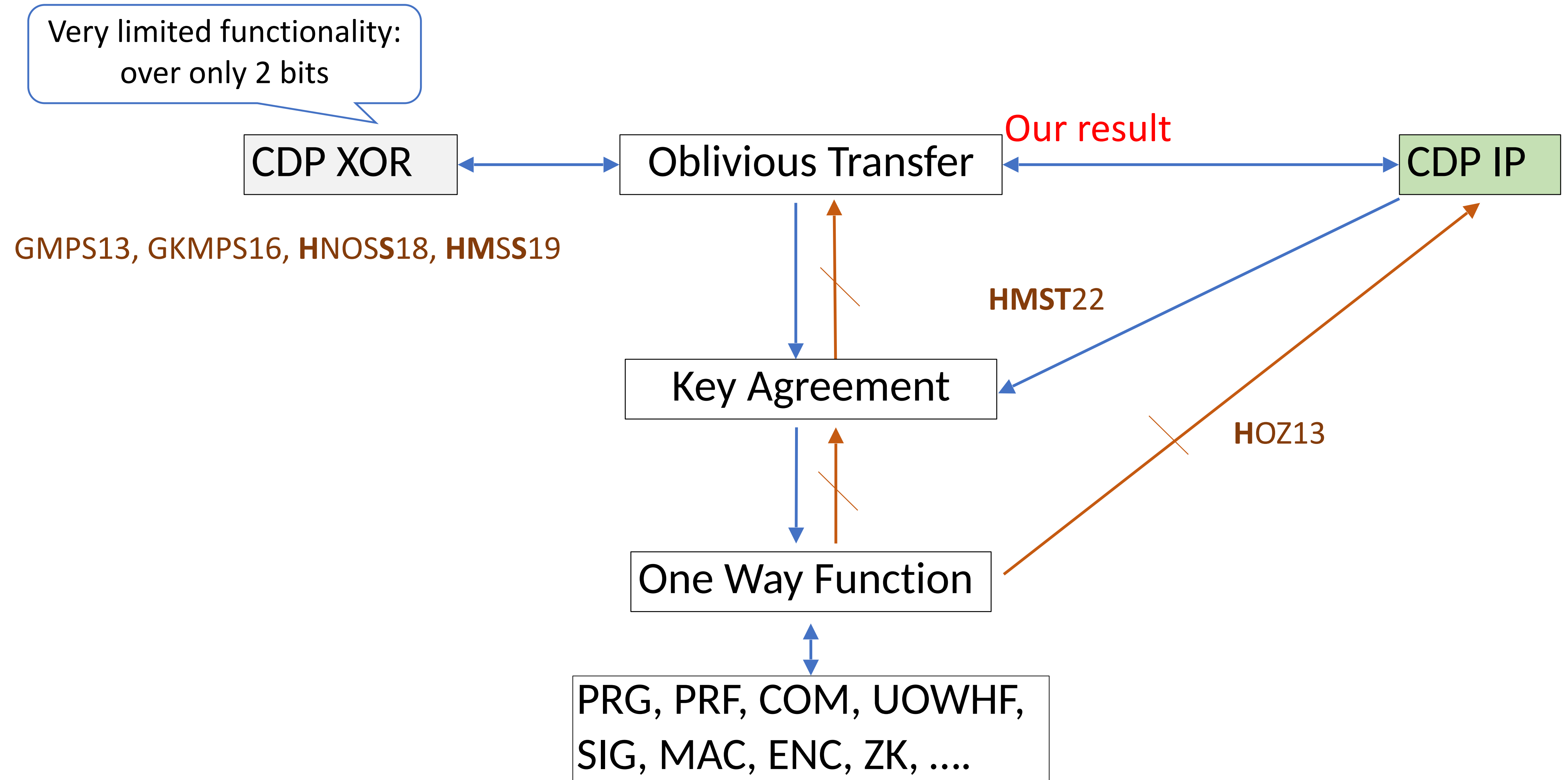
Complexity Hierarchy



Complexity Hierarchy



Complexity Hierarchy



Main Result

Main Theorem (informal):

Mildly accurate (ϵ, δ) -CDP for inner-product implies OT.

Main Result

Main Theorem (informal):

Mildly accurate (ε, δ) -CDP for inner-product implies OT.

- $\varepsilon = O(1)$
- $\delta = \left(\frac{1}{n}\right)$

Main Result

Main Theorem (informal):

Mildly accurate (ε, δ) -CDP for inner-product implies OT.

- $\varepsilon = O(1)$
- $\delta = \left(\frac{1}{n}\right)$
- $\Pr_{\substack{\mathbf{x}, \mathbf{y} \leftarrow \{-1, 1\}^n \\ out \leftarrow \Pi(\mathbf{x}, \mathbf{y})}} \left[\left| out - \langle \mathbf{x}, \mathbf{y} \rangle \right| < n^{1/6} \right] > 0.99$

Main Result

Main Theorem (informal):

Mildly accurate (ε, δ) -CDP for inner-product implies OT.

- $\varepsilon = O(1)$
- $\delta = \left(\frac{1}{n}\right)$
- $\Pr_{\substack{\mathbf{x}, \mathbf{y} \leftarrow \{-1, 1\}^n \\ \text{out} \leftarrow \Pi(\mathbf{x}, \mathbf{y})}} \left[\left| \text{out} - \langle \mathbf{x}, \mathbf{y} \rangle \right| < n^{1/6} \right] > 0.99$

Information theoretic barrier: $n^{1/2}$.

Main Result

Main Theorem (informal):

Mildly accurate (ε, δ) -CDP for inner-product implies OT.

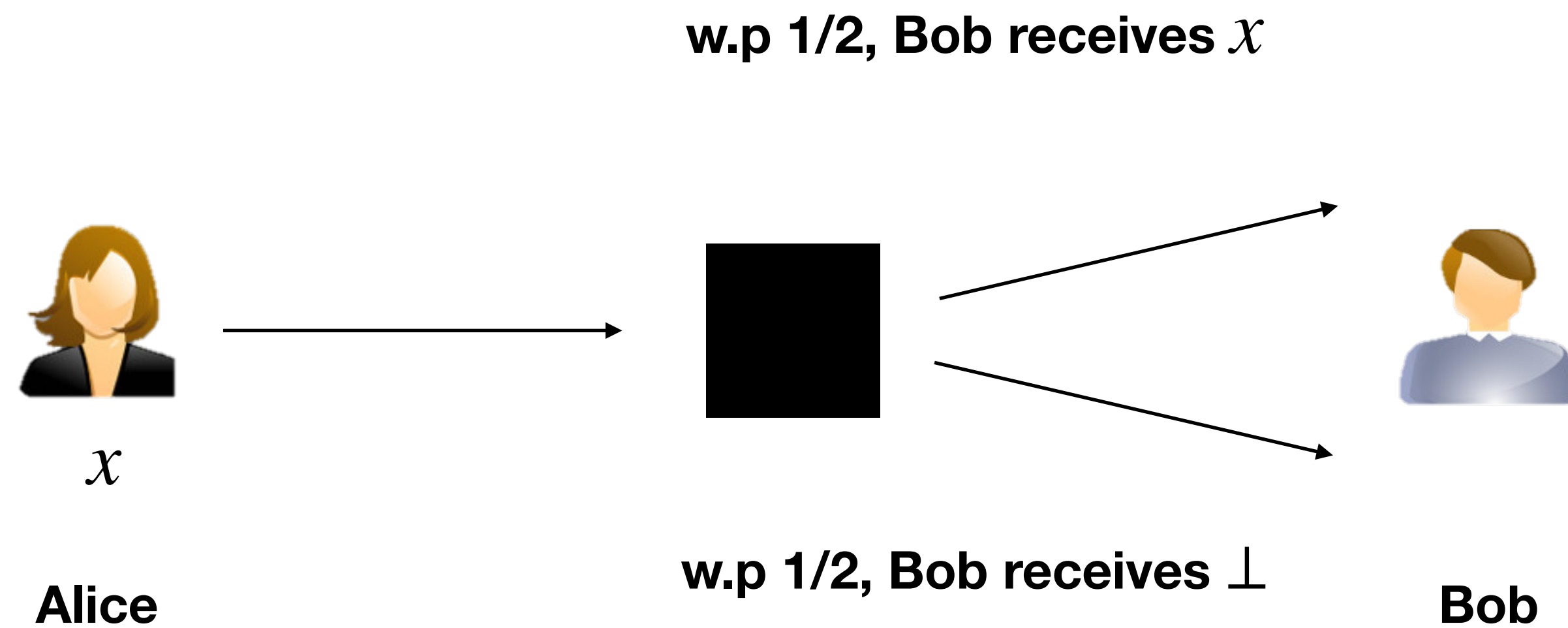
- $\varepsilon = O(1)$
- $\delta = \left(\frac{1}{n}\right)$
- $\Pr_{\substack{\mathbf{x}, \mathbf{y} \leftarrow \{-1, 1\}^n \\ \text{out} \leftarrow \Pi(\mathbf{x}, \mathbf{y})}} \left[\left| \text{out} - \langle \mathbf{x}, \mathbf{y} \rangle \right| < n^{1/6} \right] > 0.99$
- Mildly accurate $\Rightarrow$ Very accurate, $O(1)$

Information theoretic barrier: $n^{1/2}$.

Proof Overview

Rabin's OT

Rabin 1981



Alice doesn't know what Bob receives

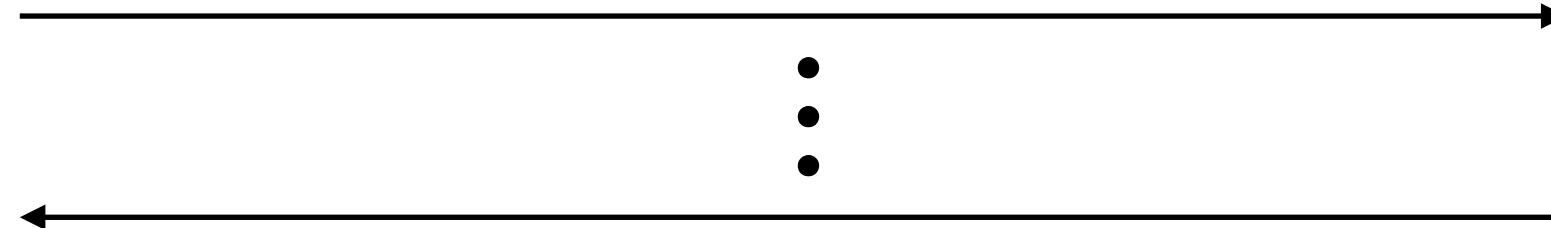
Weak Erasure Channel

Wullschleger 2009



$out_A \in \{0,1\}$

$view_A$

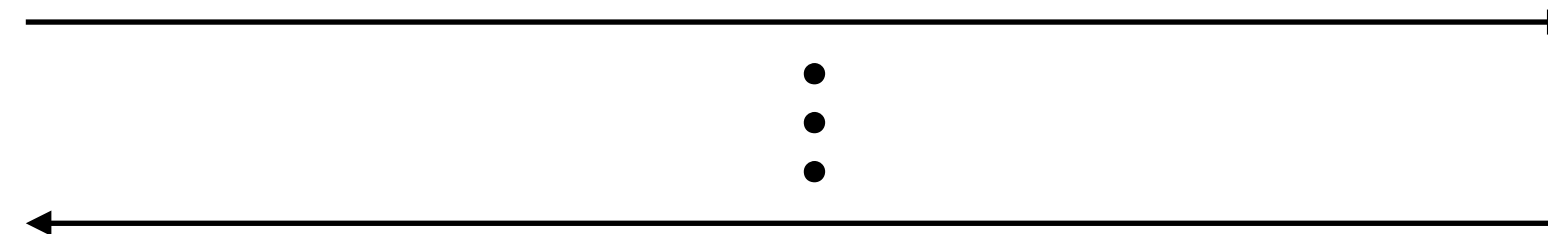


$out_B \in \{0,1,\perp\}$

$view_B$

Weak Erasure Channel

Wullschleger 2009



$out_A \in \{0,1\}$

$view_A$

$out_B \in \{0,1,\perp\}$

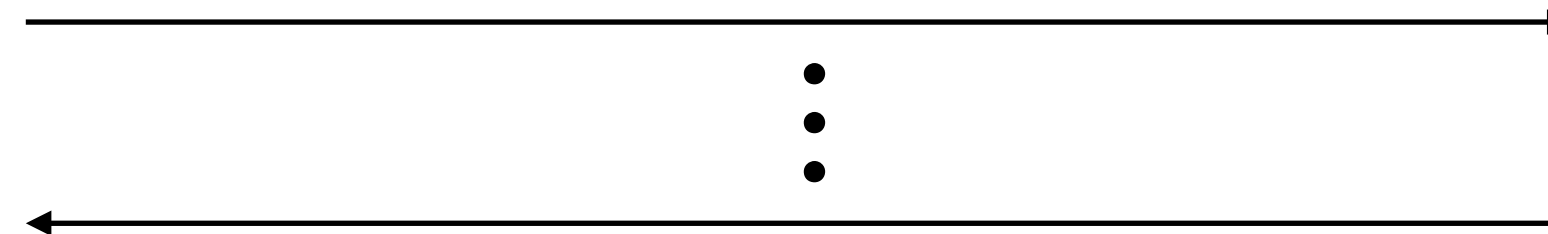
$view_B$

w.p 1/2 (non-erasure) $out_B \neq \perp$

- Agreement: $\Pr[out_A = out_B] \geq 0.99$

Weak Erasure Channel

Wullschleger 2009



$out_A \in \{0,1\}$

$view_A$

$out_B \in \{0,1,\perp\}$

$view_B$

w.p 1/2 (non-erasure) $out_B \neq \perp$

- Agreement: $\Pr[out_A = out_B] \geq 0.99$

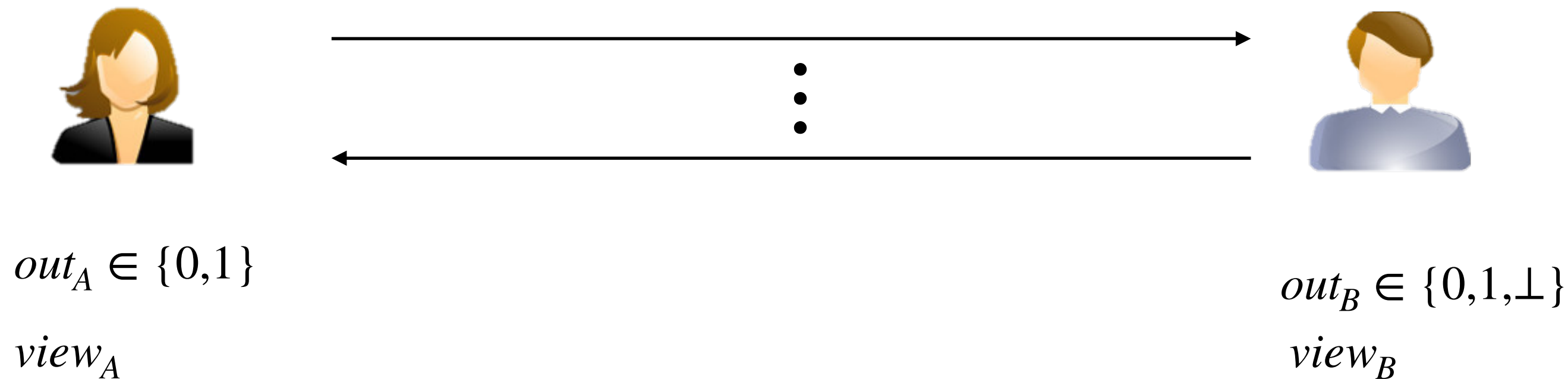
w.p 1/2 (erasure) $out_B = \perp$

- Secrecy: (If $out_B = \perp$, Bob doesn't know what the value of out_A)

$$\forall \text{ PPT } D, \Pr[D(view_B) = out_A] \leq 0.01$$

Weak Erasure Channel

Wullschleger 2009



w.p 1/2 (non-erasure) $out_B \neq \perp$

- Agreement: $\Pr[out_A = out_B] \geq 0.99$

w.p 1/2 (erasure) $out_B = \perp$

- Secrecy: (If $out_B = \perp$, Bob doesn't know what the value of out_A)

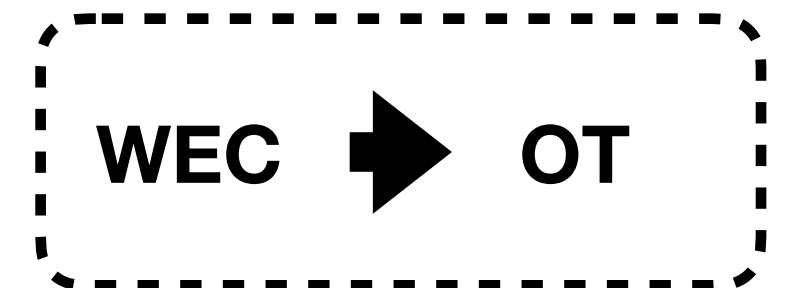
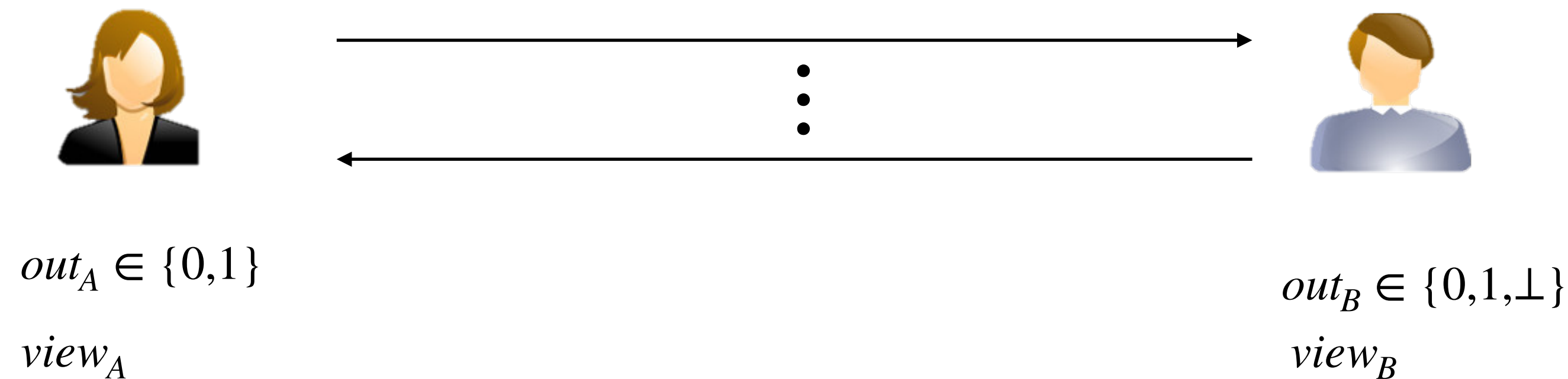
$$\forall \text{ PPT } D, \Pr[D(view_B) = out_A] \leq 0.01$$

- (Alice doesn't know if $out_B = \perp$)

$$\forall \text{ PPT } D, |\Pr[D(view_A) = 1 \mid out_B \neq \perp] - \Pr[D(view_A) = 1 \mid out_B = \perp]| \leq 0.01$$

Weak Erasure Channel

Wullschleger 2009



w.p 1/2 (non-erasure) $out_B \neq \perp$

- Agreement: $\Pr[out_A = out_B] \geq 0.99$

w.p 1/2 (erasure) $out_B = \perp$

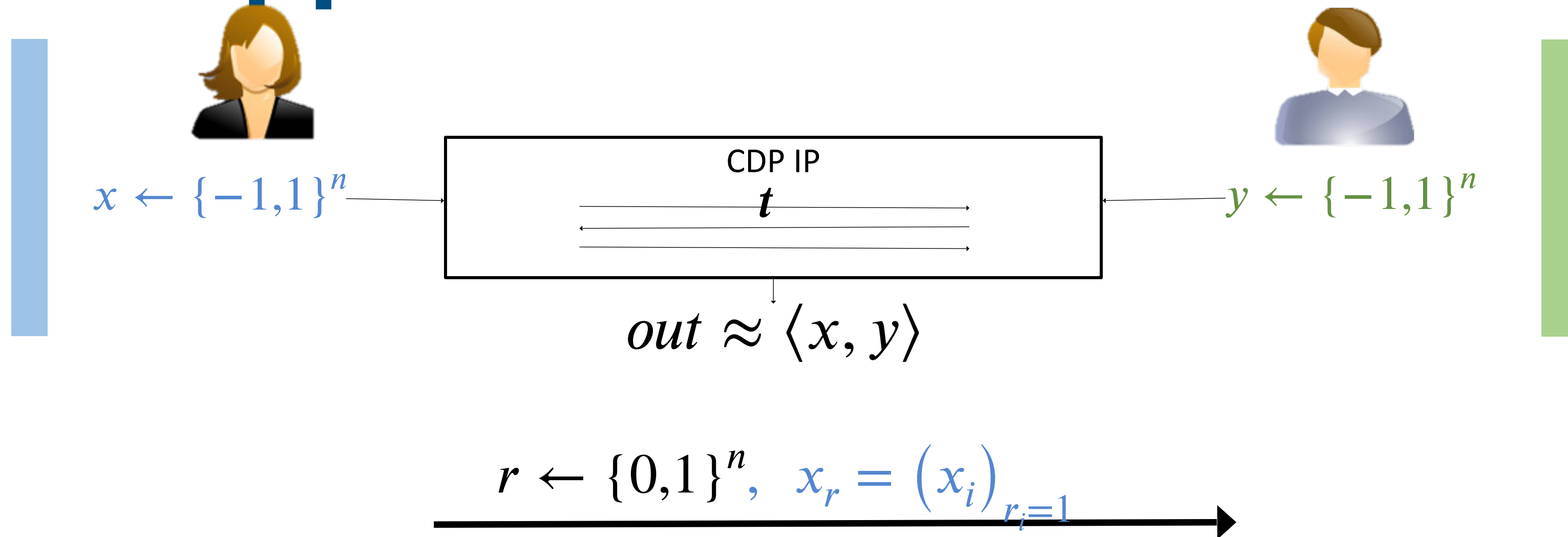
- Secrecy: (If $out_B = \perp$, Bob doesn't know what the value of out_A)

$$\forall \text{ PPT } D, \Pr[D(view_B) = out_A] \leq 0.01$$

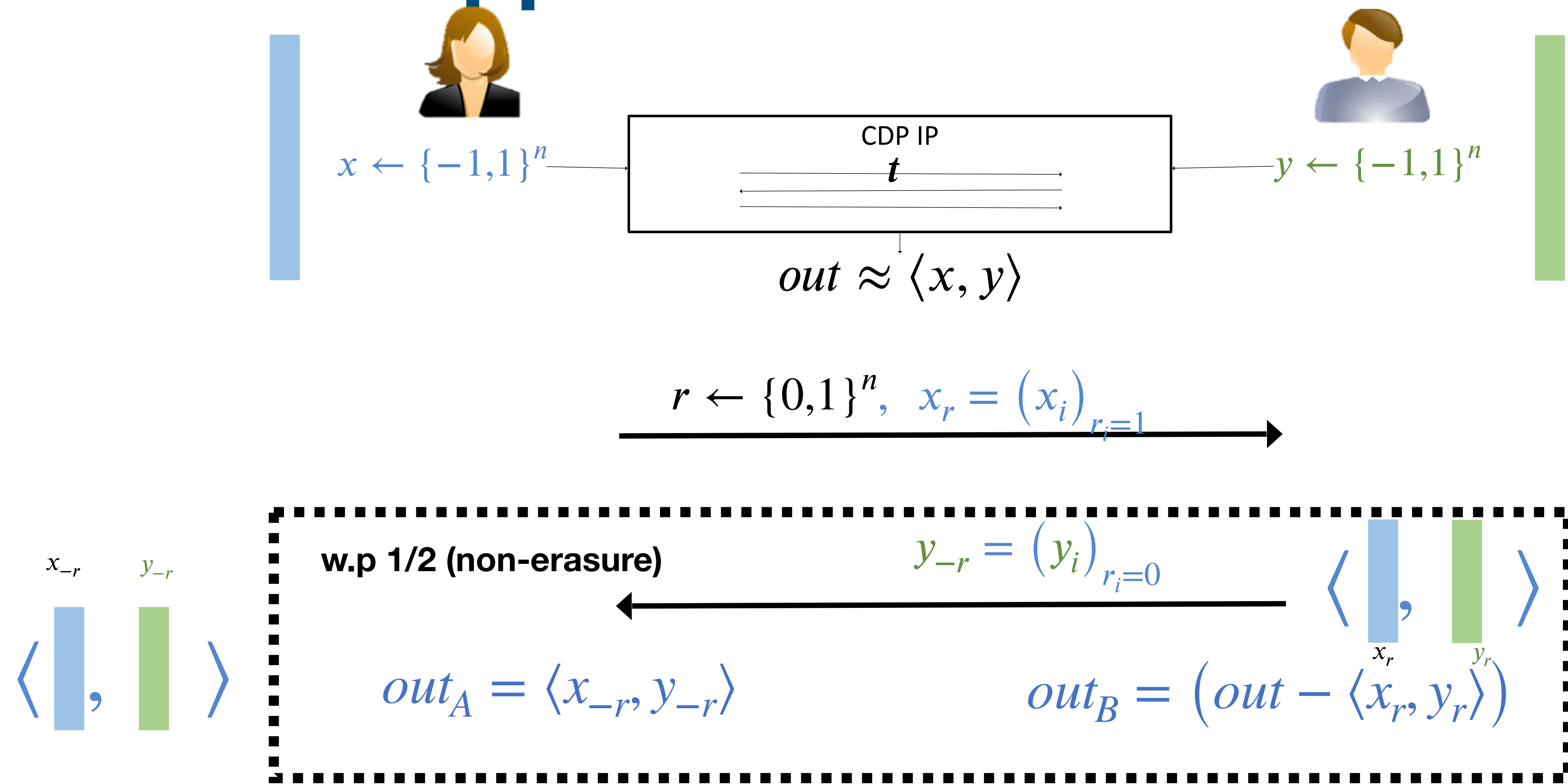
- (Alice doesn't know if $out_B = \perp$)

$$\forall \text{ PPT } D, |\Pr[D(view_A) = 1 \mid out_B \neq \perp] - \Pr[D(view_A) = 1 \mid out_B = \perp]| \leq 0.01$$

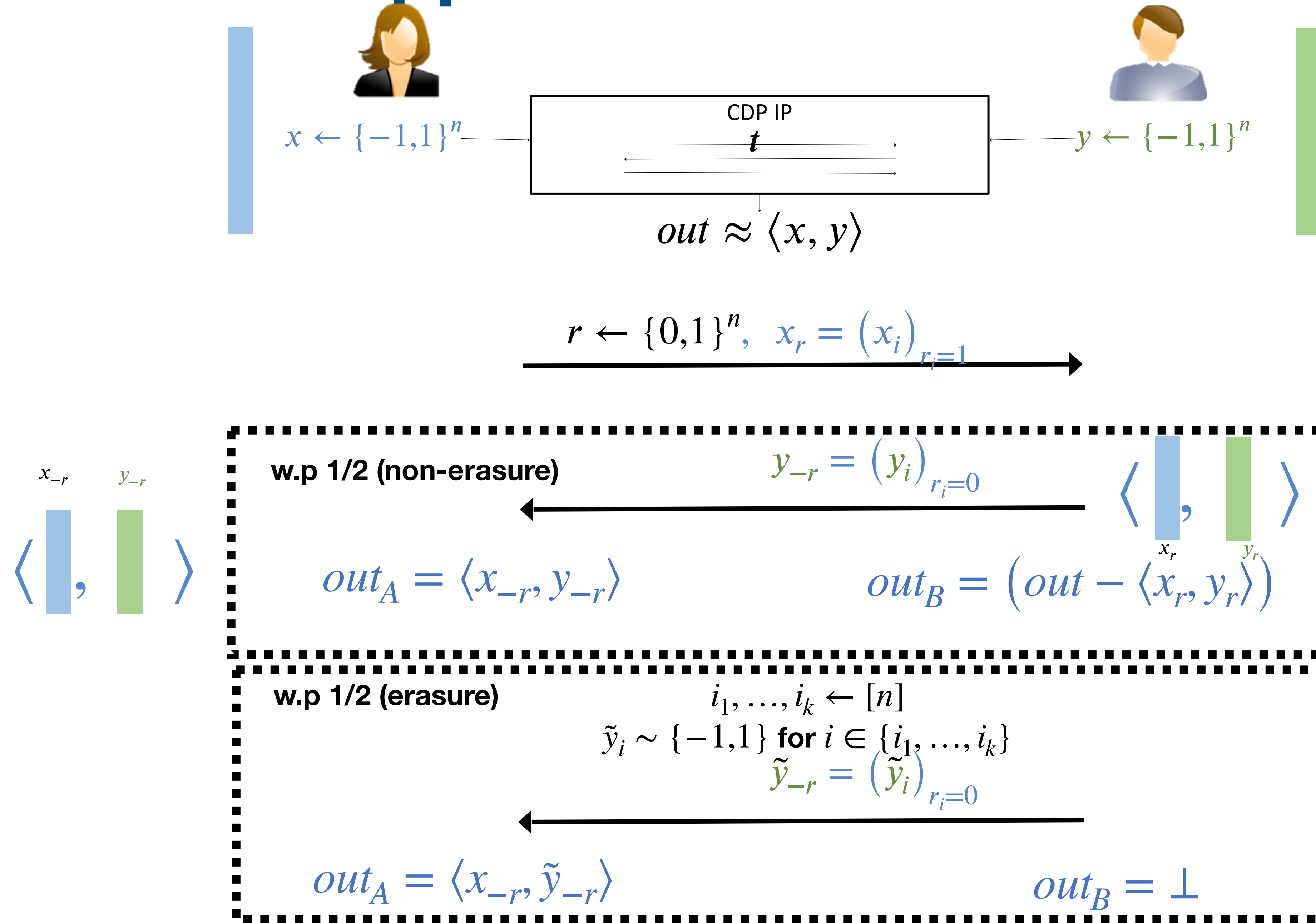
CDP IP to Approximate Weak Erasure Channel



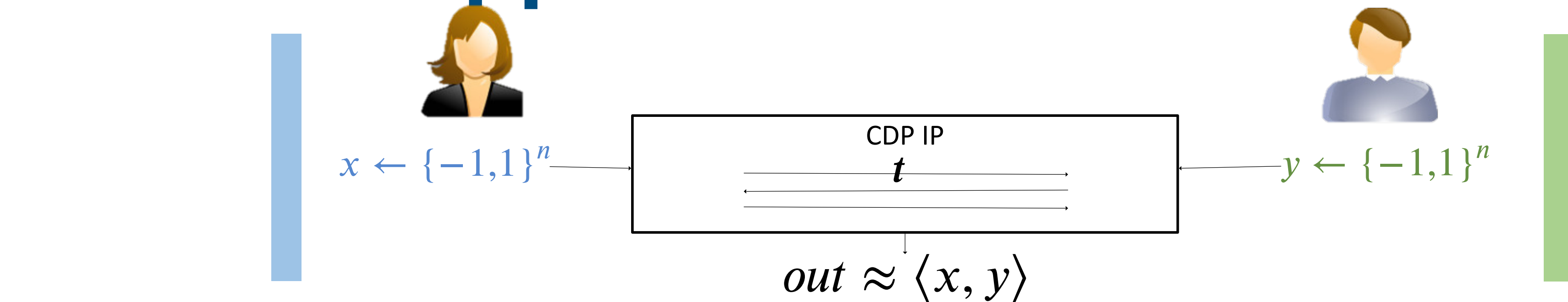
CDP IP to Approximate Weak Erasure Channel



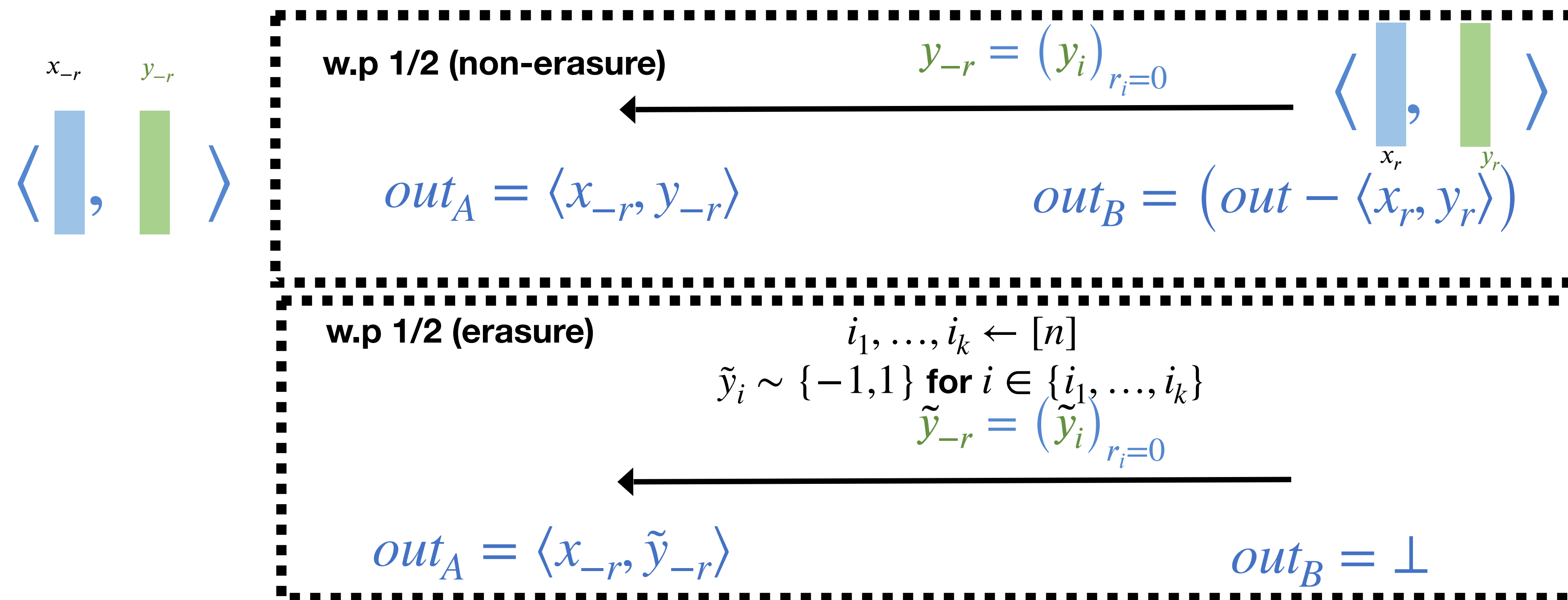
CDP IP to Approximate Weak Erasure Channel



CDP IP to Approximate Weak Erasure Channel

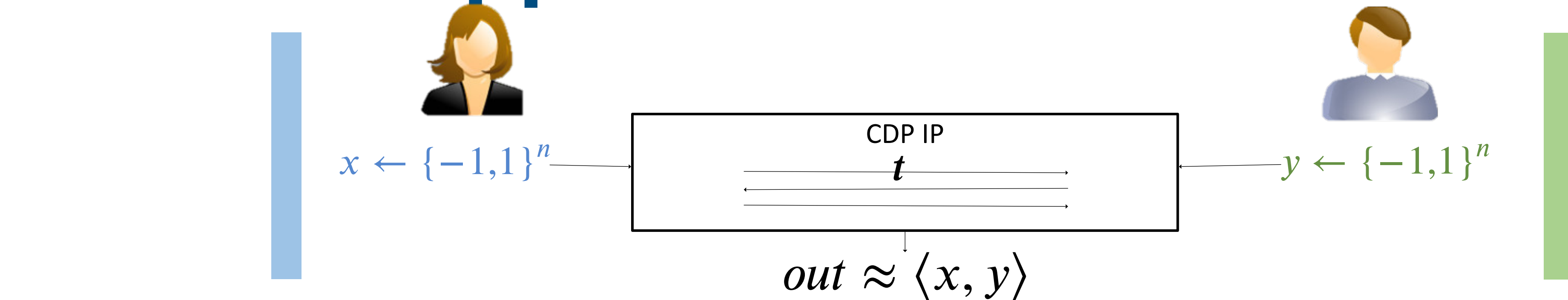


$$r \leftarrow \{0,1\}^n, \quad x_r = (x_i)_{r_i=1}$$

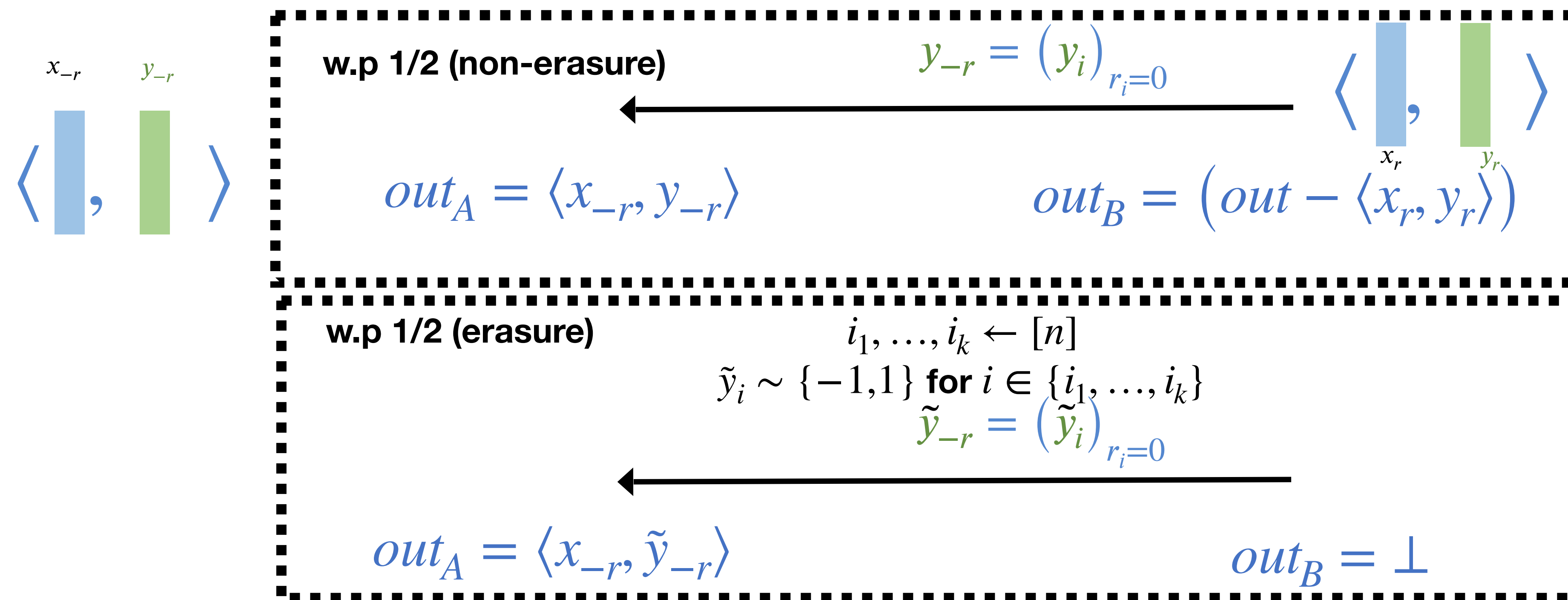


- Alice doesn't know if $out_B = \perp$
- If $out_B = \perp$, Bob doesn't know what the value of out_A

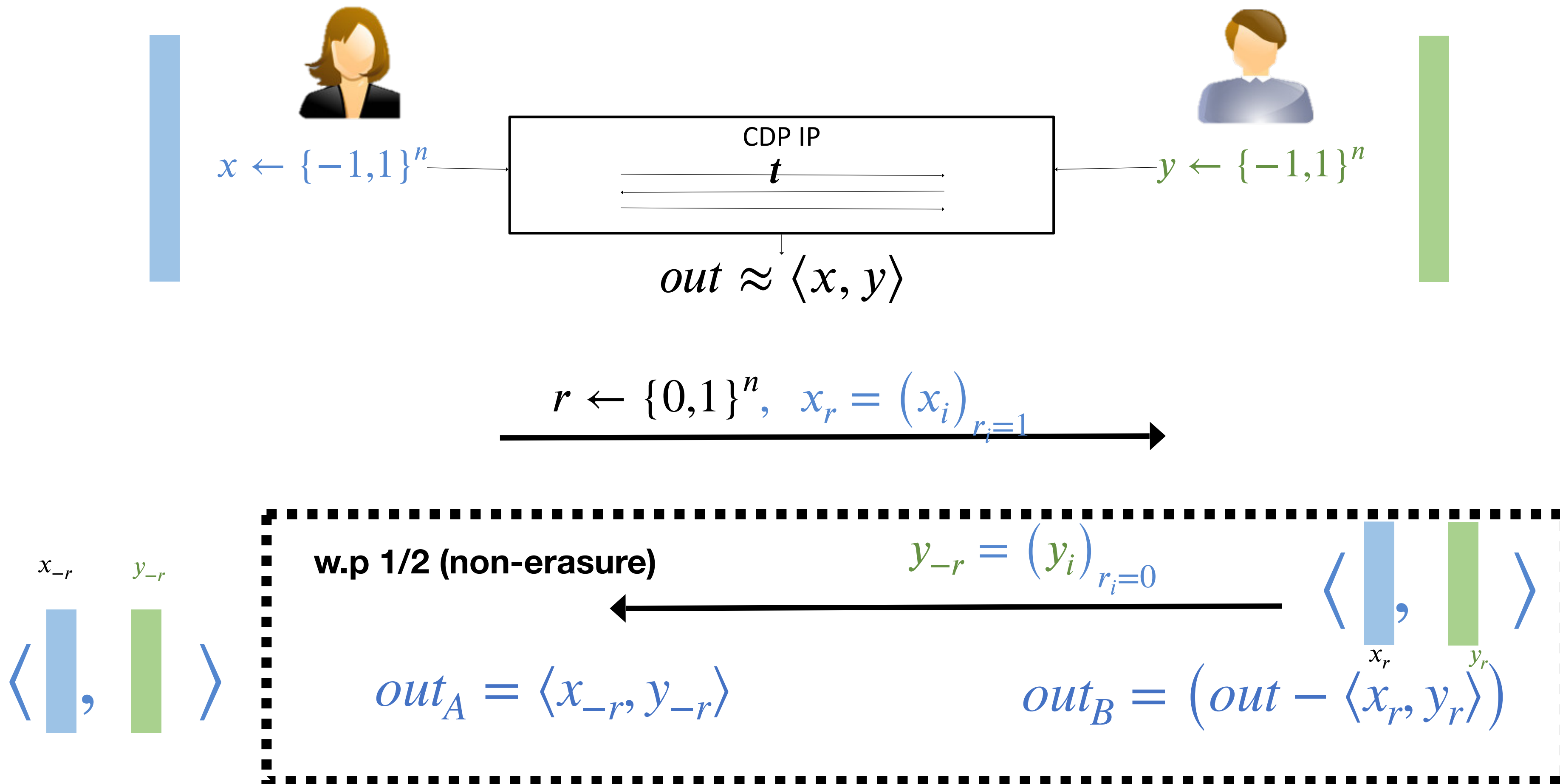
CDP IP to Approximate Weak Erasure Channel



$$r \leftarrow \{0,1\}^n, \quad x_r = (x_i)_{r_i=1}$$

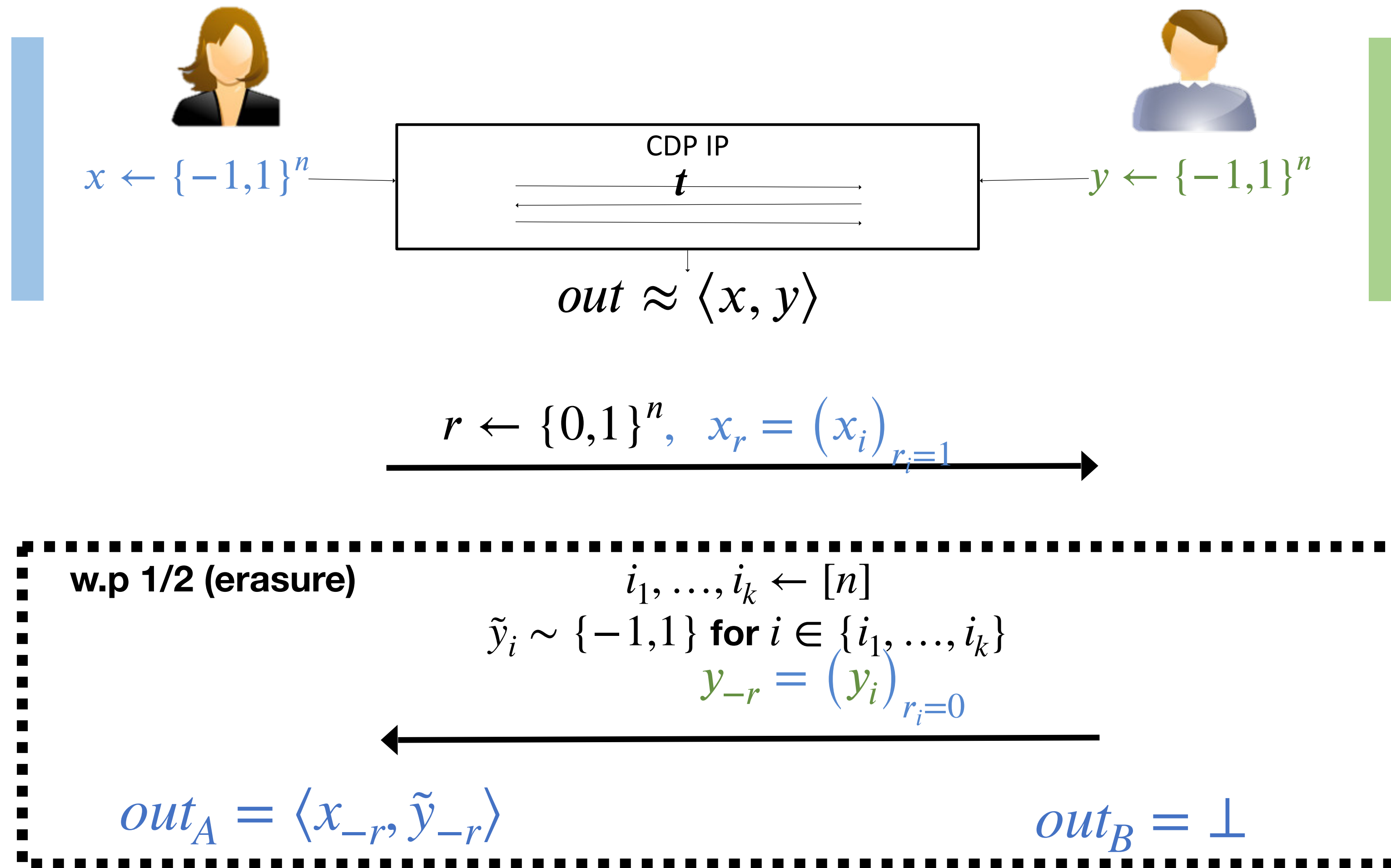


- Alice doesn't know if $out_B = \perp$
- If $out_B = \perp$, Bob doesn't know what the value of out_A
- AWEC can be amplified to WEC



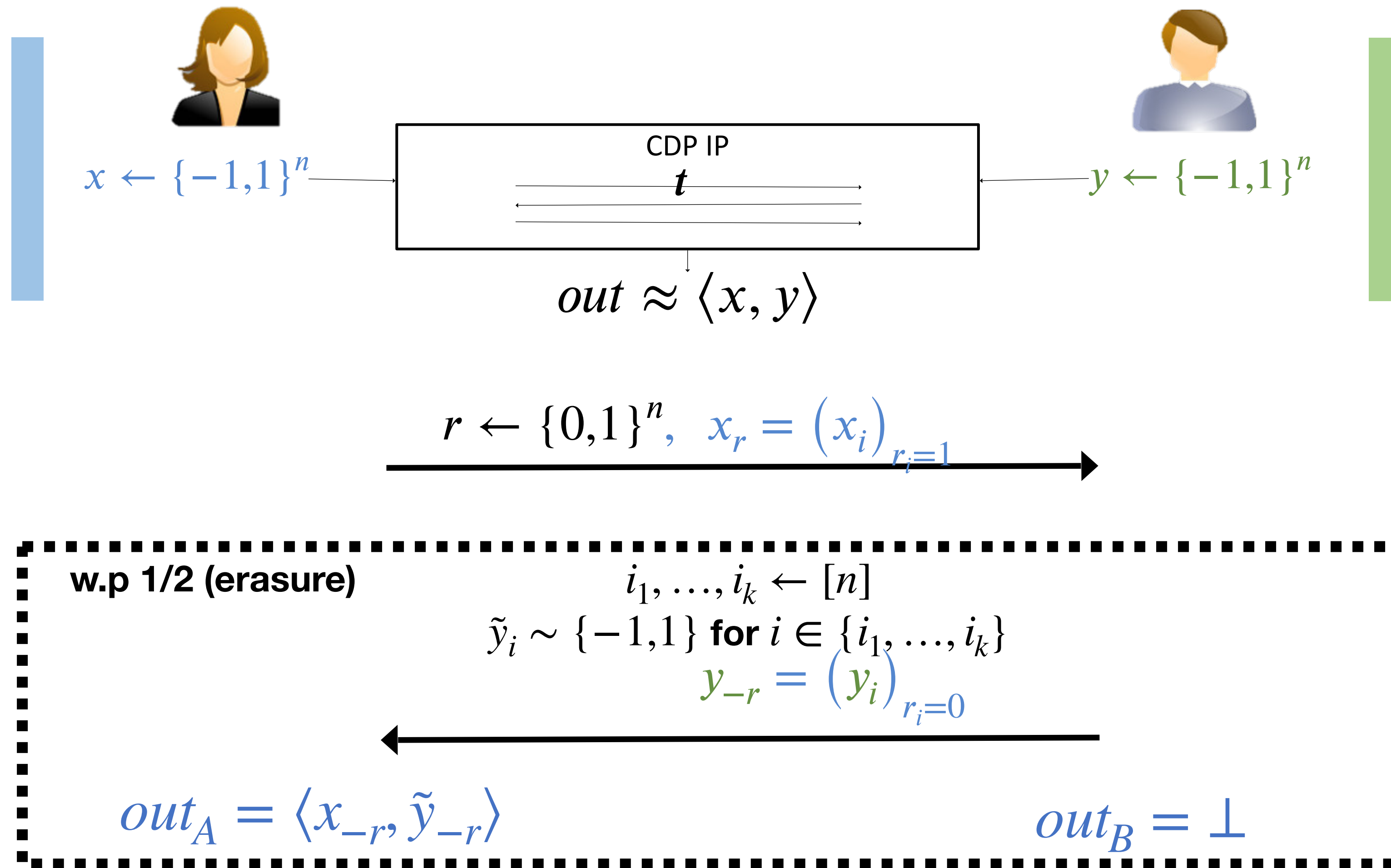
- Agreement:

$$Out \approx \langle X, Y \rangle \implies Out_A \approx Out_B$$



- Secrecy of Alice:

HMST22: $\forall$ PPT Eve, $Eve(T, R, X_R, Y_{-R})$ is far from out_A .



- Secrecy of Alice:

HMST22: $\forall$ PPT Eve, $Eve(T, R, X_R, Y_{-R})$ is far from out_A .

➤ If $out_B = \perp$, Bob doesn't know what the value of out_A

Secrecy of Bob

Toy case

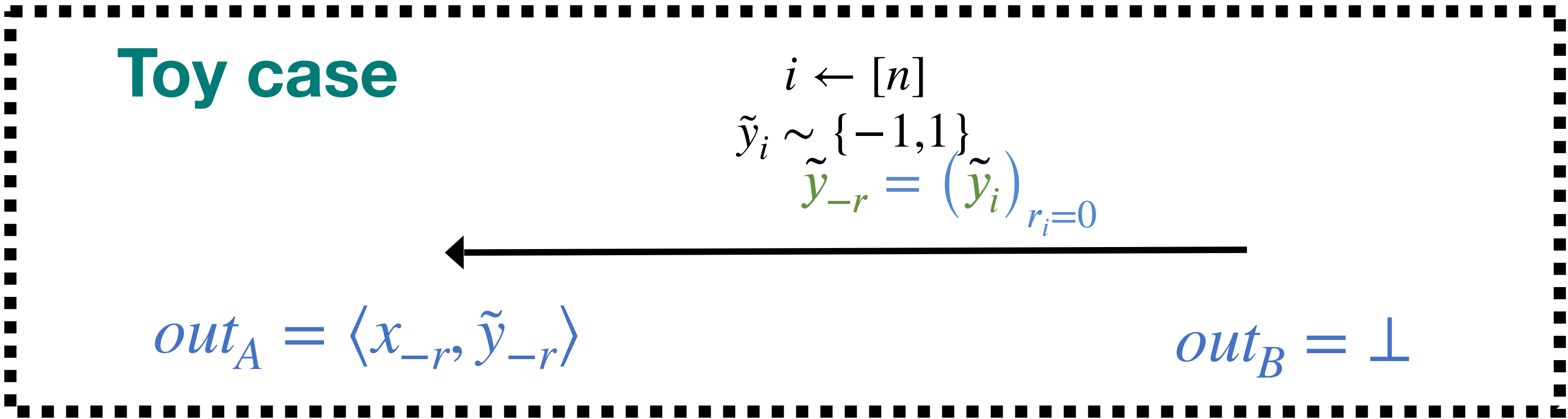
$$\begin{aligned} i &\leftarrow [n] \\ \tilde{y}_i &\sim \{-1, 1\} \\ \tilde{y}_{-r} &= (\tilde{y}_i)_{r_i=0} \end{aligned}$$



$$out_A = \langle x_{-r}, \tilde{y}_{-r} \rangle$$

$$out_B = \perp$$

Secrecy of Bob



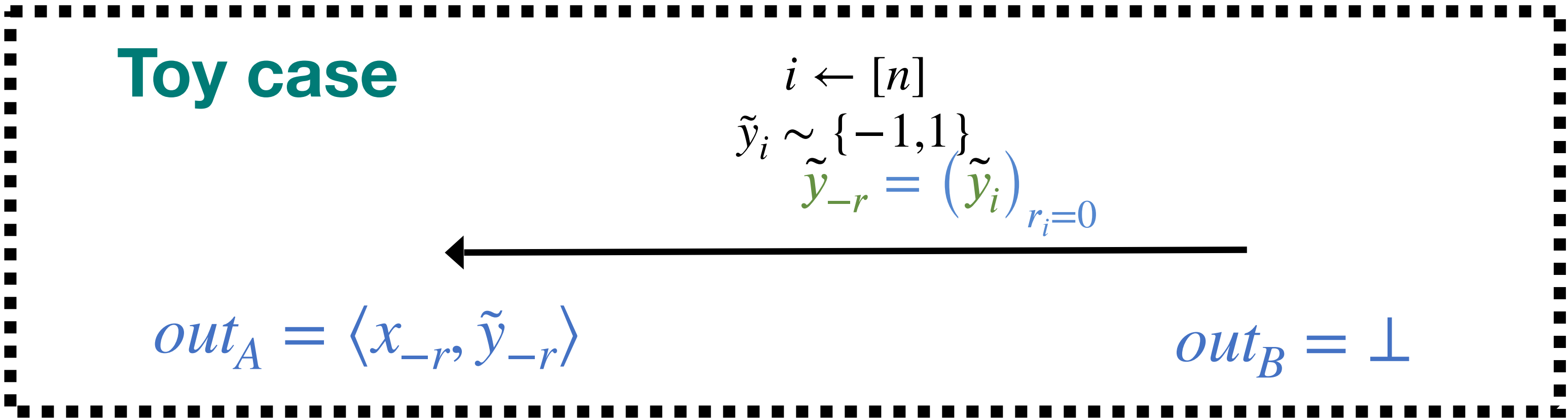
$\epsilon = O(1), \delta = \frac{1}{n}$

$x \leftarrow \{-1, 1\}^n$

Key observation:

Given an (ϵ, δ) -DP mechanism M , and database $x = (x_1, \cdots, x_n)$ that has high min-entropy.

Secrecy of Bob



$$\epsilon = O(1), \delta = \frac{1}{n}$$

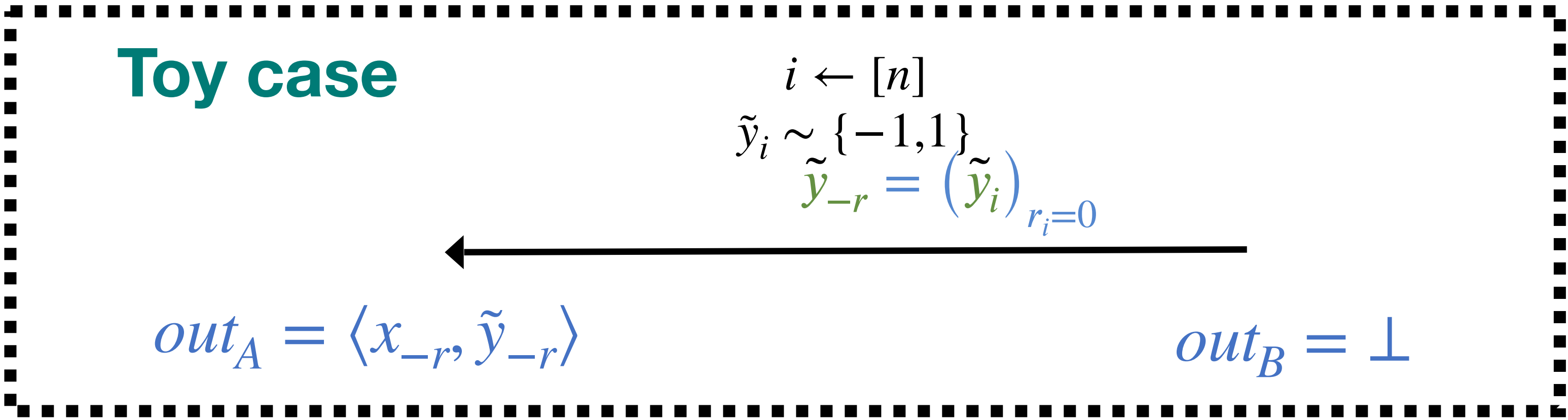
$$x \leftarrow \{-1, 1\}^n$$

Key observation:

Given an (ϵ, δ) -DP mechanism M , and database $x = (x_1, \dots, x_n)$ that has high min-entropy.

- Let $i \leftarrow [n]$, and let $x^i = (x_1, \dots, -x_i, \dots, x_n)$.
- For a random $r \leftarrow \{0, 1\}^n$, $x_r = (x_i)_{r_i=1}$.

Secrecy of Bob



$\epsilon = O(1), \delta = \frac{1}{n}$

$x \leftarrow \{-1, 1\}^n$

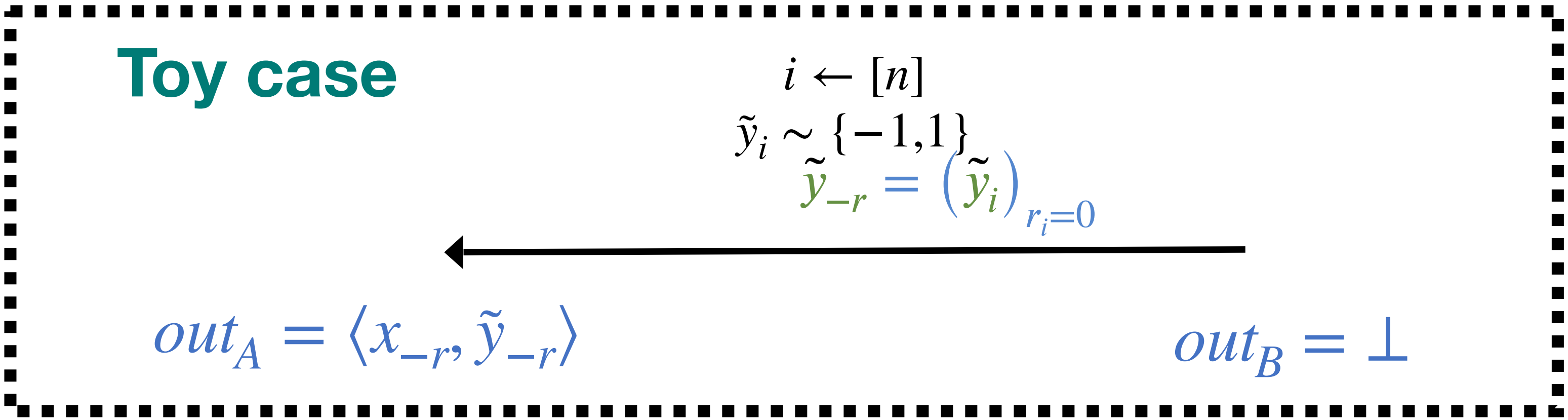
Key observation:

Given an (ϵ, δ) -DP mechanism M , and database $x = (x_1, \dots, x_n)$ that has high min-entropy.

- Let $i \leftarrow [n]$, and let $x^i = (x_1, \dots, -x_i, \dots, x_n)$.
- For a random $r \leftarrow \{0, 1\}^n$, $x_r = (x_i)_{r_i=1}$.
- For any adversary A :

$$|\Pr[A(x_r, r, M(x))=1] - \Pr[A(x_r^i, r, M(x))]| > 1/n^{\frac{1}{3}}$$

Secrecy of Bob



$\epsilon = O(1), \delta = \frac{1}{n}$

$x \leftarrow \{-1, 1\}^n$

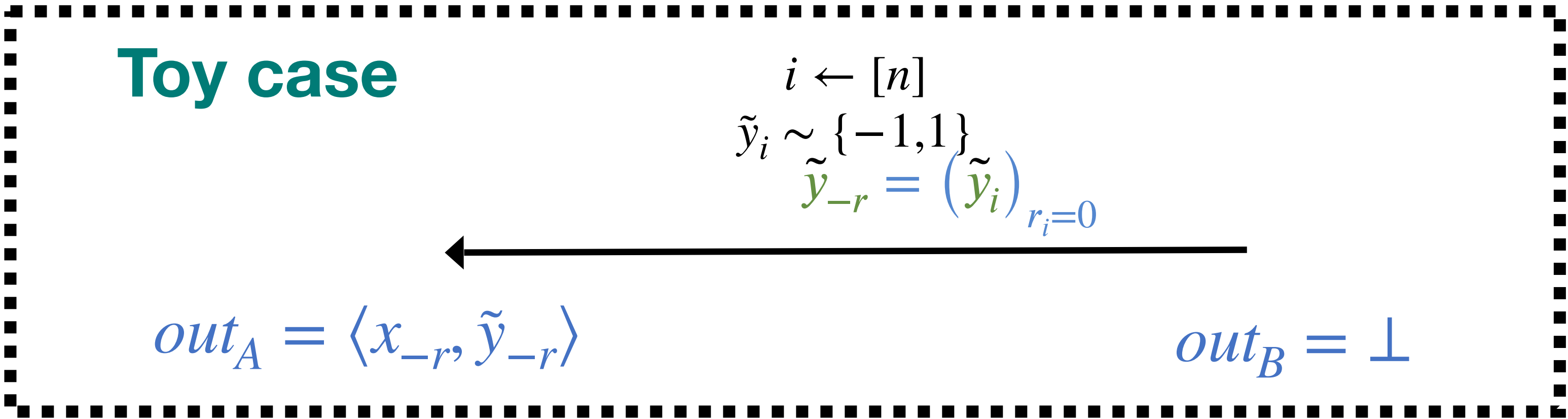
Key observation:

Given an (ϵ, δ) -DP mechanism M , and database $x = (x_1, \dots, x_n)$ that has high min-entropy.

- Let $i \leftarrow [n]$, and let $x^i = (x_1, \dots, -x_i, \dots, x_n)$.
- For a random $r \leftarrow \{0, 1\}^n$, $x_r = (x_i)_{r_i=1}$.
- For any adversary A :

$|\Pr[A(x_r, r, M(x))=1] - \Pr[A(x_r^i, r, M(x))]| > 1/n^{\frac{1}{3}} \quad \Rightarrow \quad A \text{ can be used to contradict the } (\epsilon, \delta)\text{-DP of } M.$

Secrecy of Bob



$$\epsilon = O(1), \delta = \frac{1}{n}$$

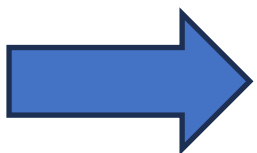
$$x \leftarrow \{-1, 1\}^n$$

Key observation:

Given an (ϵ, δ) -DP mechanism M , and database $x = (x_1, \dots, x_n)$ that has high min-entropy.

- Let $i \leftarrow [n]$, and let $x^i = (x_1, \dots, -x_i, \dots, x_n)$.
- For a random $r \leftarrow \{0, 1\}^n$, $x_r = (x_i)_{r_i=1}$.
- For any adversary A :

$$|\Pr[A(x_r, r, M(x))=1] - \Pr[A(x_r^i, r, M(x))]| > 1/n^{\frac{1}{3}}$$



A can be used to contradict the (ϵ, δ) -DP of M .

A constructive proof:
works also for **CDP**

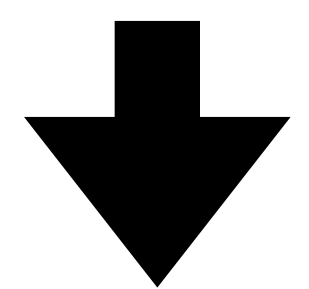
Toy case

$$i \leftarrow [n]$$
$$\tilde{y}_i \sim \{-1, 1\}$$
$$\tilde{y}_{-r} = (\tilde{y}_i)_{r_i=0}$$



$$out_A = \langle x_{-r}, \tilde{y}_{-r} \rangle$$

$$out_B = \perp$$



can be extended to $k = O(n^{1/3})$ random indexes

w.p 1/2 (erasure)

$$i_1, \dots, i_k \leftarrow [n]$$
$$\tilde{y}_i \sim \{-1, 1\} \text{ for } i \in \{i_1, \dots, i_k\}$$
$$y_{-r} = (y_i)_{r_i=0}$$



$$out_A = \langle x_{-r}, \tilde{y}_{-r} \rangle$$

$$out_B = \perp$$

Secrecy of Bob

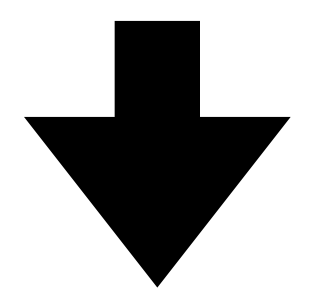
Toy case

$$i \leftarrow [n]$$
$$\tilde{y}_i \sim \{-1, 1\}$$
$$\tilde{y}_{-r} = (\tilde{y}_i)_{r_i=0}$$



$$out_A = \langle x_{-r}, \tilde{y}_{-r} \rangle$$

$$out_B = \perp$$



can be extended to $k = O(n^{1/3})$ random indexes

w.p 1/2 (erasure)

$$i_1, \dots, i_k \leftarrow [n]$$
$$\tilde{y}_i \sim \{-1, 1\} \text{ for } i \in \{i_1, \dots, i_k\}$$
$$y_{-r} = (y_i)_{r_i=0}$$



$$out_A = \langle x_{-r}, \tilde{y}_{-r} \rangle$$

$$out_B = \perp$$

$\Rightarrow$ Alice doesn't know if $out_B = \perp$

Conclusions & Open Problems

Mildly accurate CDP-IP $\Rightarrow$ OT

Open Questions:

- Answer for the error in the range $(n^{1/6}, \sqrt{n})$

Conclusions & Open Problems

Mildly accurate CDP-IP $\Rightarrow$ OT

Open Questions:

- Answer for the error in the range $(n^{1/6}, \sqrt{n})$
- Finding a more general characterization that captures more functionalities.

Conclusions & Open Problems

Mildly accurate CDP-IP $\Rightarrow$ OT

Open Questions:

- Answer for the error in the range $(n^{1/6}, \sqrt{n})$
- Finding a more general characterization that captures more functionalities.

Thank You