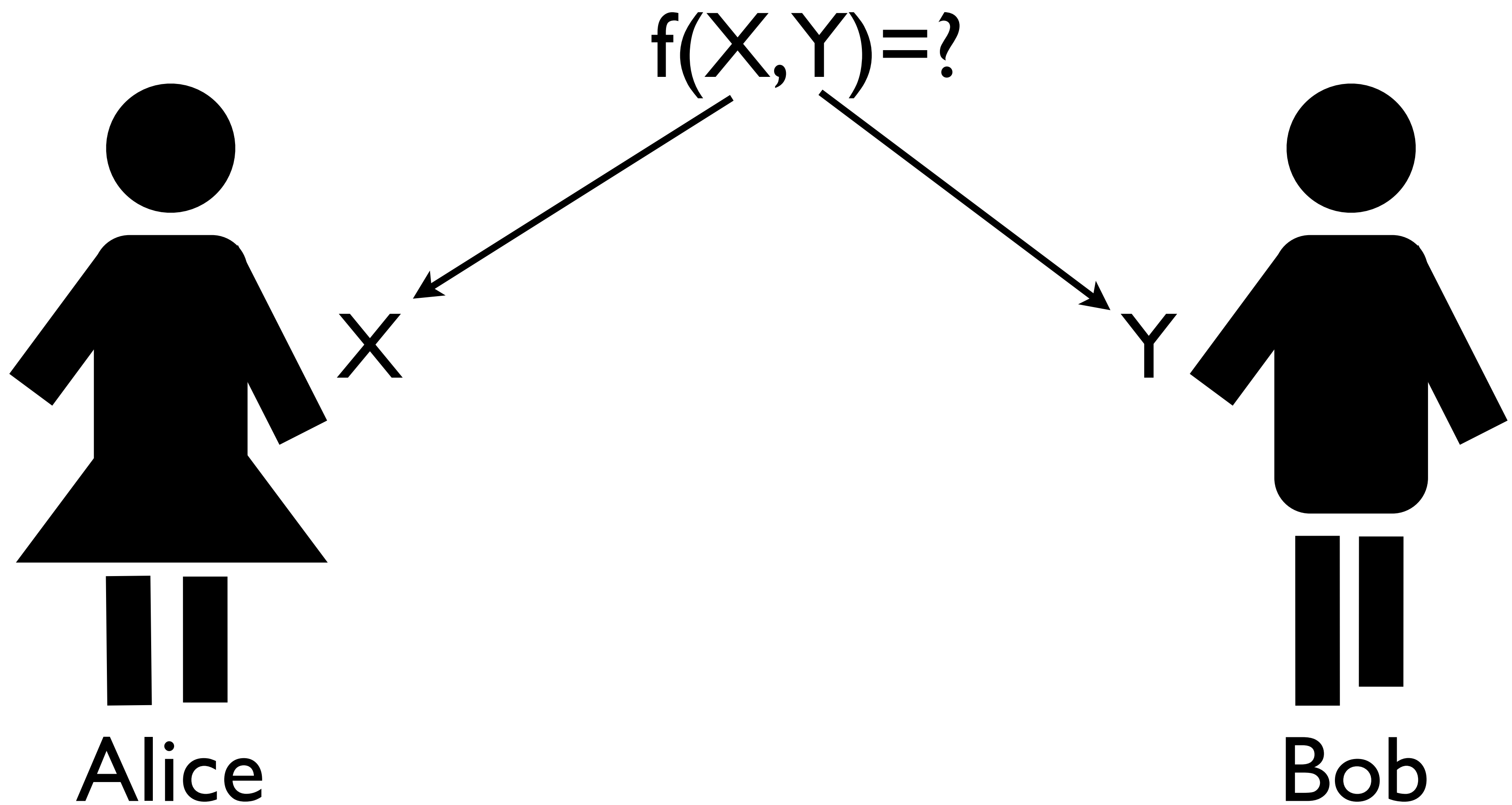
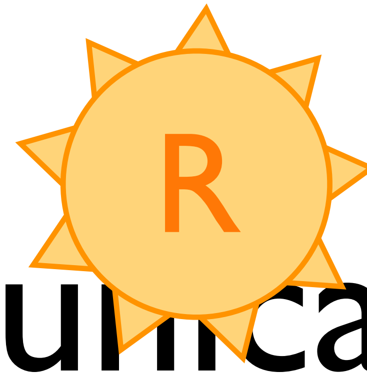
On the communication complexity of sparse set disjointness and exists-equal problems

Mert Saglam
University of Washington

Joint work with Gábor Tardos

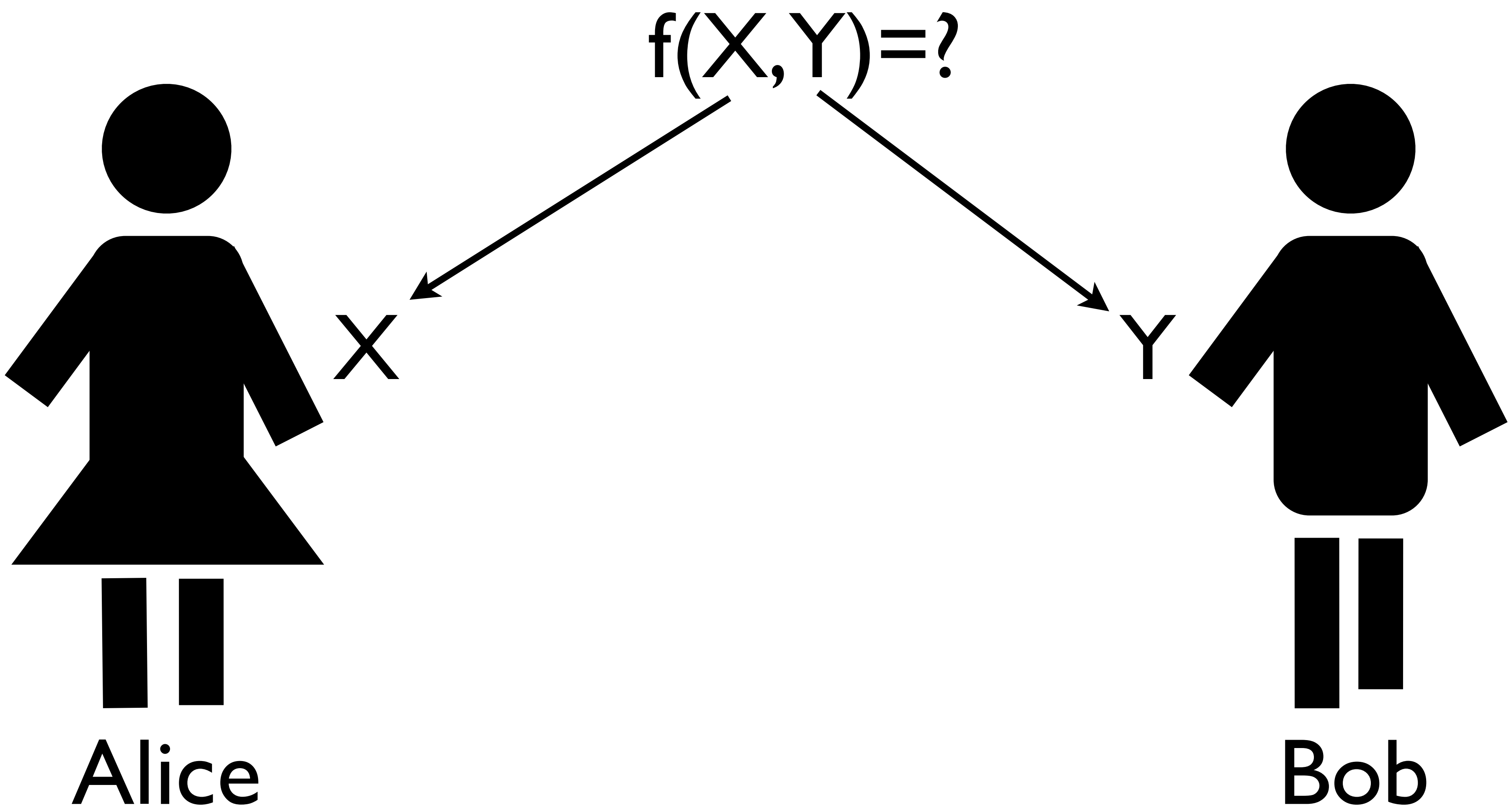
Communication complexity



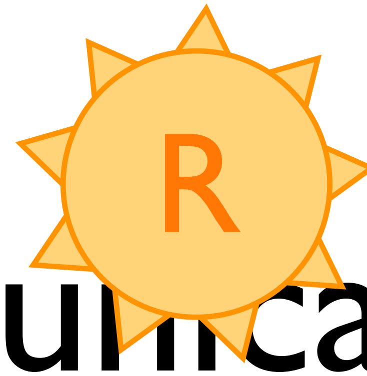


R: Shared random source

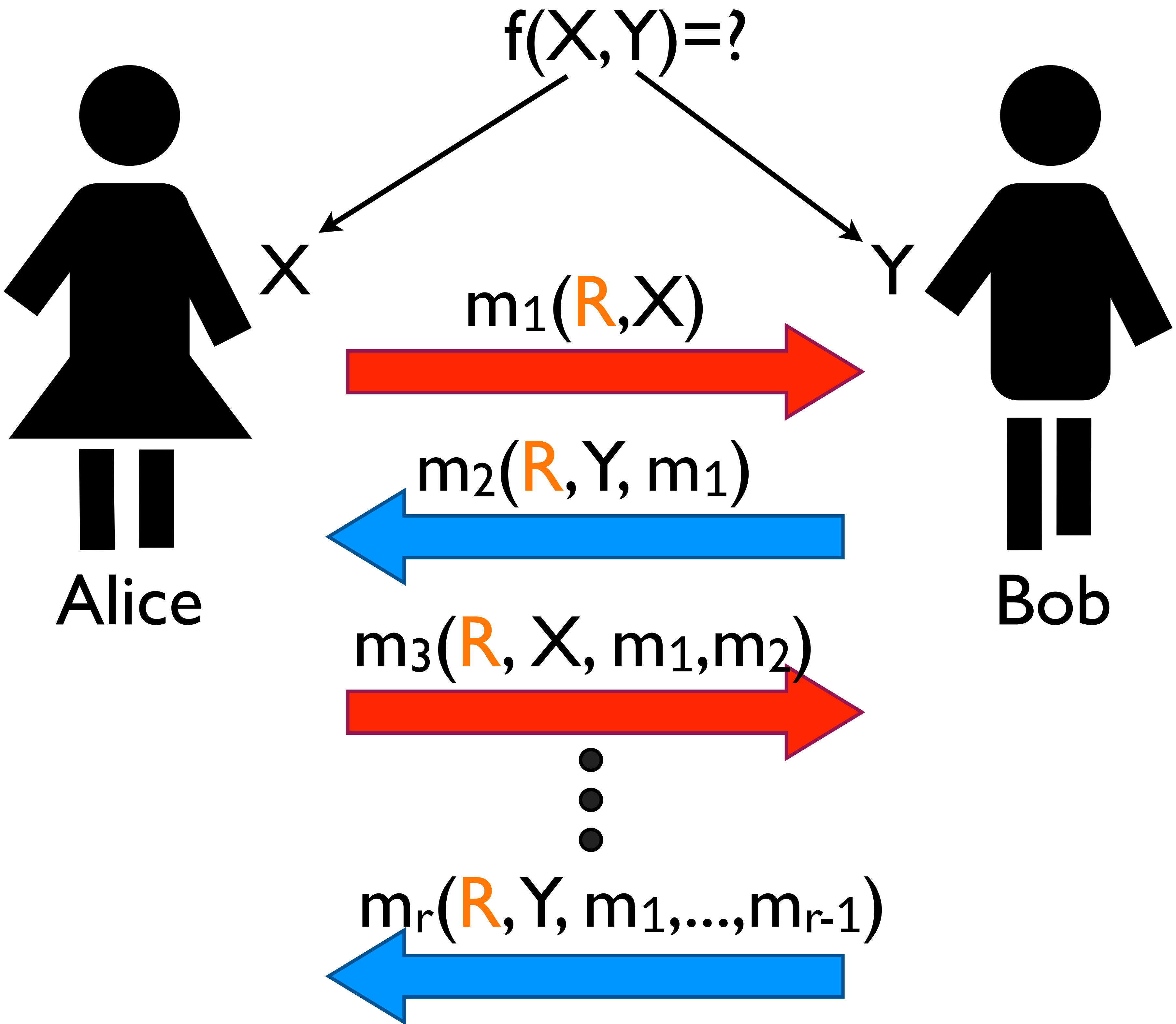
Communication complexity

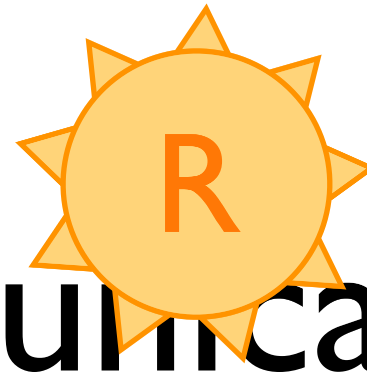


R: Shared random source



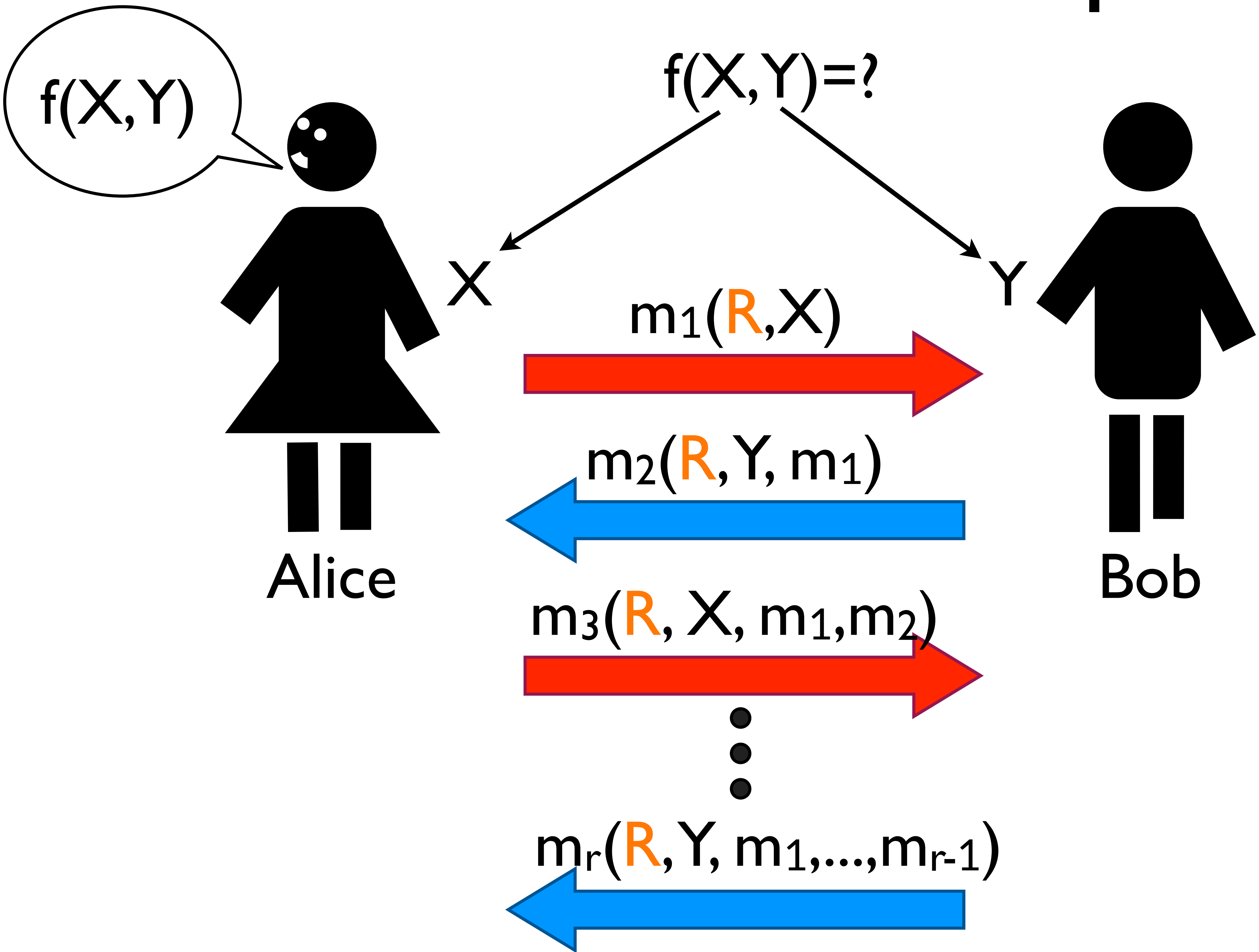
Communication complexity





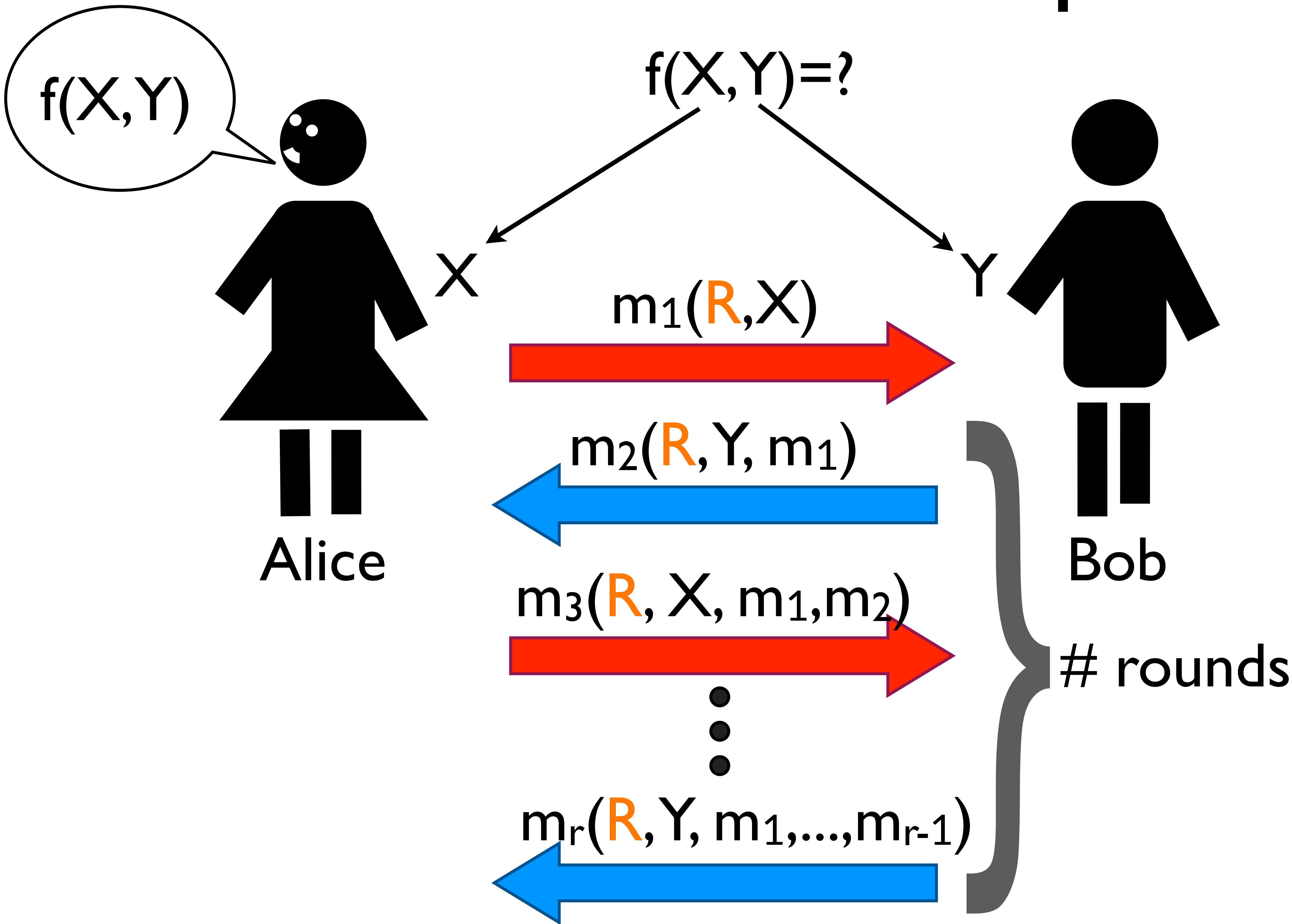
R: Shared random source

Communication complexity



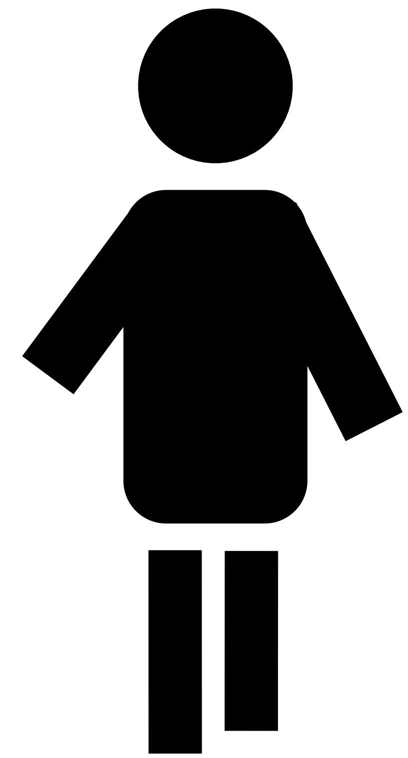
R: Shared random source

Communication complexity



Disjointness problem

$S=\{3,7,8,11\}$

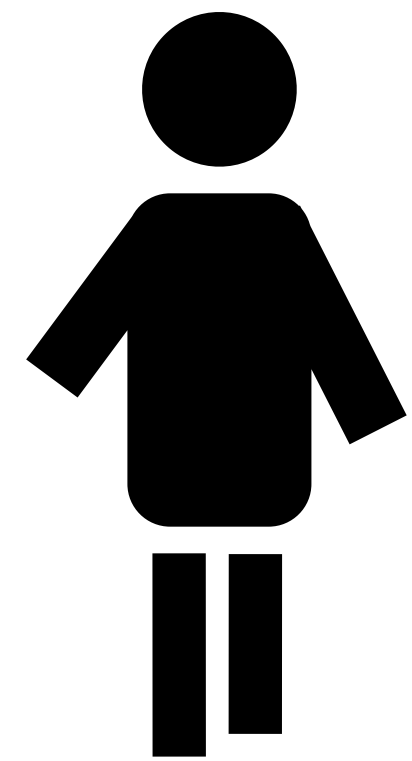


$T=\{2,5,8,14\}$

$S,T \subseteq [m]$

Disjointness problem

$$|S \cap T| \stackrel{?}{=} 0$$



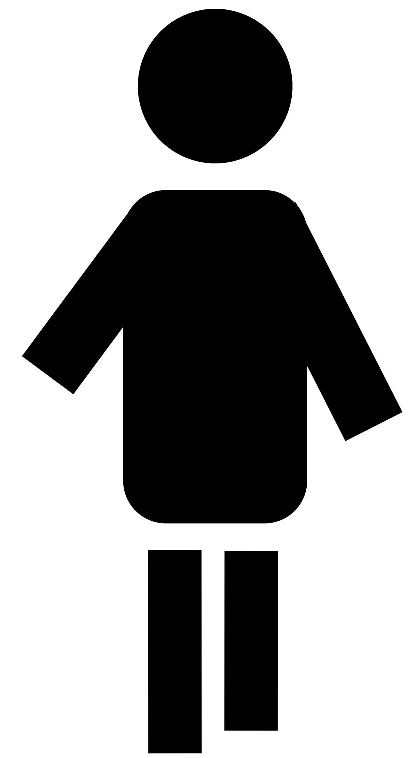
$S=\{3,7,8,11\}$

$T=\{2,5,8,14\}$

$S,T \subseteq [m]$

Disjointness problem

$|S \cap T| \stackrel{?}{=} 0$

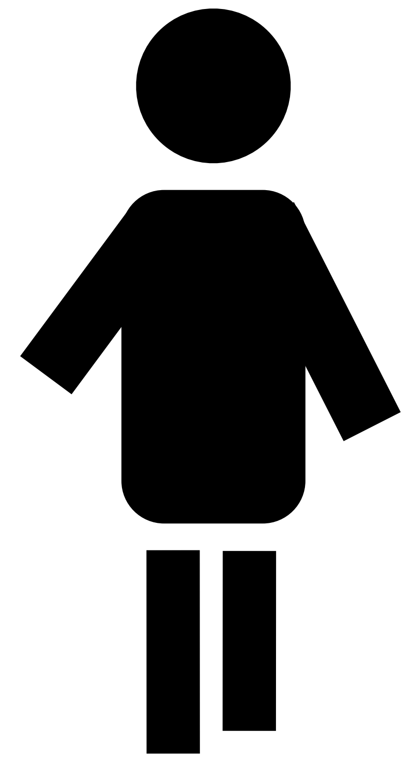


$S=\{3,7,8,11\}$ $T=\{2,5,8,14\}$

$S,T \subseteq [m]$

Disjointness problem

$|S \cap T| \stackrel{?}{=} 0$



$S=\{3,7,8,11\}$ $T=\{2,5,8,14\}$

$S,T \subseteq [m]$

In Sparse Set Disjointness DISJ_k^m **$|T|, |S| \leq k$**

Previous work

Total Bits	Rounds	Error	
$\Omega(k)$, $m \geq k^2$	Arbitrary	1/3	Babai, Frankl, Simon 86
$\Omega(k)$	Arbitrary	1/3	Kalyanasundaram, Schnitger 92, Razborov 92, Bar-Yossef et al. 02
$O(k \log k)$	1	1/k	Folklore
$\Omega(k \log k)$	1	1/3	Folklore, Buhrman et al. 13, Woodruff 08
$O(k)$	$O(\log k)$	0.01	Håstad, Wigderson 93

Our contributions

Bits	Rounds	Error	Best Previous
$O(k \log^{(r)} k)$	r	$1/\exp^{(r)}(c \log^{(r)} k)$	$O(k \log k)$, for $r=1$
$O(k)$	$\log^* k$	$\exp(-k^{1-\varepsilon})$	$O(\log k)$ rounds, 0.01 error
$\Omega(k \log^{(r)} k)$	r	1/3 error	$\Omega(k)$ $\Omega(k \log k)$, for $r=1$

Defn: $\exp^{(r)}(x) = 2^{2^{\cdot^{\cdot^{\cdot^{2^x}}}}}$

Our contributions

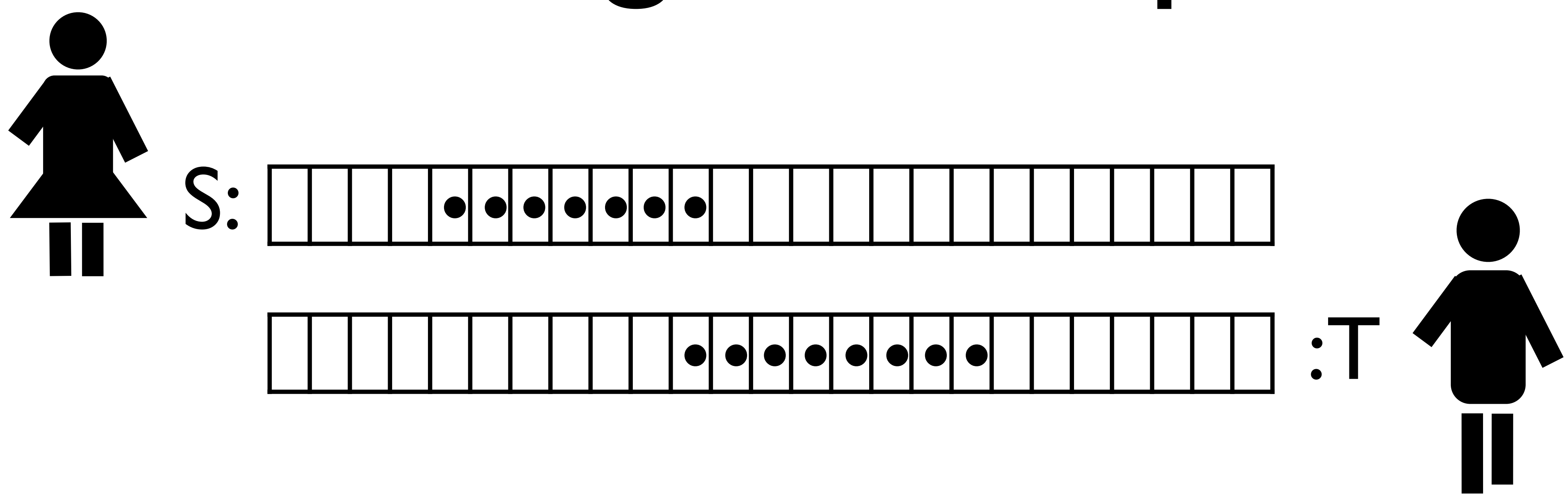
Bits	Rounds	Error	Best Previous
$O(k \log^{(r)} k)$	r	$1/\exp^{(r)}(c \log^{(r)} k)$	$O(k \log k)$, for $r=1$
$O(k)$	$\log^* k$	$\exp(-k^{1-\varepsilon})$	$O(\log k)$ rounds, 0.01 error
$\Omega(k \log^{(r)} k)$	r	1/3 error	$\Omega(k)$
			$\Omega(k \log k)$, for $r=1$

Holds for any $r \leq \log^* k$

Defn: $\exp^{(r)}(x) = 2^{2^{\cdot^{\cdot^{\cdot^{2^x}}}}}$

The upper bound

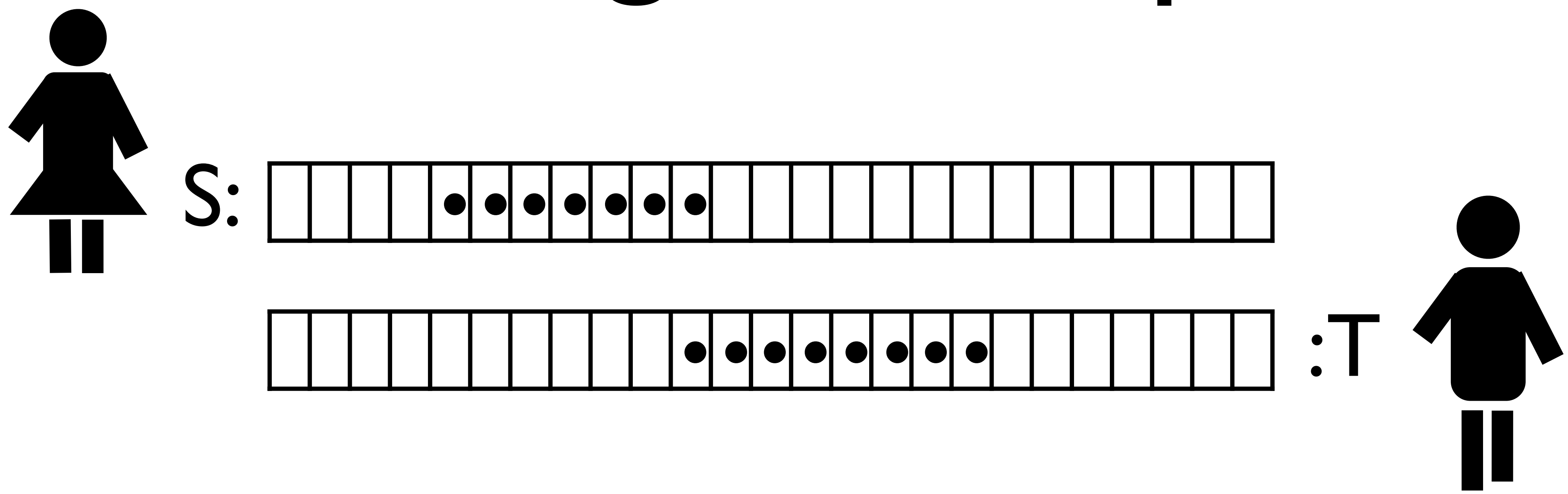
Håstad-Wigderson protocol



Z_1 Z_2 Z_3 $\cdots$ Z_k $\cdots$

- Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$

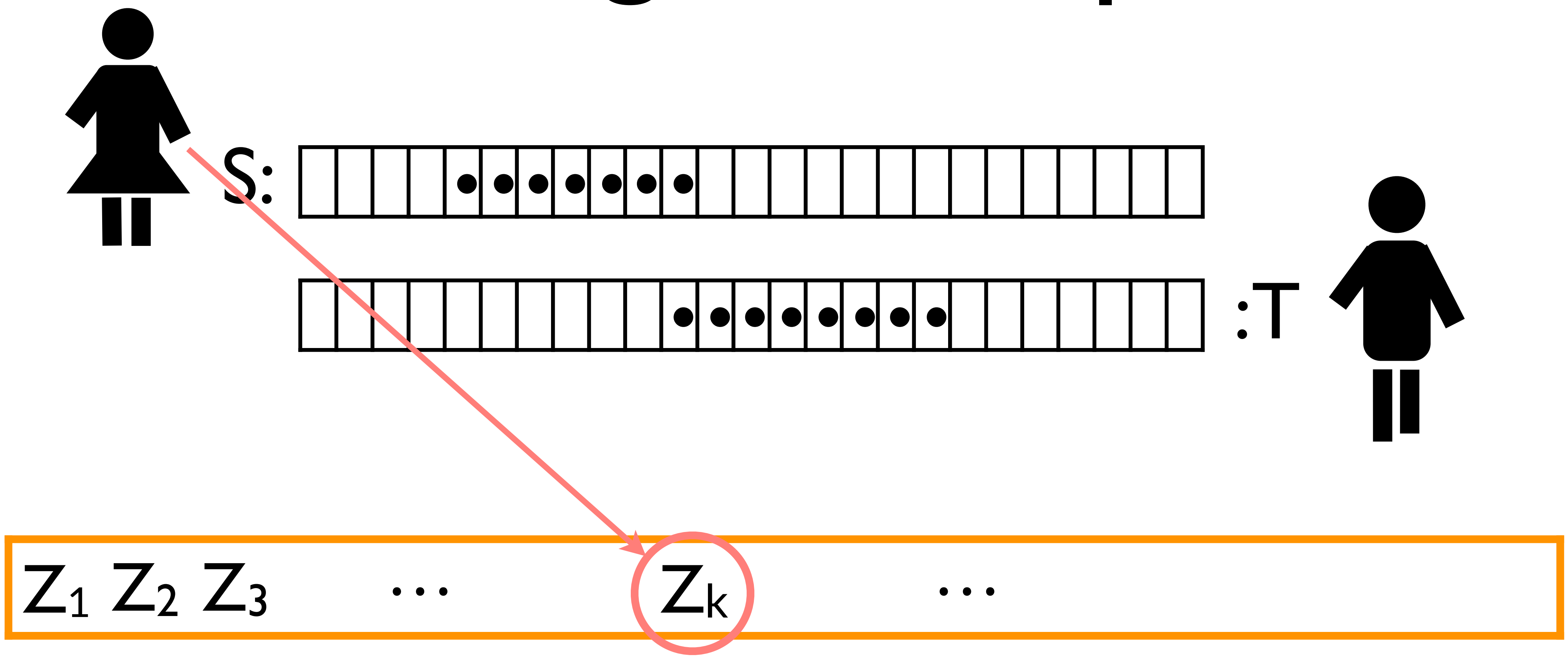
Håstad-Wigderson protocol



$Z_1\ Z_2\ Z_3\ \cdots\ Z_k\ \cdots$

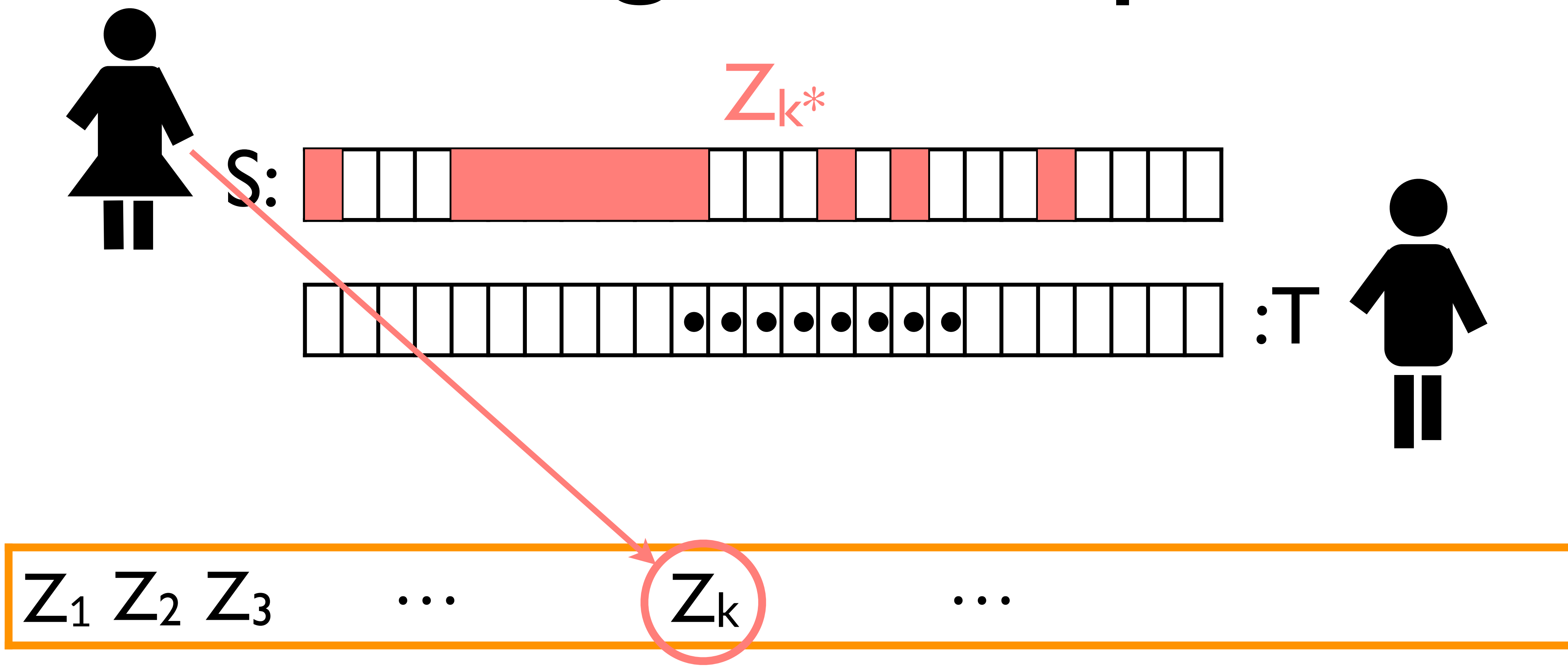
- Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$
- Finds the first k^*, $Z_{k^*} \supseteq S$

Håstad-Wigderson protocol



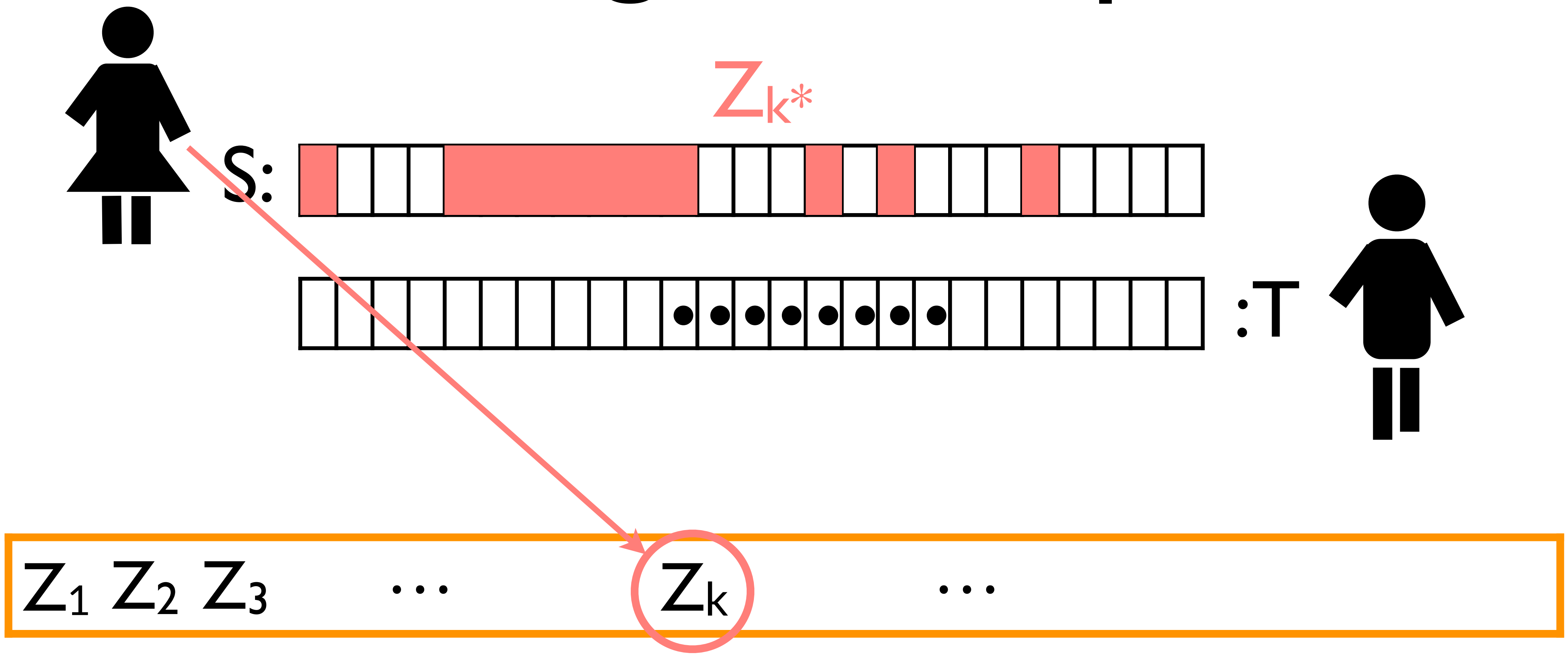
- Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$
- Finds the first k*, $Z_{k^*} \supseteq S$

Håstad-Wigderson protocol



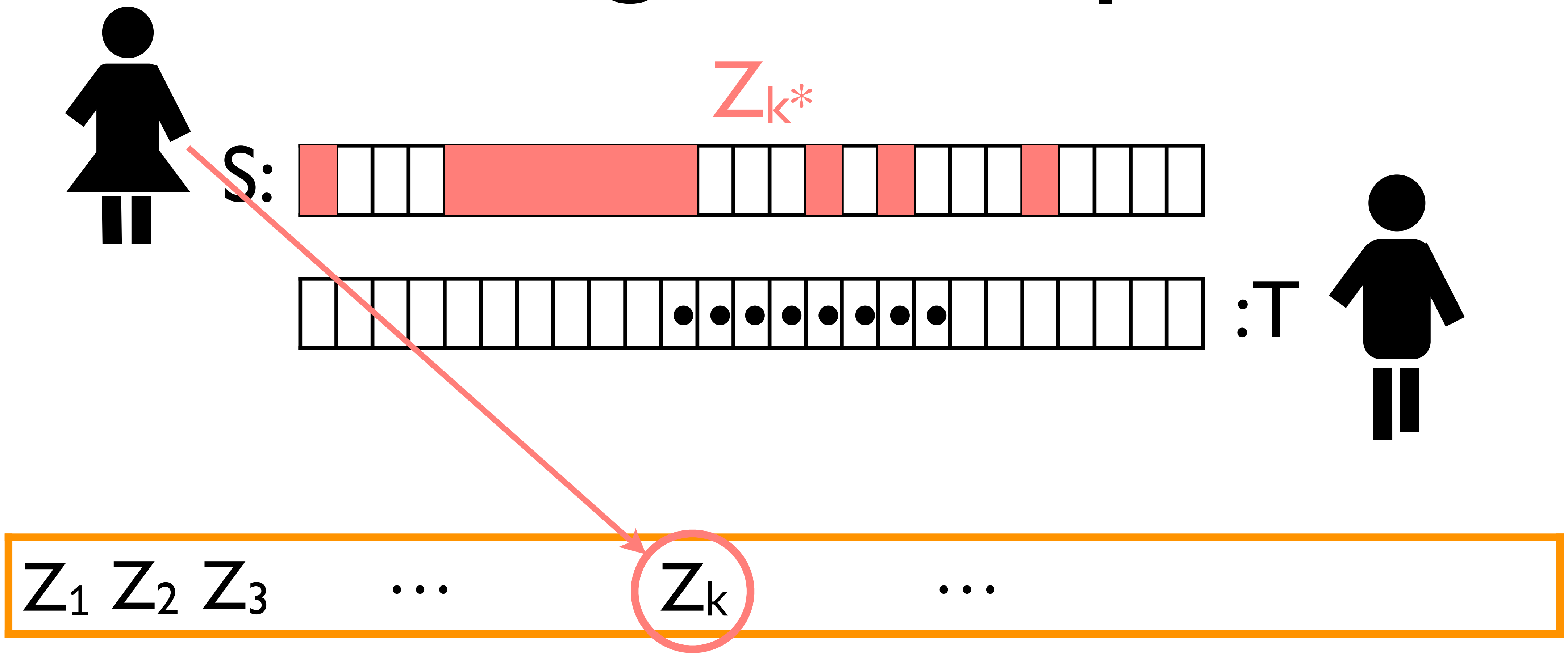
- Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$
- Finds the first k^*, $Z_{k^*} \supseteq S$

Håstad-Wigderson protocol



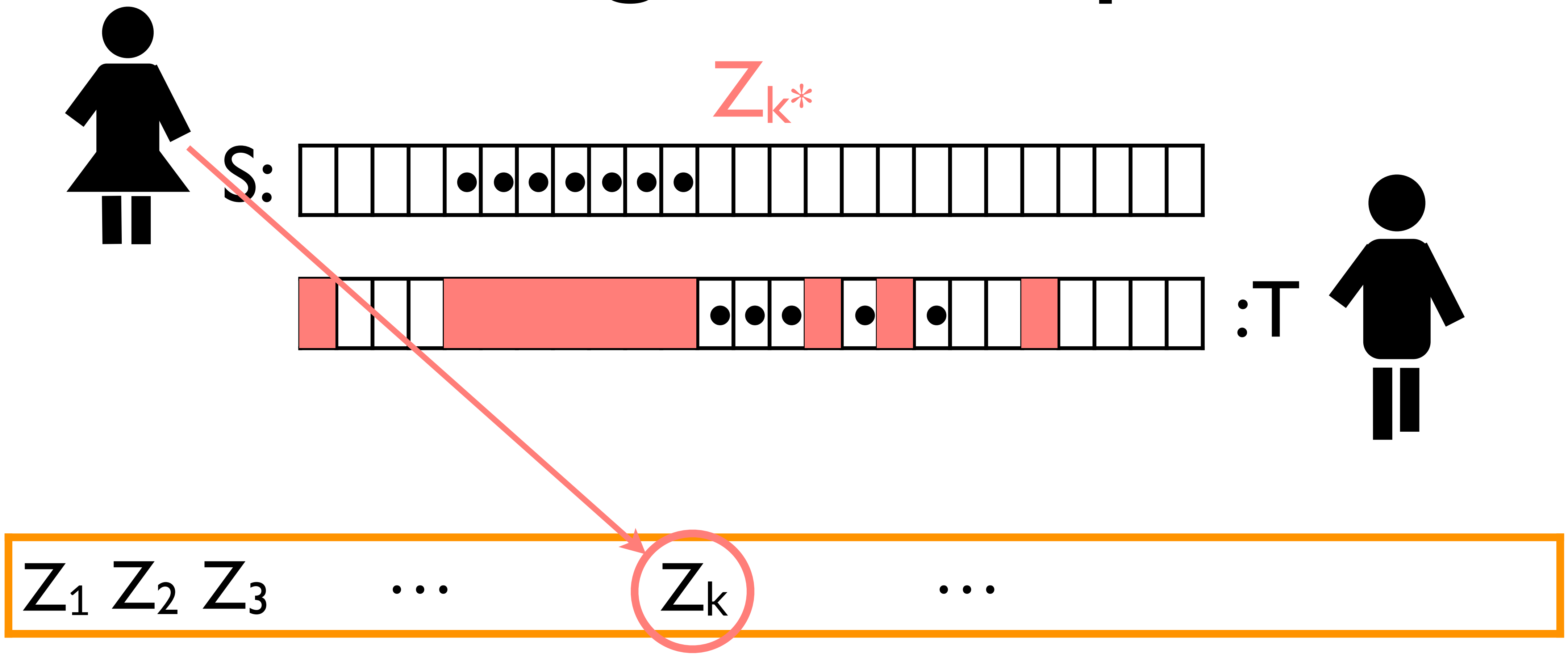
- Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$
- Finds the first k*, $Z_{k^*} \supseteq S$
- $\Pr[Z_i \supseteq S] = p^{|S|}$, so

Håstad-Wigderson protocol



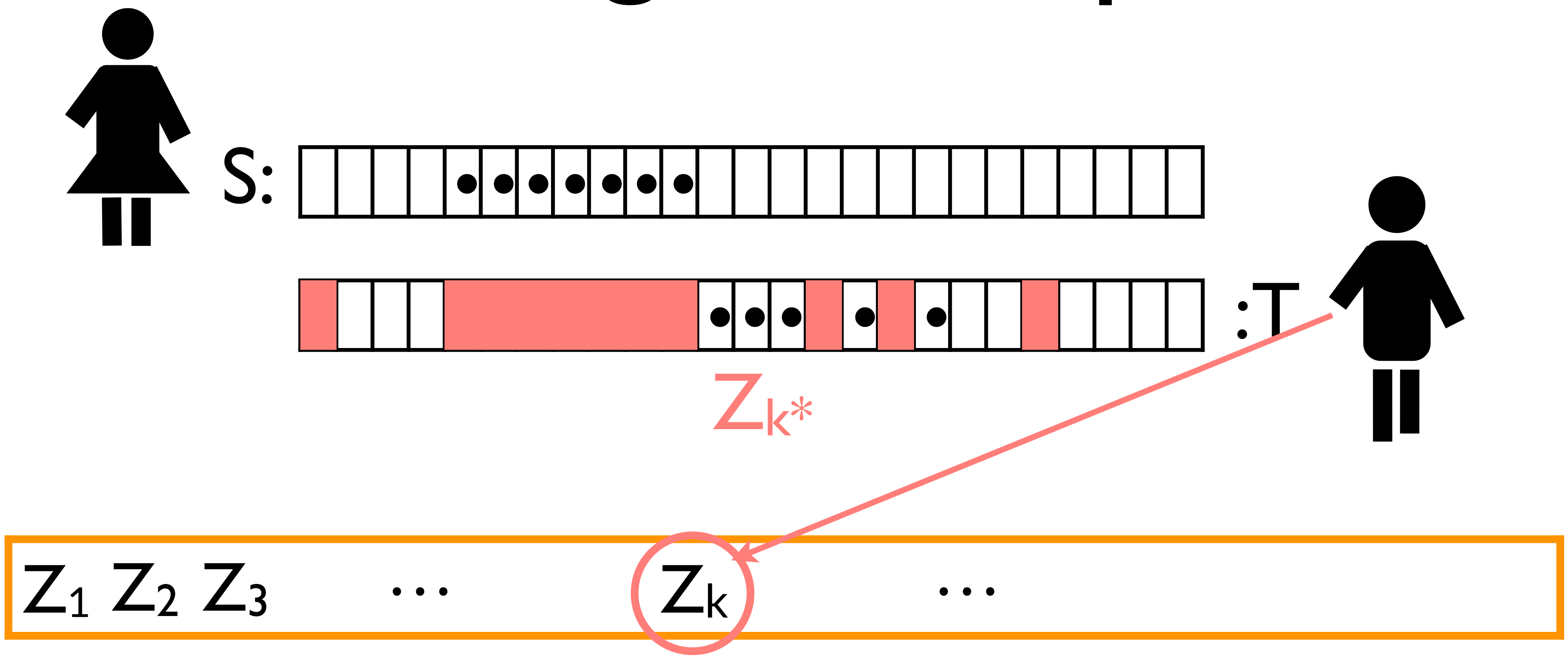
- Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$
- Finds the first k^*, $Z_{k^*} \supseteq S$
- $\Pr[Z_i \supseteq S] = p^{|S|}$, so $E[\, k^* \,] = 1/p^{|S|}$

Håstad-Wigderson protocol



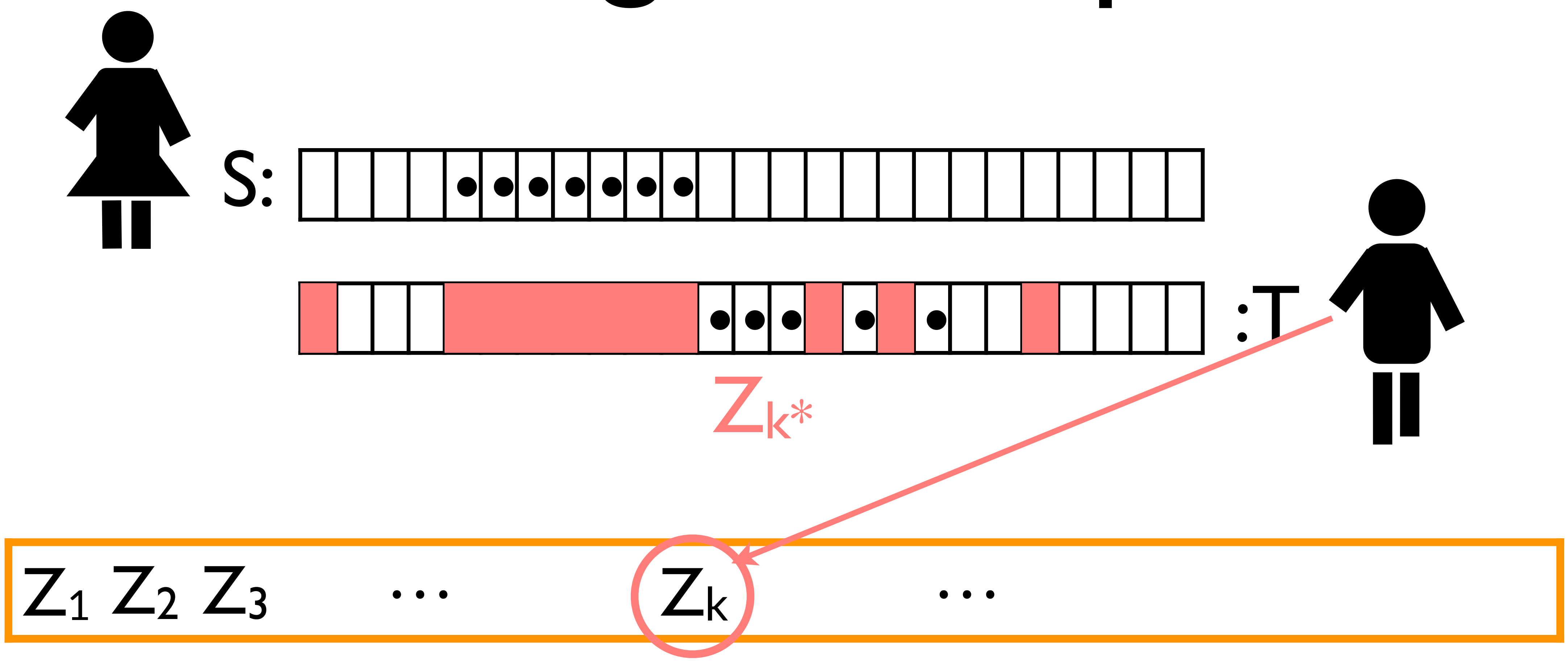
- Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$
- Finds the first k^*, $Z_{k^*} \supseteq S$
- $\Pr[Z_i \supseteq S] = p^{|S|}$, so $E[\, k^* \,] = 1/p^{|S|}$
- Send k^* to Bob: $|S| \log 1/p$ bits

Håstad-Wigderson protocol



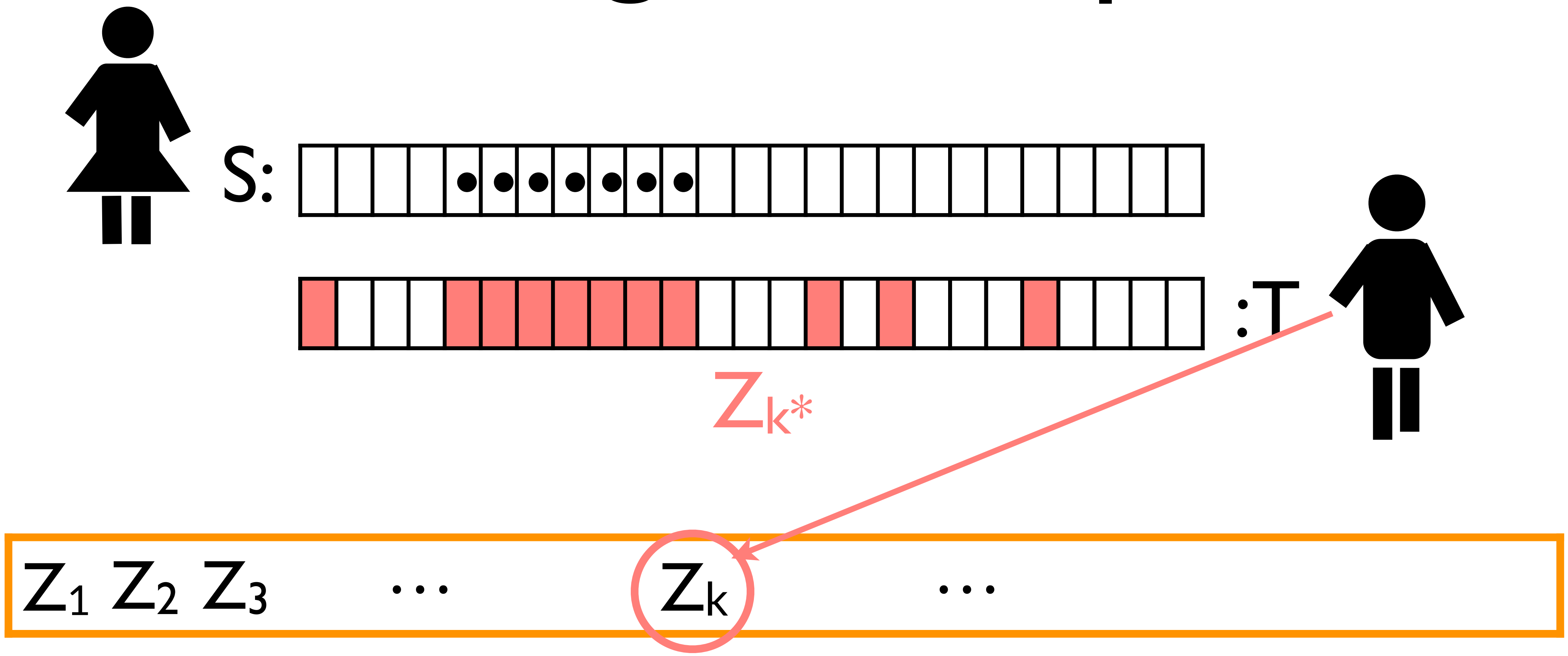
- If $a \in S \cap T \Rightarrow a \in Z_{k^*}$,

Håstad-Wigderson protocol



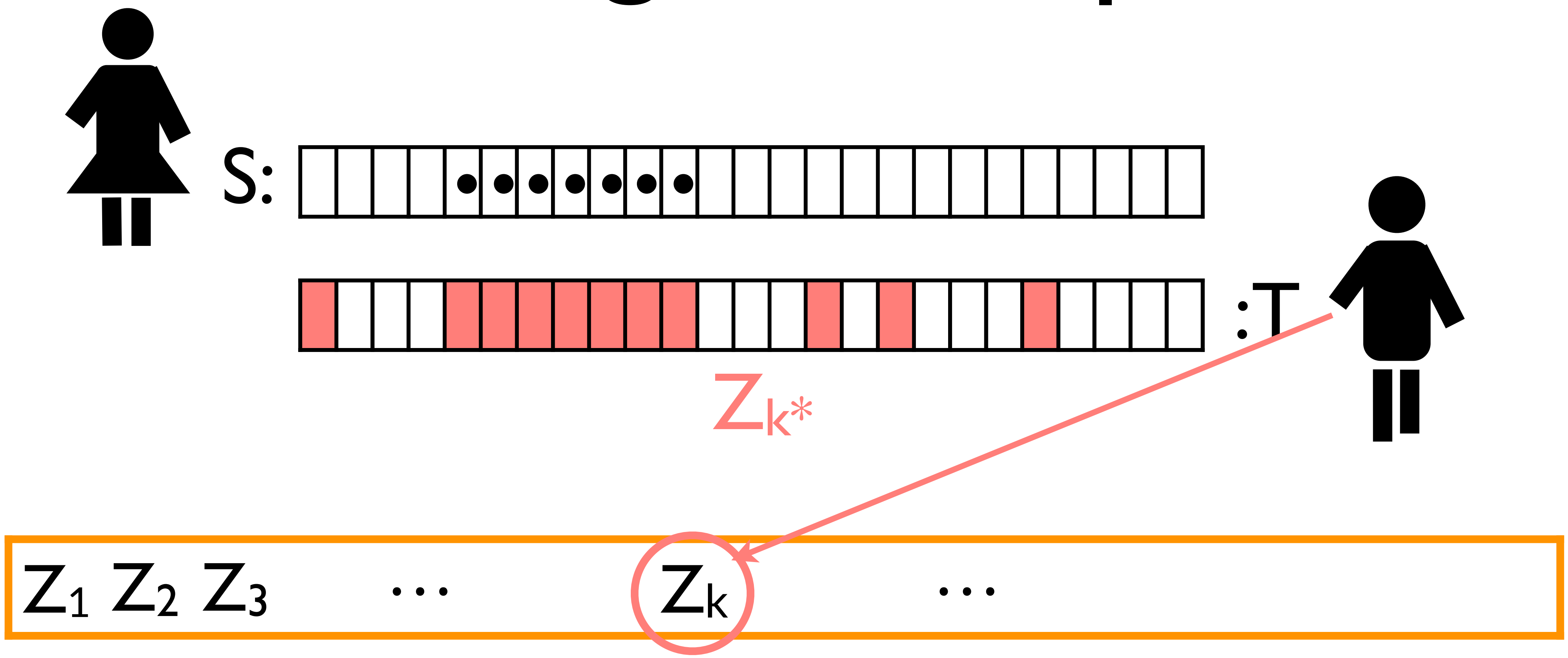
- If $a \in S \cap T \Rightarrow a \in Z_{k^*}$, so set $T' = T \cap Z_{k^*}$

Håstad-Wigderson protocol



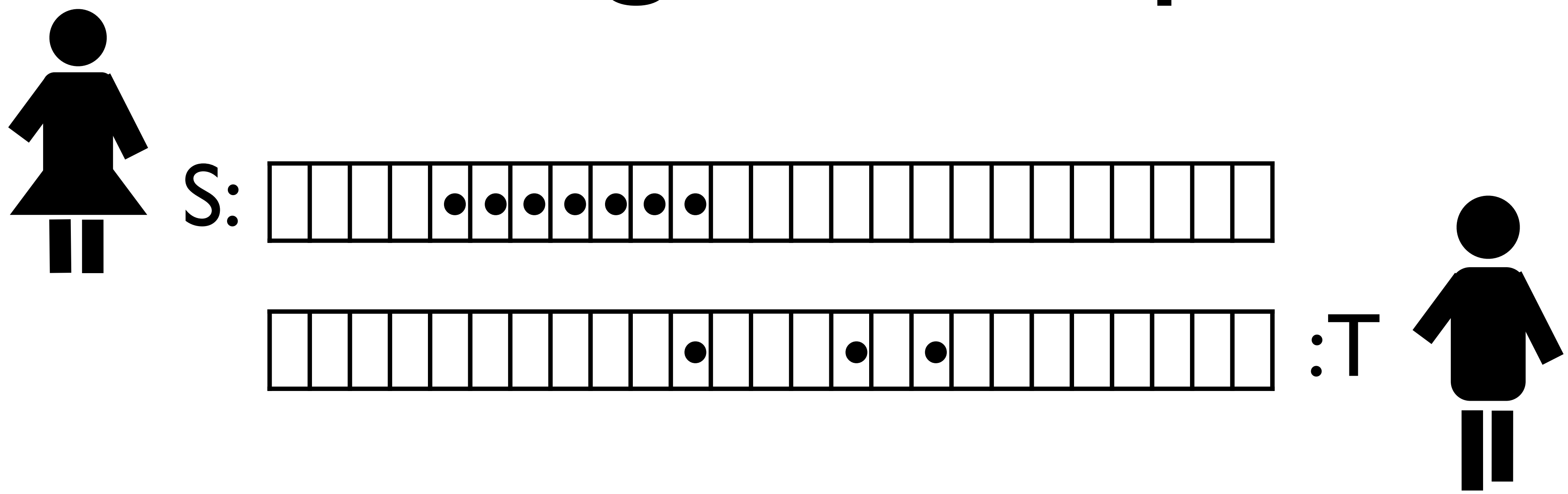
- If $a \in S \cap T \Rightarrow a \in Z_{k^*}$, so set $T' = T \cap Z_{k^*}$

Håstad-Wigderson protocol



- If $a \in S \cap T \Rightarrow a \in$ Z_{k^*}, so set $T' = T \cap Z_{k^*}$
- If $S \cap T = \varnothing$, $E[|T'|] = p|T|$

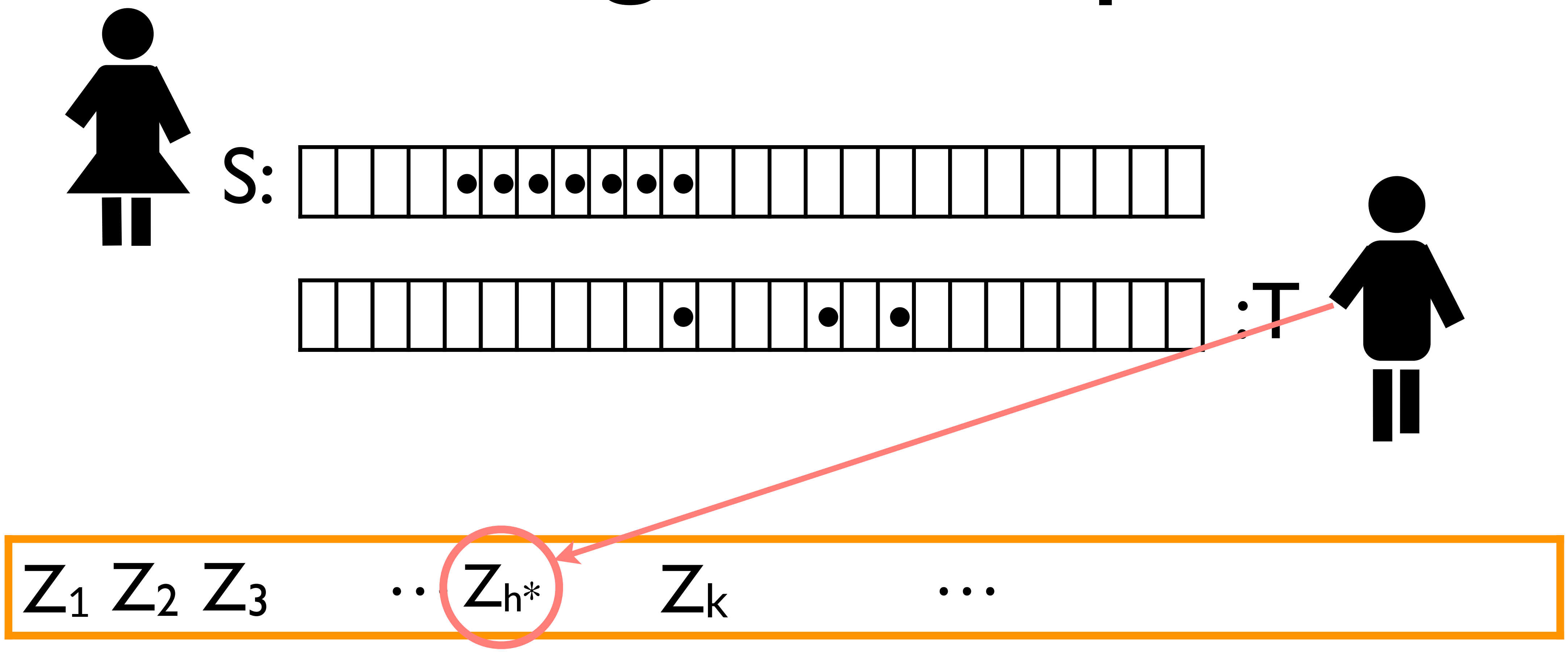
Håstad-Wigderson protocol



$Z_1\ Z_2\ Z_3 \quad \cdots \quad Z_k \quad \cdots$

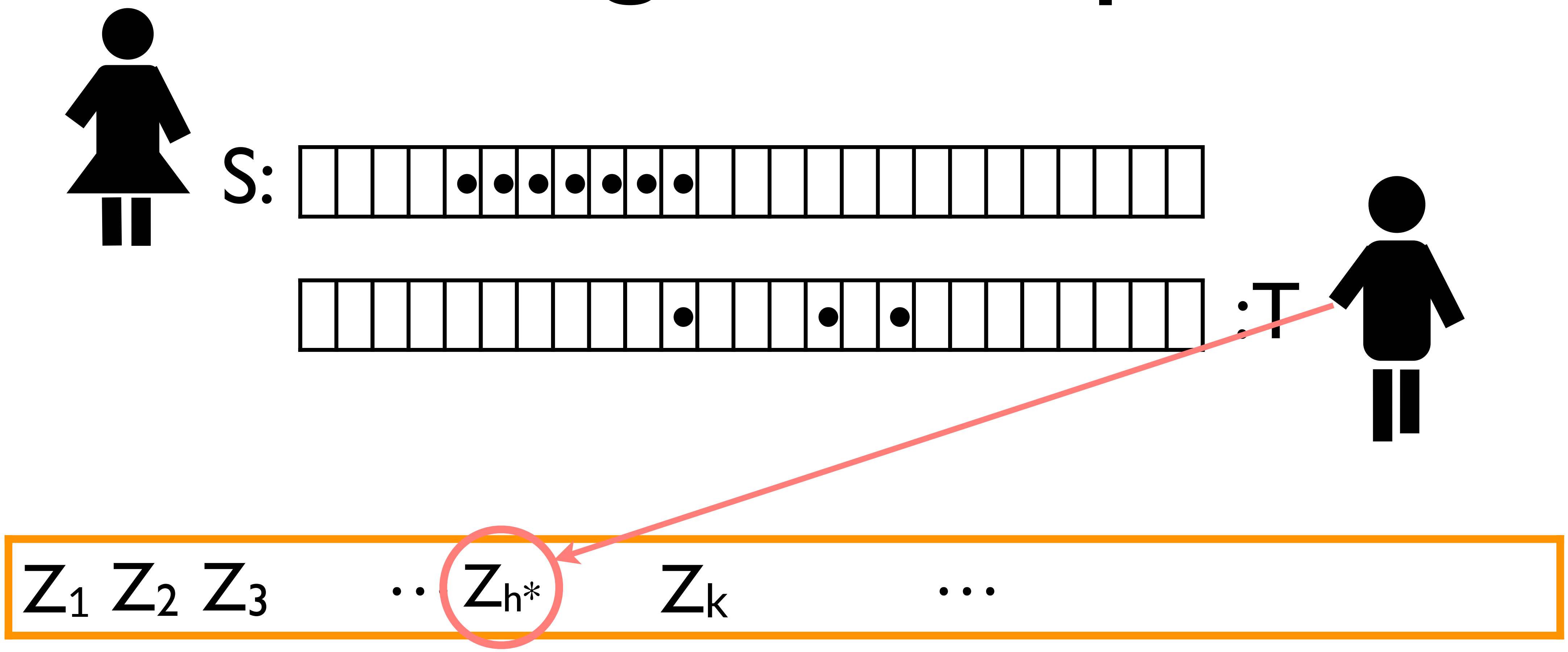
- If $a \in S \cap T \Rightarrow a \in Z_{k^*}$, so set $T' = T \cap Z_{k^*}$
- If $S \cap T = \varnothing$, $E[|T'|] = p|T|$
- Bob repeats for T'

Håstad-Wigderson protocol



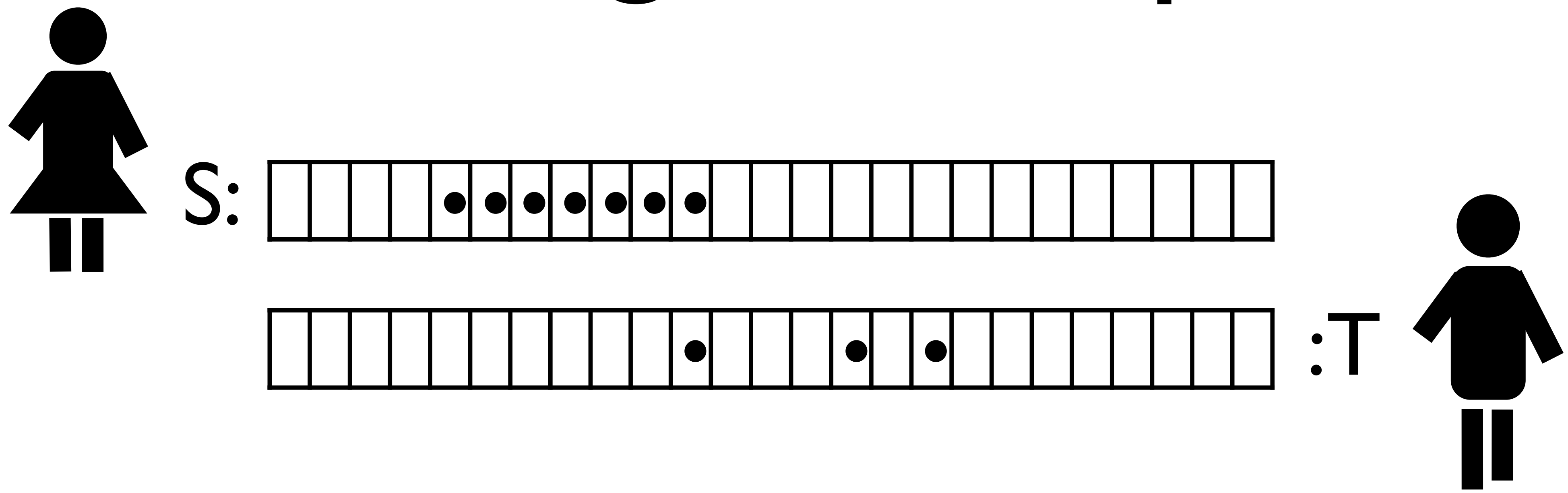
- If $a \in S \cap T \Rightarrow a \in Z_{k^*}$, so set $T' = T \cap Z_{k^*}$
- If $S \cap T = \emptyset$, $E[|T'|] = p|T|$
- Bob repeats for T' $Z_{h^*} \supseteq T'$

Håstad-Wigderson protocol



- If $a \in S \cap T \Rightarrow a \in Z_{k*}$, so set $T' = T \cap Z_{k*}$
- If $S \cap T = \varnothing$, $E[|T'|] = p|T|$
- Bob repeats for T' $\quad Z_{h*} \supseteq T'$
- For $p=1/2$, if $S \cap T = \varnothing$, in $O(\log k)$ rounds $S' = T' = \varnothing$

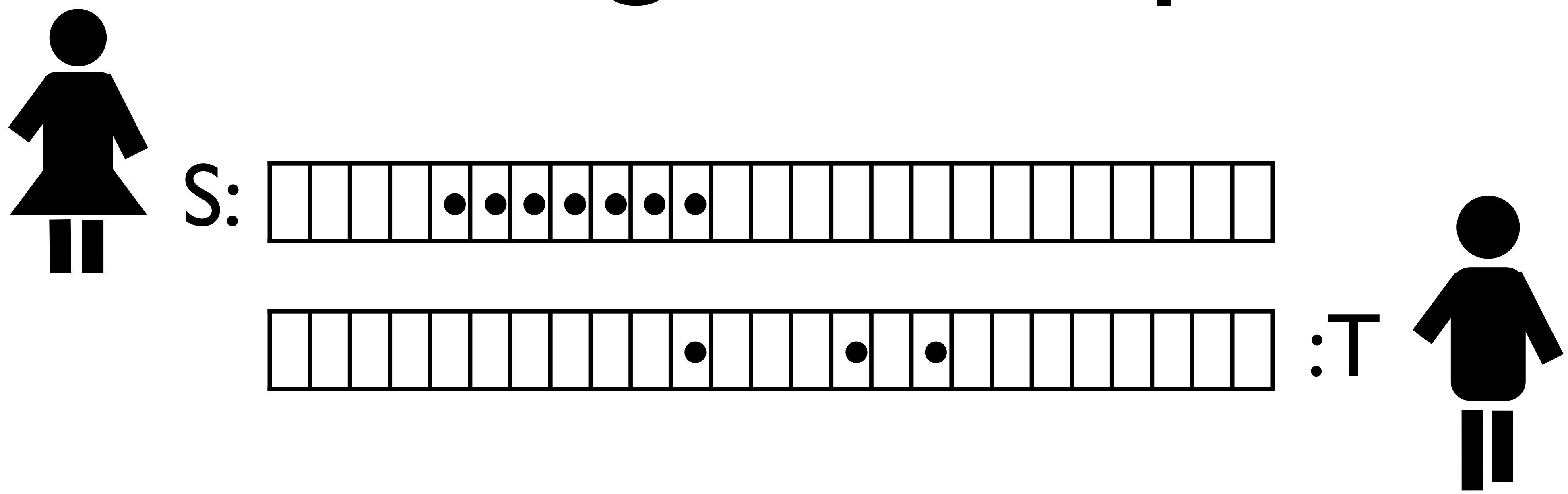
Håstad-Wigderson protocol



Run O(log k) rounds

- ➡ If S'=T'=∅, declare DISJOINT
- ➡ Otherwise, declare INTERSECT

Håstad-Wigderson protocol



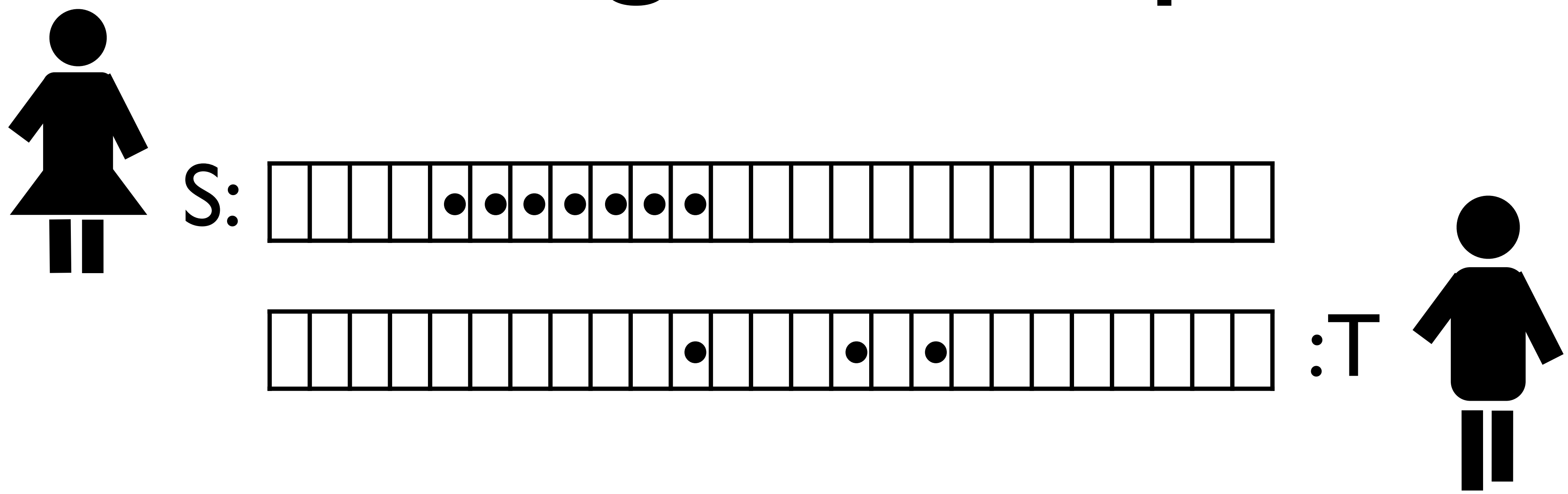
Run O(log k) rounds

➡ If S'=T'=∅, declare DISJOINT

➡ Otherwise, declare INTERSECT

Cost: |S'| bits per round

Håstad-Wigderson protocol



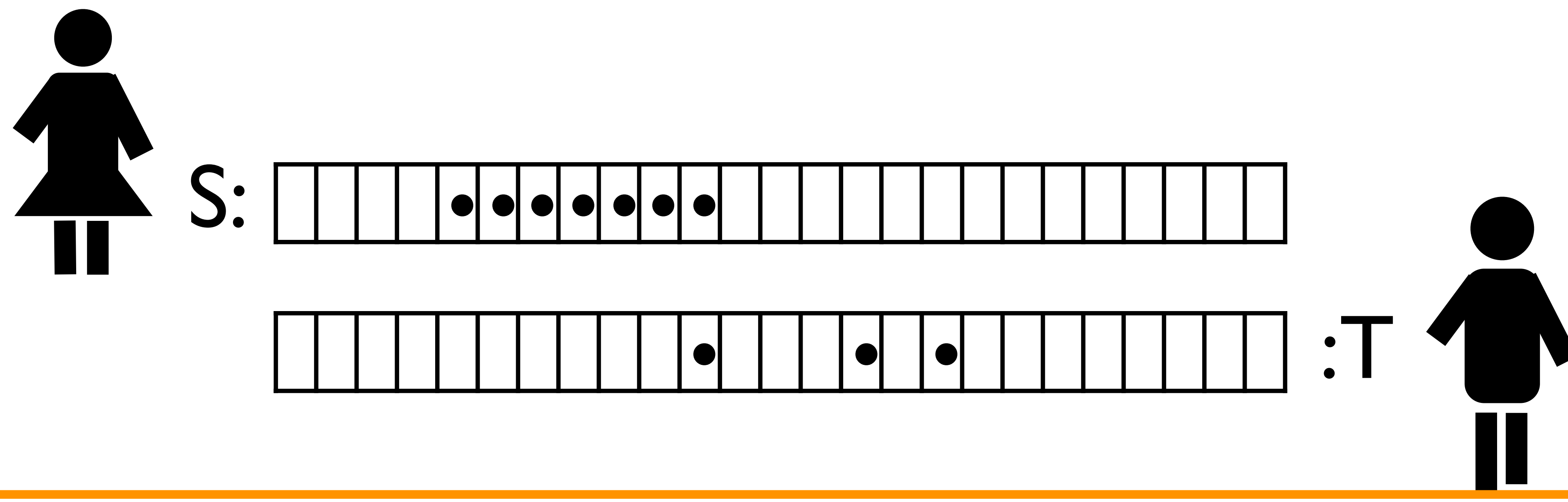
Run O(log k) rounds

- ➡ If S'=T'=∅, declare DISJOINT
- ➡ Otherwise, declare INTERSECT

Cost: |S'| bits per round

Total = $k + k/2 + k/4 + \cdots = O(k)$

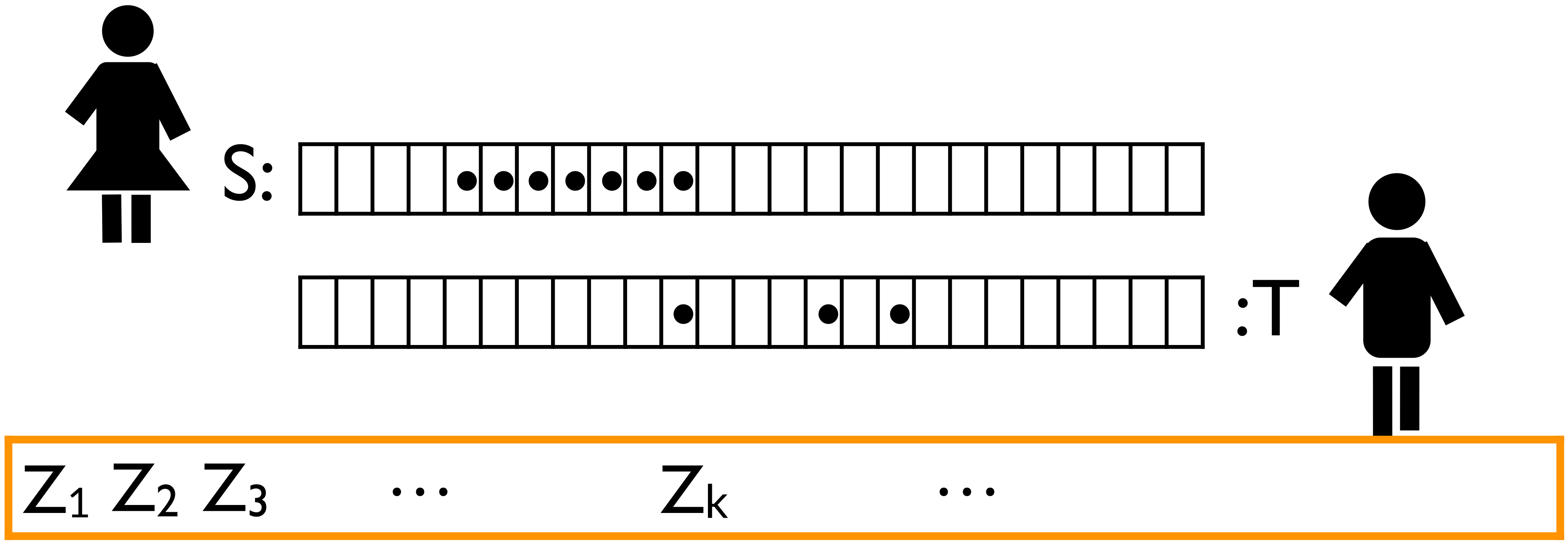
Our bound: $R^r(DISJ_k) \leq O(k \log^{(r)} k)$



Z_1 Z_2 Z_3 $\cdots$ Z_k $\cdots$

Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$

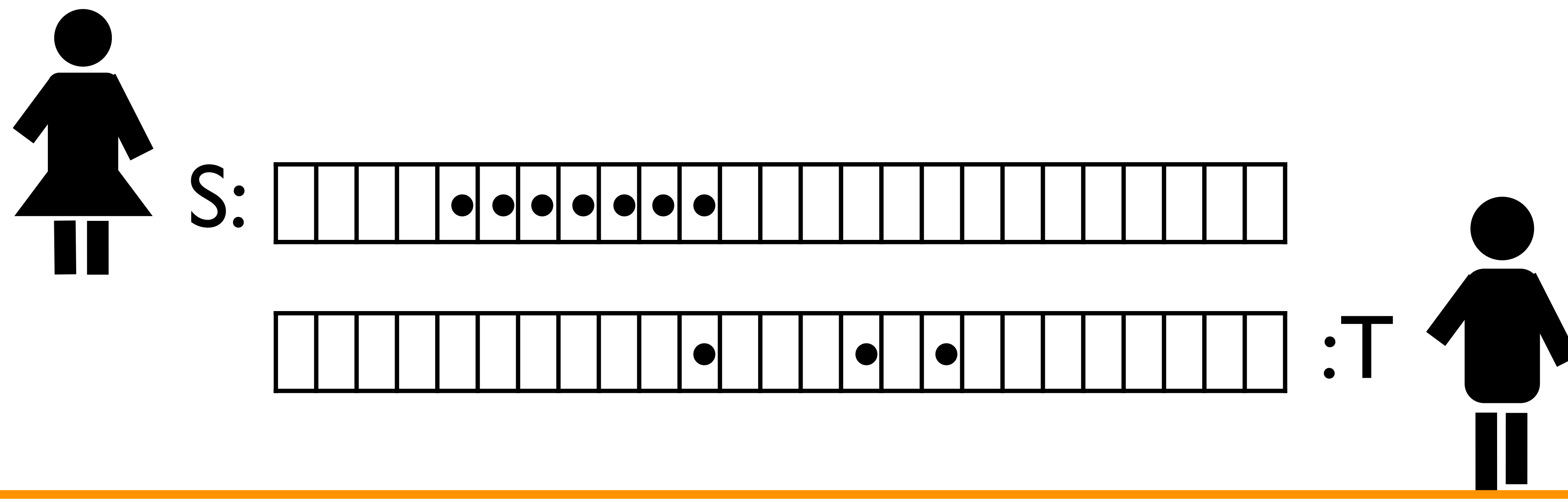
Our bound: $R^r(\mathrm{DISJ}_k) \leq O(k \log^{(r)} k)$



Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$

Bits per round: $|S'| \log 1/p$

Our bound: $R^r(\mathrm{DISJ}_k) \le O(k \log^{(r)} k)$



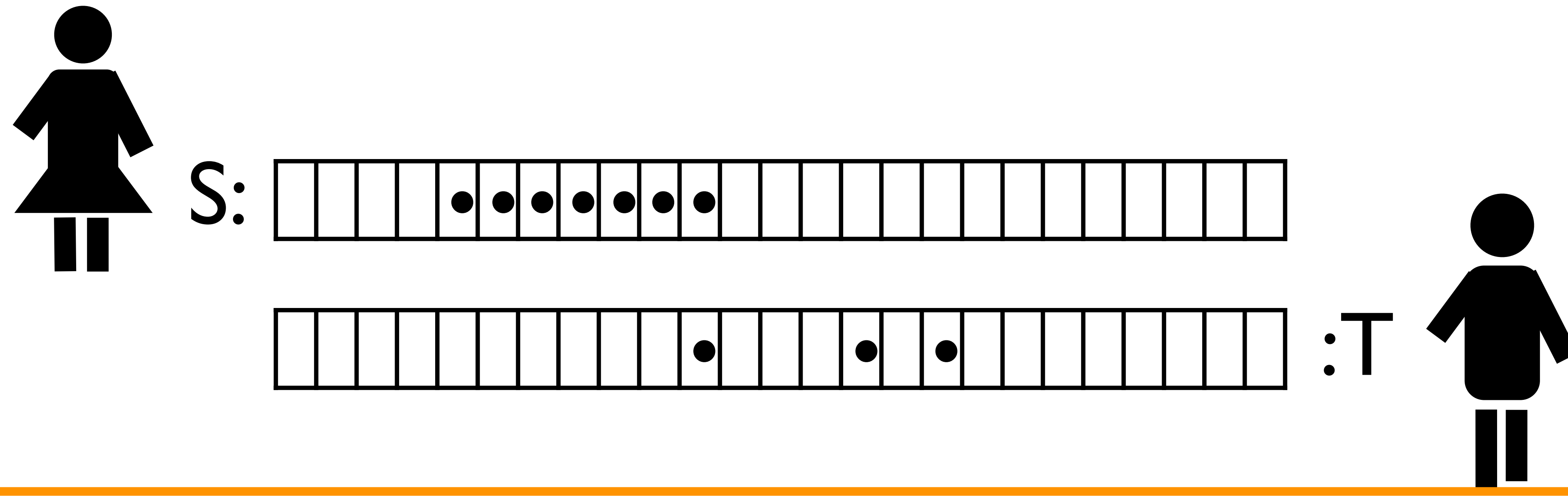
$Z_1\ Z_2\ Z_3\ \cdots\ Z_k\ \cdots$

Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$

Bits per round: $|S'| \log 1/p$

Observation: Sets get smaller $\Rightarrow$ can afford smaller p each round

Our bound: $R^r(DISJ_k) \leq O(k \log^{(r)} k)$



$Z_1\ Z_2\ Z_3 \quad \cdots \quad Z_k \quad \cdots$

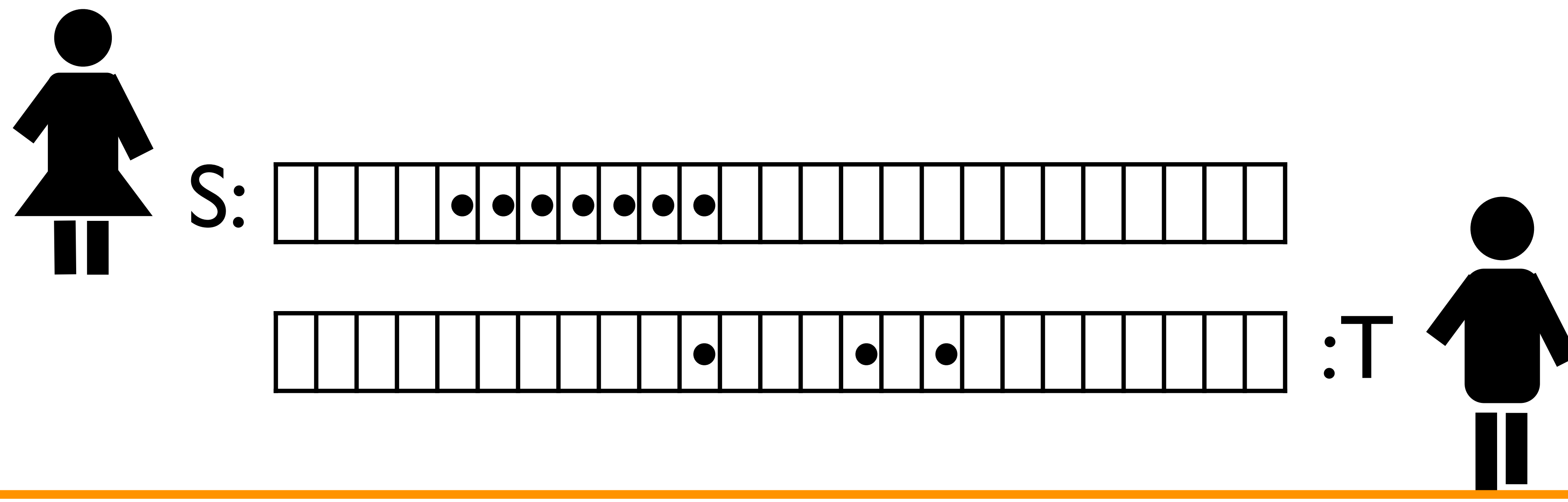
Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$

Bits per round: $|S'| \log 1/p$

Observation: Sets get smaller $\Rightarrow$ can afford smaller p each round

Defn: $\exp^{(r)}(x) = 2^{2^{\cdot^{\cdot^{\cdot^{2^x}}}}}$

Our bound: $R^r(DISJ_k) \leq O(k \log^{(r)} k)$



$Z_1 \; Z_2 \; Z_3 \quad \cdots \quad Z_k \quad \cdots$

Z_i: random, $\forall a \in [m]$, $\Pr[a \in Z_i] = p$

Bits per round: $|S'| \log 1/p$

Observation: Sets get smaller $\Rightarrow$ can afford smaller p each round

Defn: $\exp^{(r)}(x) = 2^{2^{\cdot^{\cdot^{\cdot^{2^x}}}}}$ $\qquad p_i = 1/\exp^{(i)}(5 \log^{(r)} k)$

Our bound: $R^r(\mathrm{DISJ}_k) \leq O(k \log^{(r)} k)$

Run r rounds

- ➡ If $S' = T' = \varnothing$, declare DISJOINT
- ➡ Otherwise, declare INTERSECT

Our bound: $R^r(DISJ_k) \leq O(k \log^{(r)} k)$

Run r rounds

- ➡ If $S'=T'=\varnothing$, declare DISJOINT
- ➡ Otherwise, declare INTERSECT

Fact 1

- $p_i = 1/\exp^{(i)}(5 \log^{(r)} k)$
- if $S \cap T=\varnothing$, in r rounds $S'=T'=\varnothing$

Our bound: $R^r(DISJ_k) \leq O(k \log^{(r)} k)$

Run r rounds

➡ If $S'=T'=\emptyset$, declare DISJOINT

➡ Otherwise, declare INTERSECT

Fact 1

- $p_i = 1/\exp^{(i)}(5 \log^{(r)} k)$
- if $S \cap T=\emptyset$, in r rounds $S'=T'=\emptyset$

Fact 2 $|message_i| \leq 5k \log^{(r)} k / 2^i$

Our bound: $R^r(\mathrm{DISJ}_k) \leq O(k \log^{(r)} k)$

Run r rounds

- If $S' = T' = \varnothing$, declare DISJOINT
- Otherwise, declare INTERSECT

Fact 1

- $p_i = 1/\exp^{(i)}(5 \log^{(r)} k)$
- if $S \cap T = \varnothing$, in r rounds $S' = T' = \varnothing$

Fact 2 $|\mathrm{message}_i| \leq 5k \log^{(r)} k / 2^i$

For $i > 3$, $|\mathrm{message}_i| \leq \prod_{t=1}^{i/2} p_{i-2t+1} \; |S| \log 1/p_i$

$$\leq \frac{k}{\exp^{(i-1)} \exp^{(i-3)}} \log \exp^{(i)} \leq k / 2^i$$

The lower bound

Exists-equal problem

- Stronger lower bound: easier problem exists-equal (EE)

Exists-equal problem

- Stronger lower bound: easier problem exists-equal (EE)
- Let $x,y \in [t]^n$

Exists-equal problem

- Stronger lower bound: easier problem exists-equal (EE)
- Let $x,y \in [t]^n$
- $EE_n^t(x, y) = 1$ iff $\exists i, x_i = y_i$

Exists-equal problem

- Stronger lower bound: easier problem exists-equal (EE)
- Let $x,y \in [t]^n$
- $EE_n^t(x, y) = 1$ iff $\exists i, x_i = y_i$

x:	3	4	4	5	1
y:	2	3	4	2	4

Exists-equal problem

- Stronger lower bound: easier problem exists-equal (EE)
- Let $x,y \in [t]^n$
- $EE_n^t(x, y) = 1$ iff $\exists i$, $x_i = y_i$

x:	3	4	4	5	1
y:	2	3	4	2	4

So EE(x,y)=1

Exists-equal problem

- EE^t_n is OR of n equality problems over [t].
- $EE^t_n = DISJ^{tn}_n$

Exists-equal problem

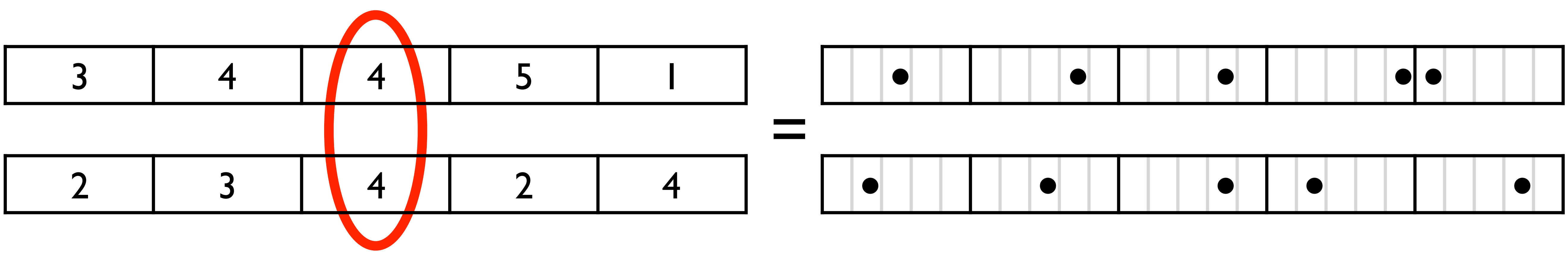
- EE^t_n is OR of n equality problems over [t].
- $EE^t_n = DISJ^{tn}_n$

3	4	4	5	1

2	3	4	2	4

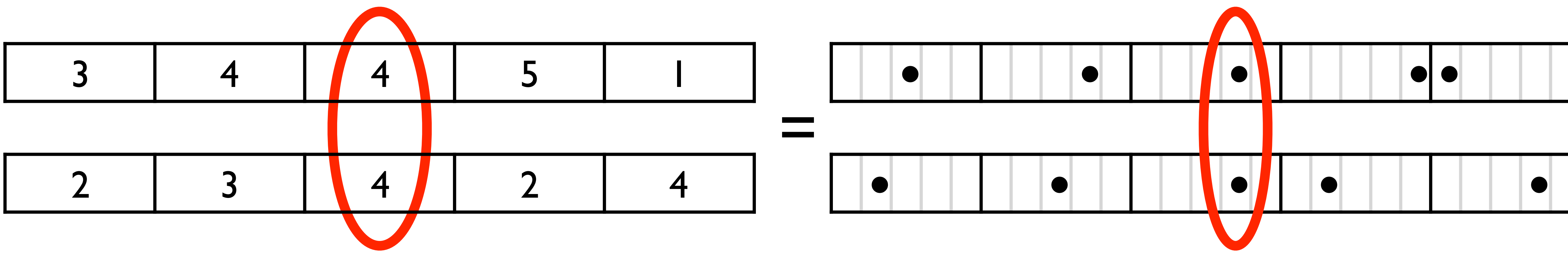
Exists-equal problem

- EE_n^t is OR of n equality problems over [t].
- $EE_n^t = DISJ_n^{tn}$



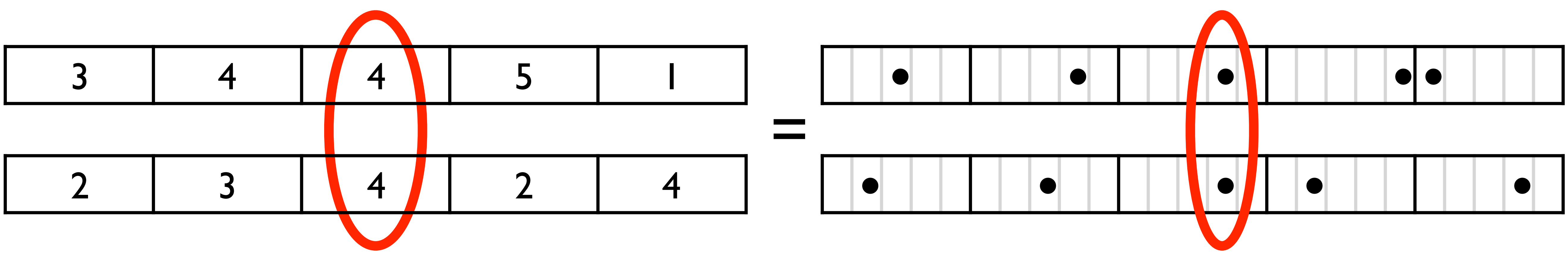
Exists-equal problem

- EE^t_n is OR of n equality problems over [t].
- $EE^t_n = DISJ^{tn}_n$



Exists-equal problem

- EE^t_n is OR of n equality problems over [t].
- $EE^t_n = DISJ^{tn}_n$



$|S| = |T| = n$

$S, T \subset [nt]$

The lower bound

- Show: any r-round EE protocol communicates $\Omega(n \log^{(r)} n)$ bits.

The lower bound

- Show: any r-round EE protocol communicates $\Omega(n \log^{(r)} n)$ bits.
- EE is OR of n equality problems, so decompose to subproblems

The lower bound

- Show: any r-round EE protocol communicates $\Omega(n \log^{(r)} n)$ bits.
- EE is OR of n equality problems, so decompose to subproblems
- Show $\Omega(\log^{(r)} n)$ lower bound per subproblem

The lower bound

- Show: any r-round EE protocol communicates $\Omega(n \log^{(r)} n)$ bits.
- EE is OR of n equality problems, so decompose to subproblems
- Show $\Omega(\log^{(r)} n)$ lower bound per subproblem
- Equality has $O(1)$ bit communication protocol

The lower bound

- Show: any r-round EE protocol communicates $\Omega(n \log^{(r)} n)$ bits.
- EE is OR of n equality problems, so decompose to subproblems
- ~~Show $\Omega(\log^{(r)} n)$ lower bound per subproblem~~
- Equality has $O(1)$ bit communication protocol

The lower bound

- Show: any r-round EE protocol communicates $\Omega(n \log^{(r)} n)$ bits.
- EE is OR of n equality problems, so decompose to subproblems
- ~~Show $\Omega(\log^{(r)} n)$ lower bound per subproblem~~
- Equality has $O(1)$ bit communication protocol
- We get **super-linear** increase in complexity!

The lower bound

By Yao's lemma, sufficient to consider deterministic protocols with random input

The lower bound

By Yao's lemma, sufficient to consider deterministic protocols with random input

- Set t=4n

The lower bound

By Yao's lemma, sufficient to consider deterministic protocols with random input

- Set $t=4n$
- Hard distribution ν: $(x,y)\in [t]^n \times [t]^n$, uniform random

The lower bound

By Yao's lemma, sufficient to consider deterministic protocols with random input

- Set $t=4n$
- Hard distribution ν: $(x,y) \in [t]^n \times [t]^n$, uniform random
- $3/4 \leqslant \Pr_{(x,y)\sim\nu}[EE(x,y)=0] = (1-1/(4n))^n \leqslant e^{-1/4} < 0.78$

The lower bound

By Yao's lemma, sufficient to consider deterministic protocols with random input

- Set $t=4n$
- Hard distribution ν: $(x,y) \in [t]^n \times [t]^n$, uniform random
- $3/4 \leqslant \Pr_{(x,y)\sim\nu}[EE(x,y)=0] = (1-1/(4n))^n \leqslant e^{-1/4} < 0.78$
- $\Rightarrow$ Any 0-round protocol has 0.22 error

Round elimination

Thm: No r-round $C = O(n \log^{(r)} n)$-bits protocol for EE_n^t

Induction on r:

Round elimination

Thm: No r-round $C = O(n \log^{(r)} n)$-bits protocol for EE_n^t

Induction on r:

r-round $C = \varepsilon n \log^{(r)} n$-bits protocol for EE_n^t

Round elimination

Thm: No r-round $C = O(n \log^{(r)} n)$-bits protocol for EE_n^t

Induction on r:

r-round $C=\varepsilon n \log^{(r)} n$-bits protocol for EE_n^t

construct $\Rightarrow$

Round elimination

Thm: No r-round $C = O(n \log^{(r)} n)$-bits protocol for EE_n^t

Induction on r:

r-round $C=\varepsilon n \log^{(r)} n$-bits protocol for EE_n^t

construct $\Rightarrow$

(r-1)-round 10C-bits protocol for $EE_{n'}^{t'}$ where $n'=n/B$ and $t'=t/B$

$\boxed{\mathbf{B}=2^{C/n}}$

Round elimination

Thm: No r-round $C = O(n \log^{(r)} n)$-bits protocol for EE_n^t

Induction on r:

r-round $C=\varepsilon n \log^{(r)} n$-bits protocol for EE_n^t

construct
$\Rightarrow$

(r-1)-round 10C-bits protocol for $EE_{n'}^{t'}$ where $n'=n/B$ and $t'=t/B$

$\boxed{B=2^{C/n}}$

Observe $10C=o(n' \log^{(r-1)} n')$

Contradicts induction hypothesis!

Round elimination

Thm: No r-round $C = O(n \log^{(r)} n)$-bits protocol for EE_n^t

Induction on r:

r-round $C=\varepsilon n \log^{(r)} n$-bits protocol for EE_n^t

construct $\Rightarrow$

(r-1)-round 10C-bits protocol for $EE_{n'}^{t'}$ where $n'=n/B$ and $t'=t/B$

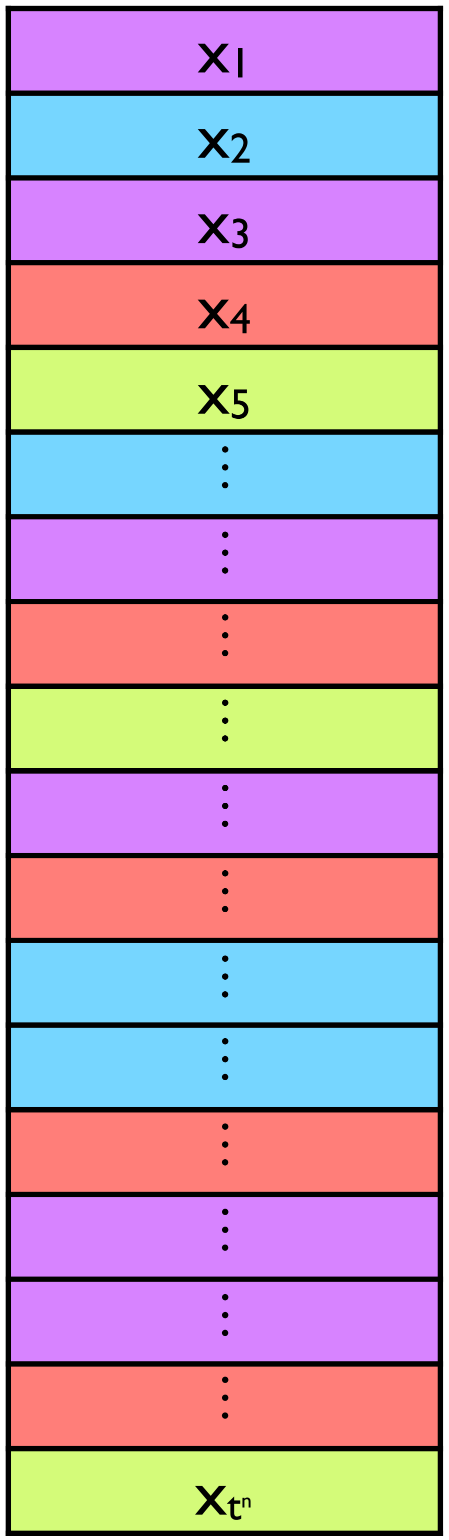
$B=2^{C/n}$

Observe $10C=o(n' \log^{(r-1)} n')$

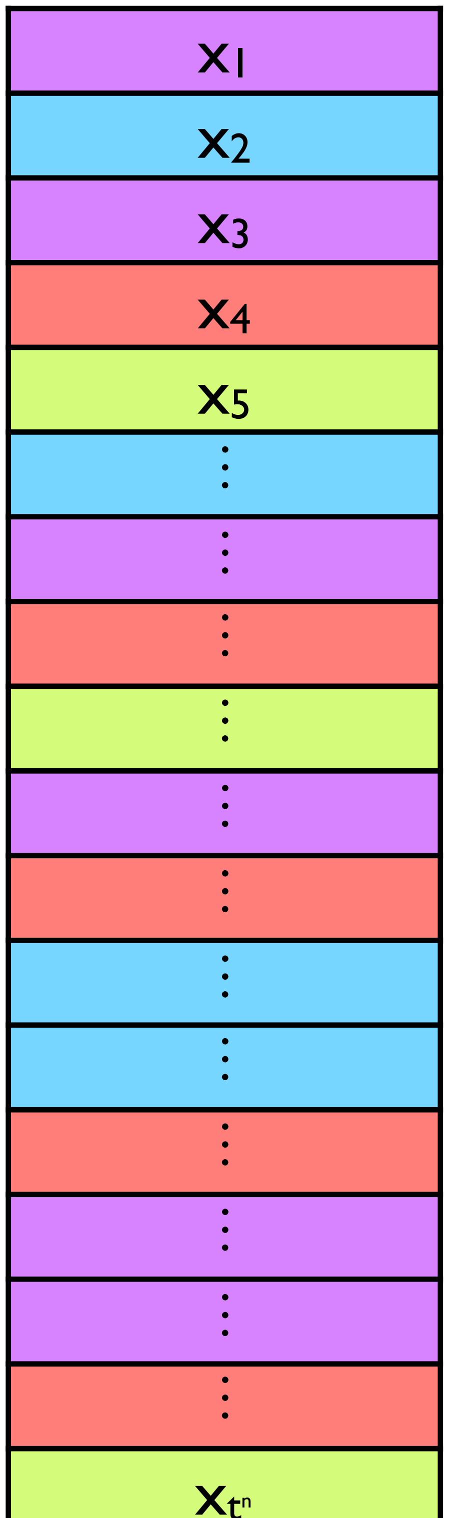
Contradicts induction hypothesis!

Note: $t'=4n'$

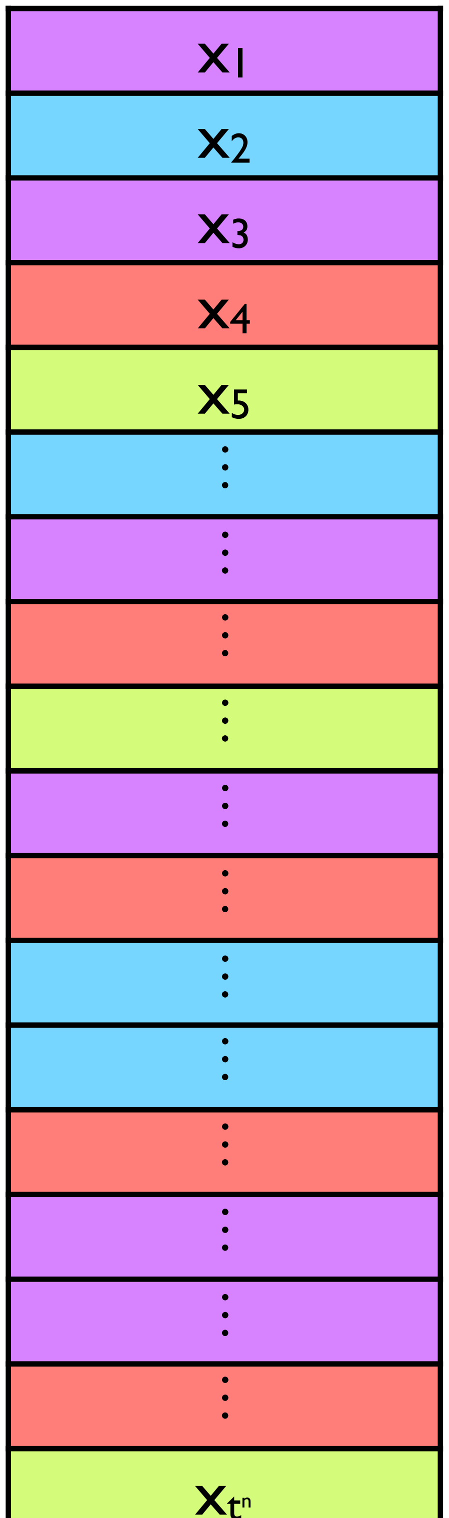
Fixing the first message



Fixing the first message

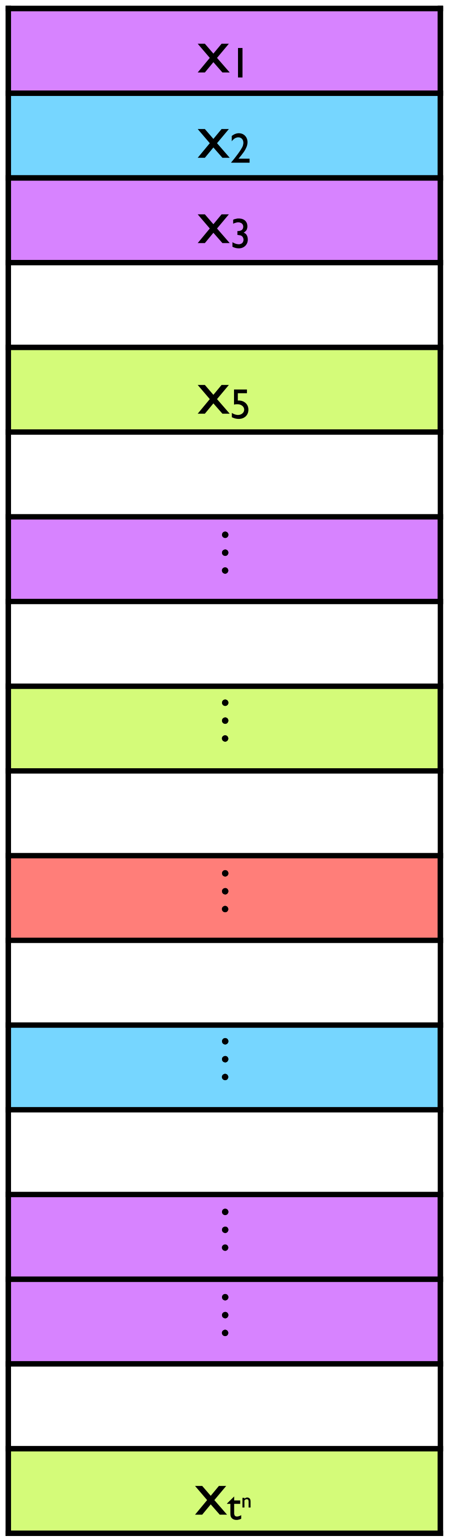


Fixing the first message



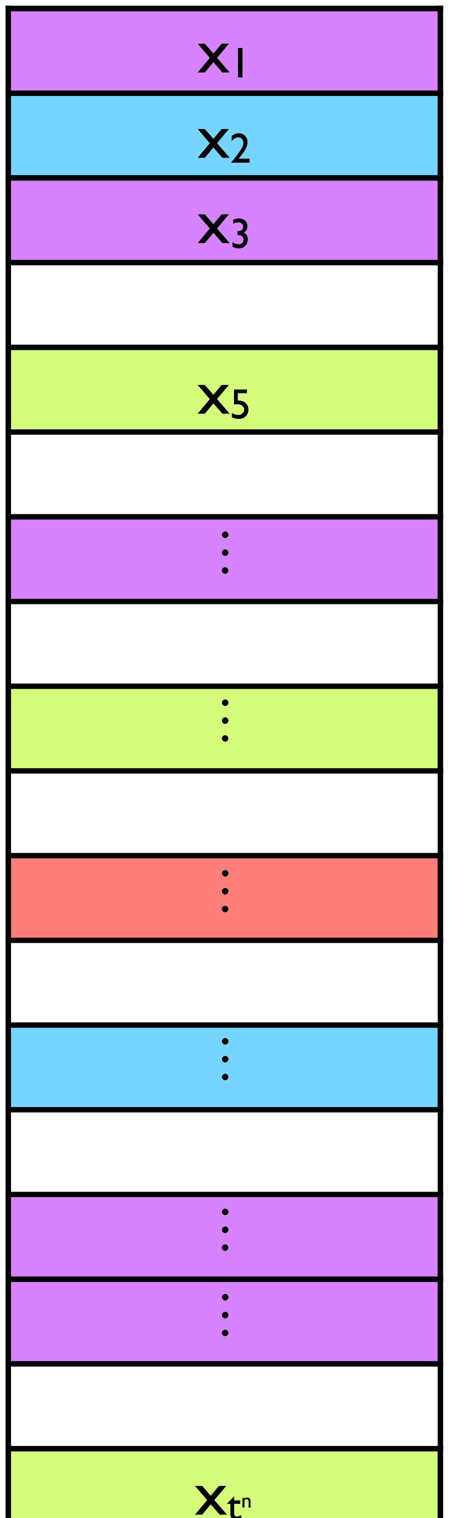
- Throw x_0, $Pr_y[P(x_0,y) \neq EE(x_0,y)] \geq 2\delta$

Fixing the first message



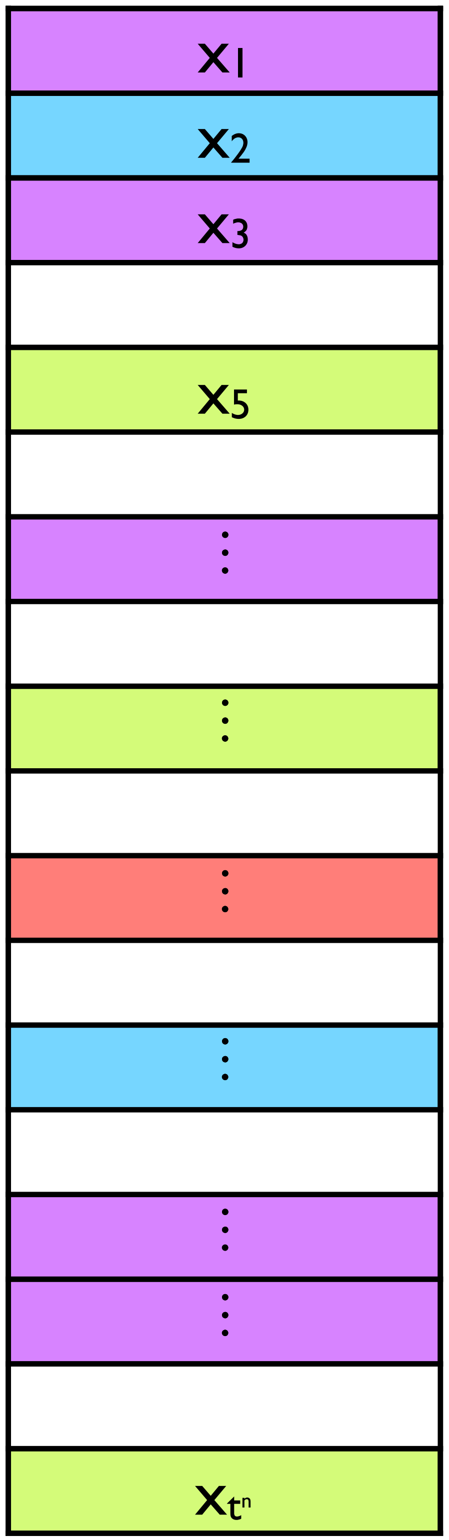
- Throw x_0, $\Pr_y[P(x_0,y) \neq EE(x_0,y)] \geq 2\delta$

Fixing the first message



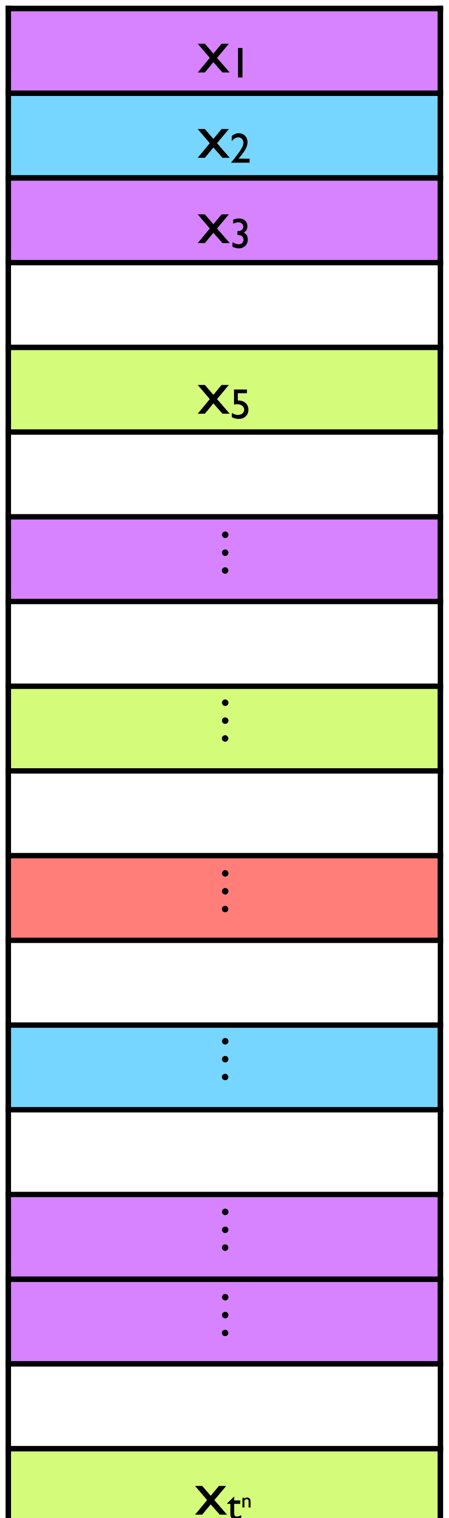
- Throw x_0, $Pr_y[P(x_0,y) \neq EE(x_0,y)] \geq 2\delta$
- At most half the inputs are gone

Fixing the first message



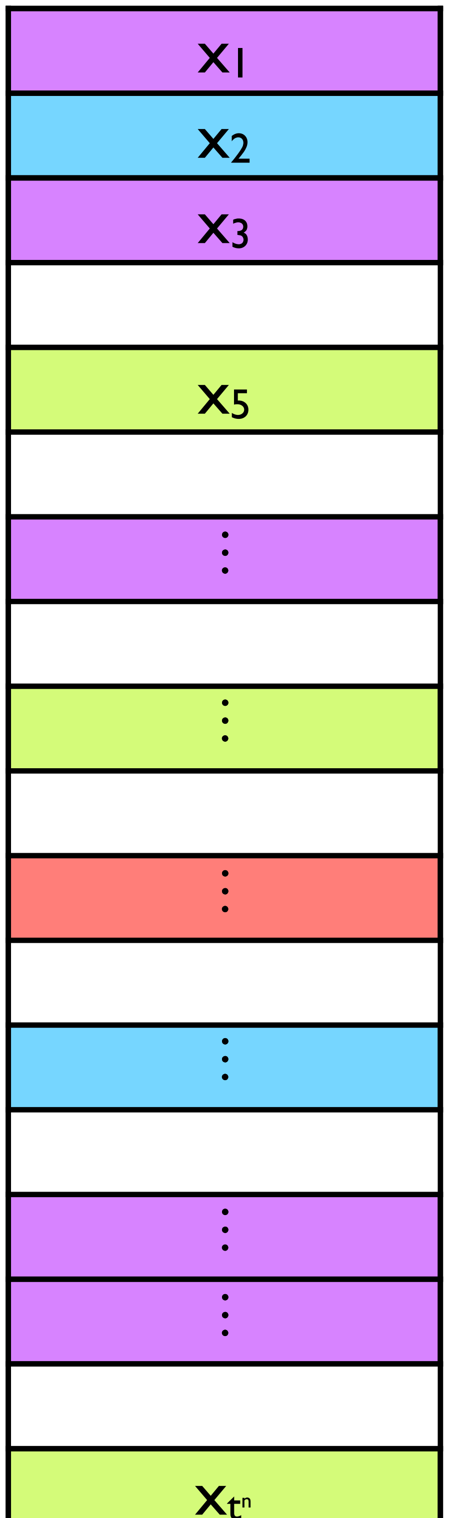
- Throw x_0, $\Pr_y[P(x_0,y) \neq EE(x_0,y)] \geq 2\delta$
- At most half the inputs are gone
- Fix the most frequent message **m***

Fixing the first message



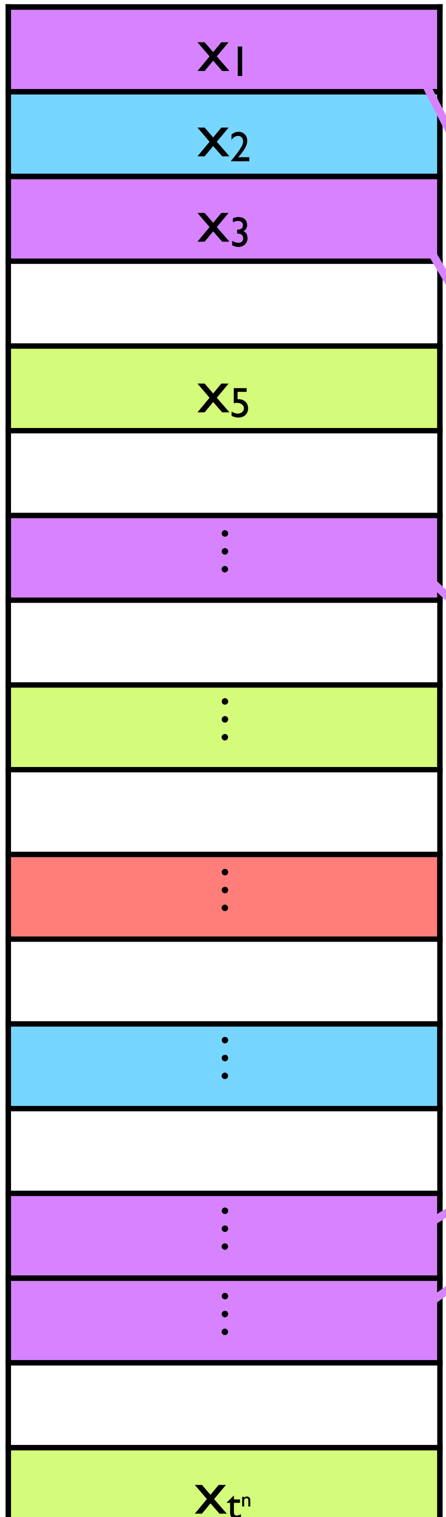
- Throw x_0, $\Pr_y[P(x_0,y) \neq EE(x_0,y)] \geq 2\delta$
- At most half the inputs are gone
- Fix the most frequent message **m***
- Let $S \subseteq [t]^n$ be inputs on which **m*** is sent

Fixing the first message



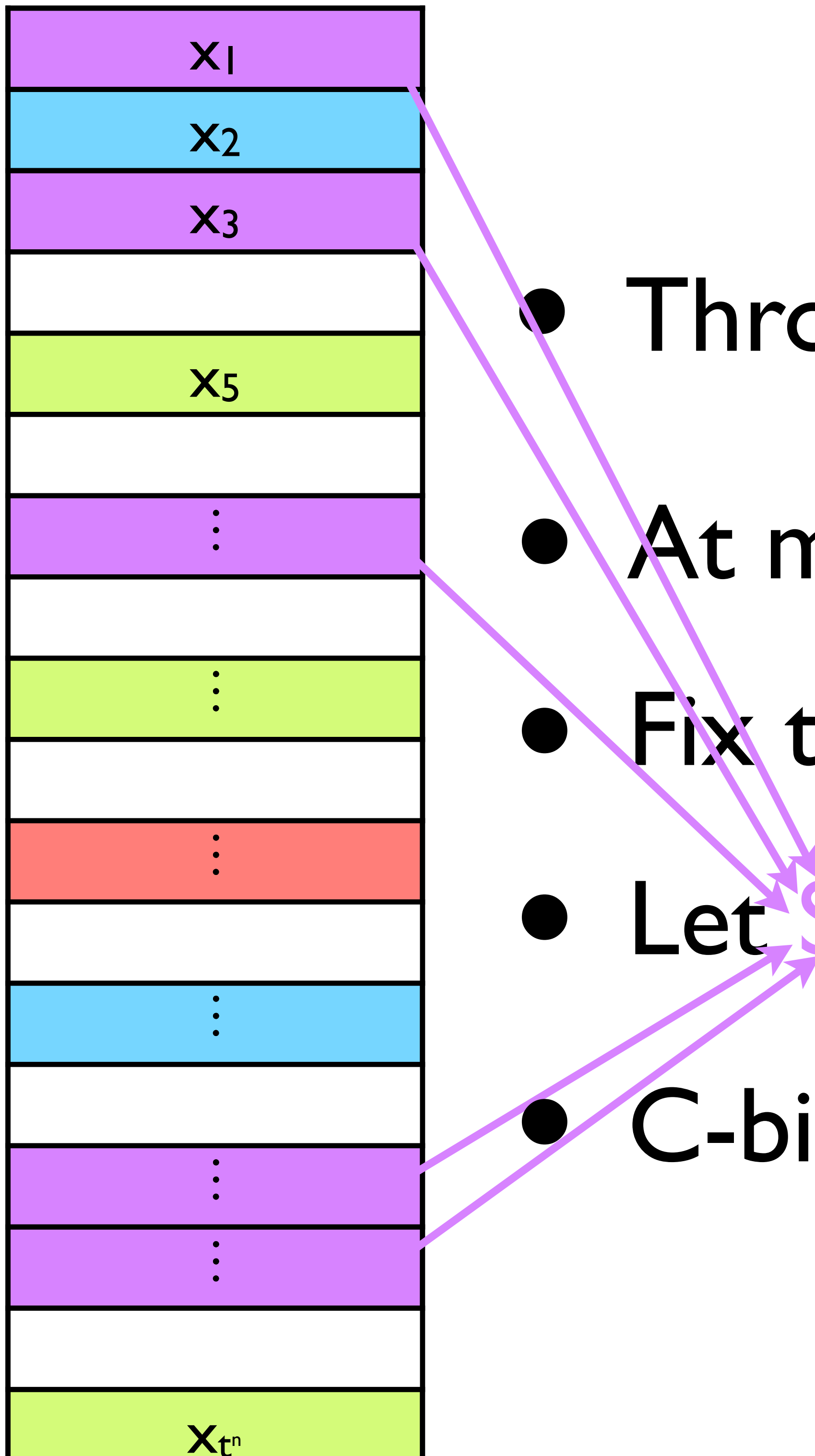
- Throw x_0, $Pr_y[P(x_0,y) \neq EE(x_0,y)] \geq 2\delta$
- At most half the inputs are gone
- Fix the most frequent message **m***
- Let $S \subseteq [t]^n$ be inputs on which **m*** is sent

Fixing the first message



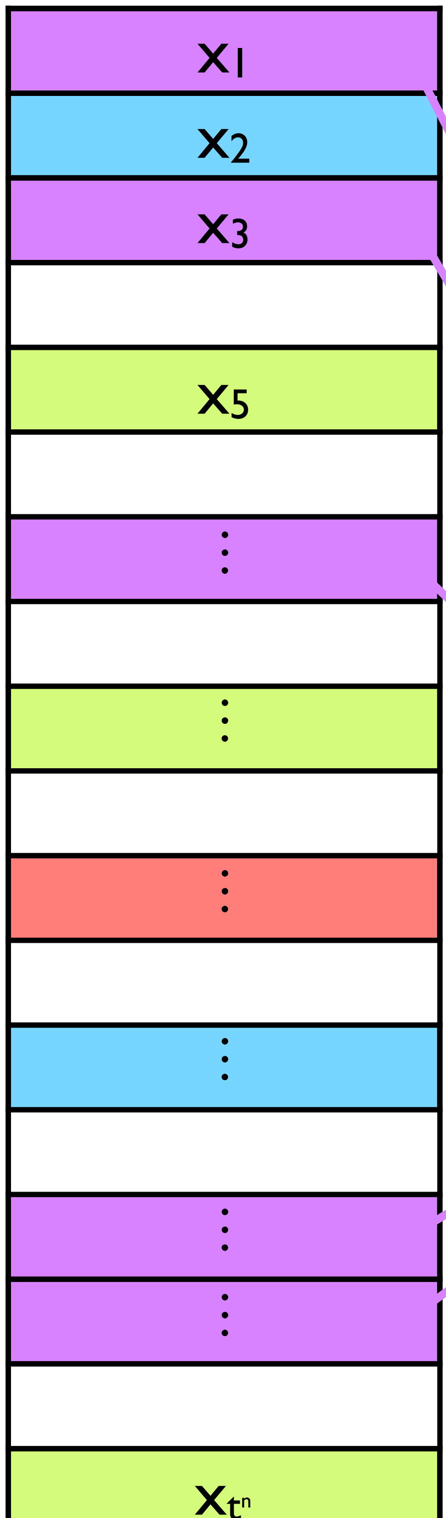
- Throw x_0, $\Pr_y[P(x_0,y) \neq EE(x_0,y)] \geq 2\delta$
- At most half the inputs are gone
- Fix the most frequent message **m***
- Let $S \subseteq [t]^n$ be inputs on which **m*** is sent

Fixing the first message



- Throw x_0, $\Pr_y[P(x_0,y) \neq EE(x_0,y)] \geq 2\delta$
- At most half the inputs are gone
- Fix the most frequent message **m***
- Let $S \subseteq [t]^n$ be inputs on which **m*** is sent
- C-bits protocol $\Rightarrow$ $\leq 2^C$ different messages

Fixing the first message



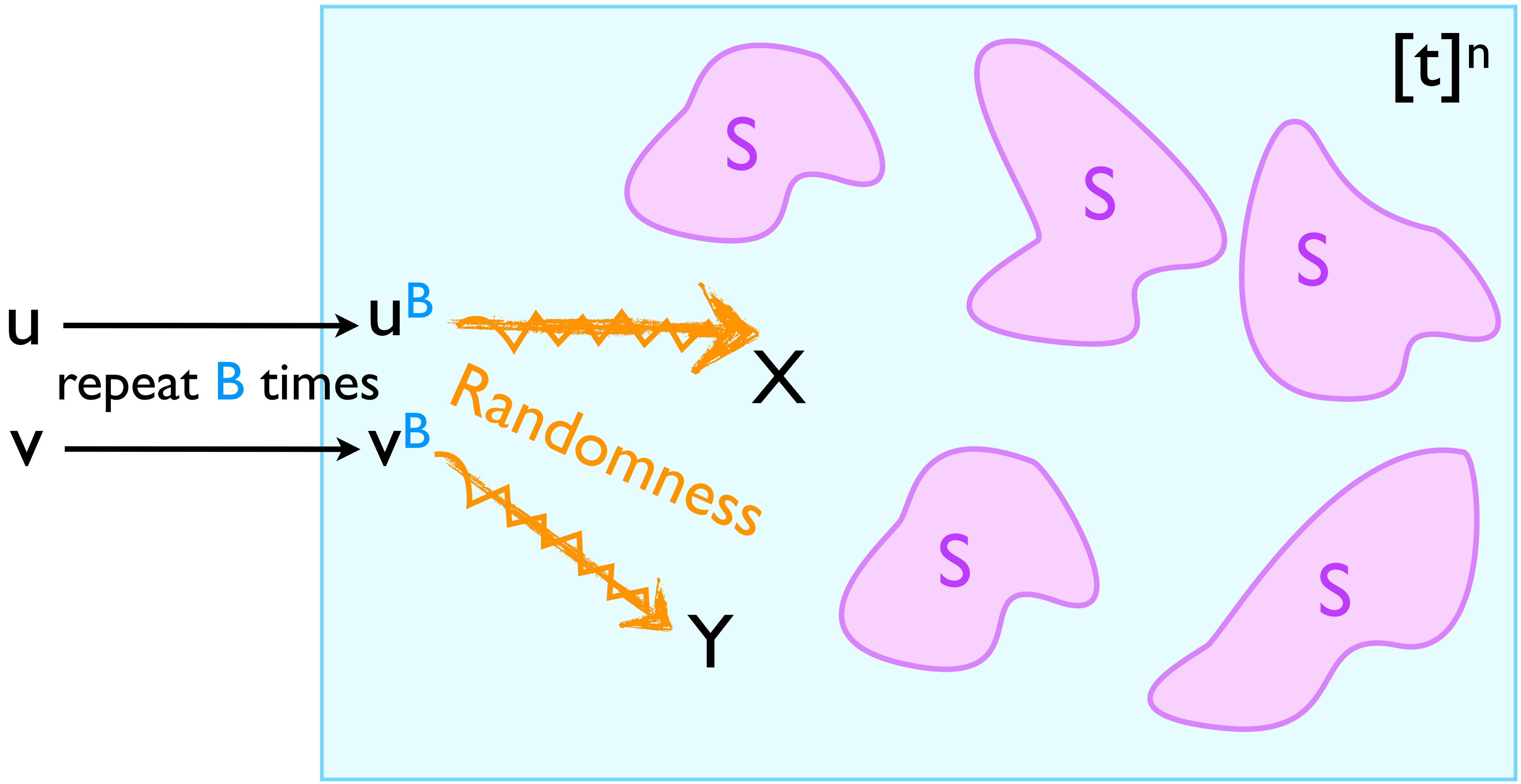
- Throw x_0, $\Pr_y[P(x_0,y) \neq EE(x_0,y)] \geq 2\delta$
- At most half the inputs are gone
- Fix the most frequent message **m***
- Let $S \subseteq [t]^n$ be inputs on which **m*** is sent
- C-bits protocol $\Rightarrow$ $\leq 2^C$ different messages
- $|S| \geq t^n/2^{C+1}$

r-round protocol for EE_n^t ⇒ (r-1)-round protocol for t'

$\mathbf{B} = 2^{C/n}$

$n' = n/B$

$u, v \in [t']^{n'}$

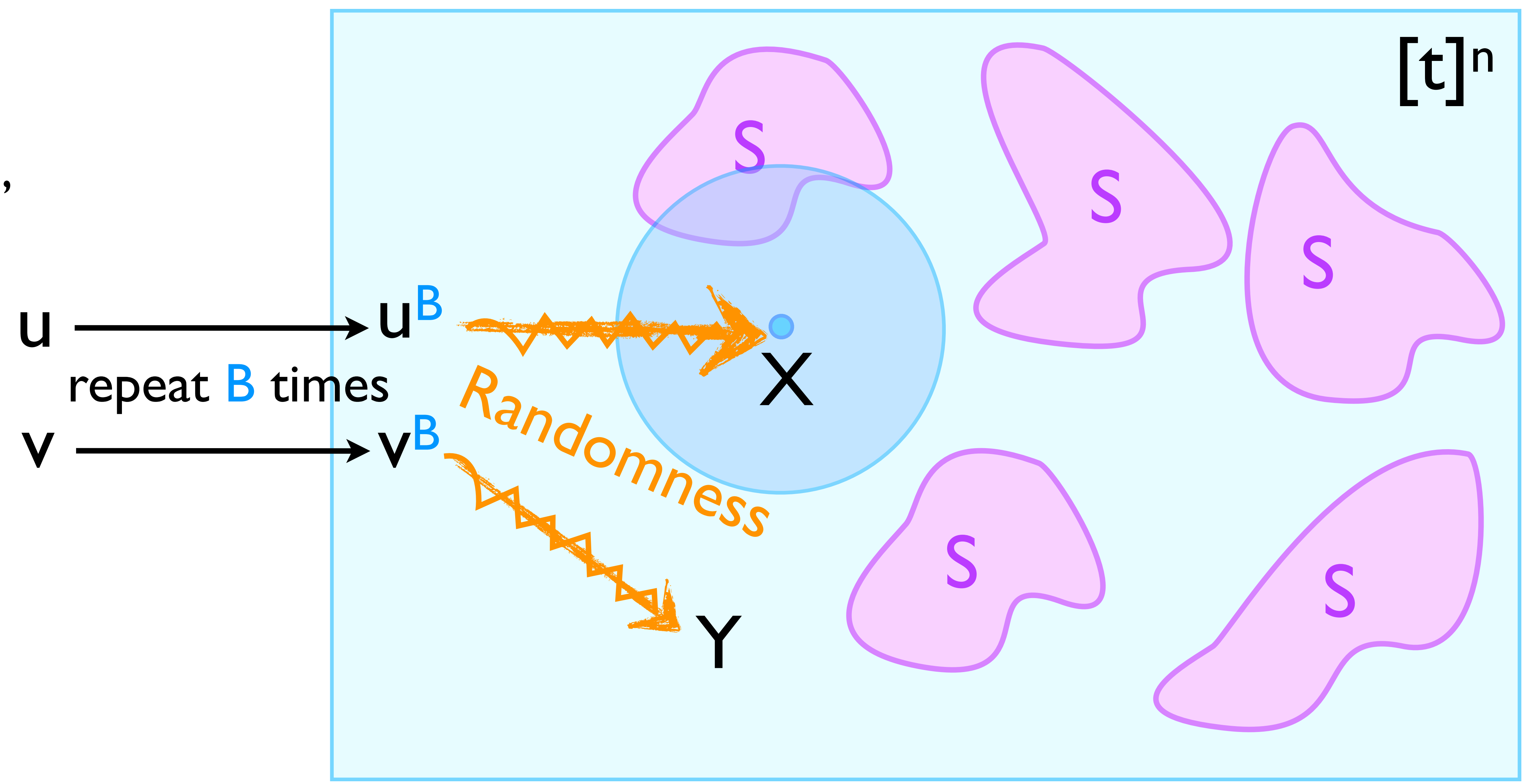


r-round protocol for $EE^t_n \Rightarrow$ (r-1)-round protocol for $^{t'}$

B $= 2^{C/n}$

$n' = n/B$

$u,v \in [t']^{n'}$

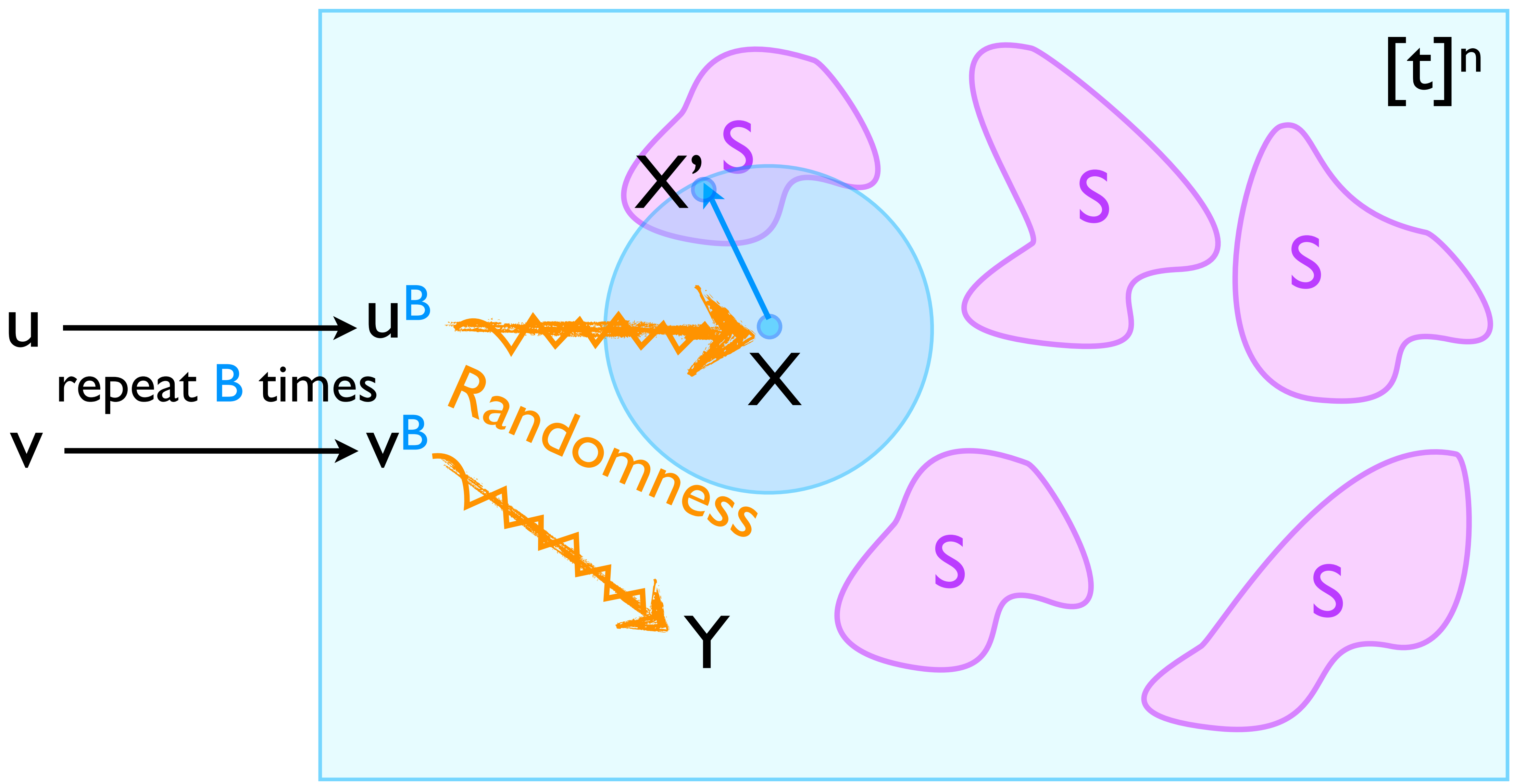


r-round protocol for EE_n^t $\Rightarrow$ (r-1)-round protocol for t'

$\mathbf{B} = 2^{C/n}$

$n' = n/B$

$u,v \in [t']^{n'}$

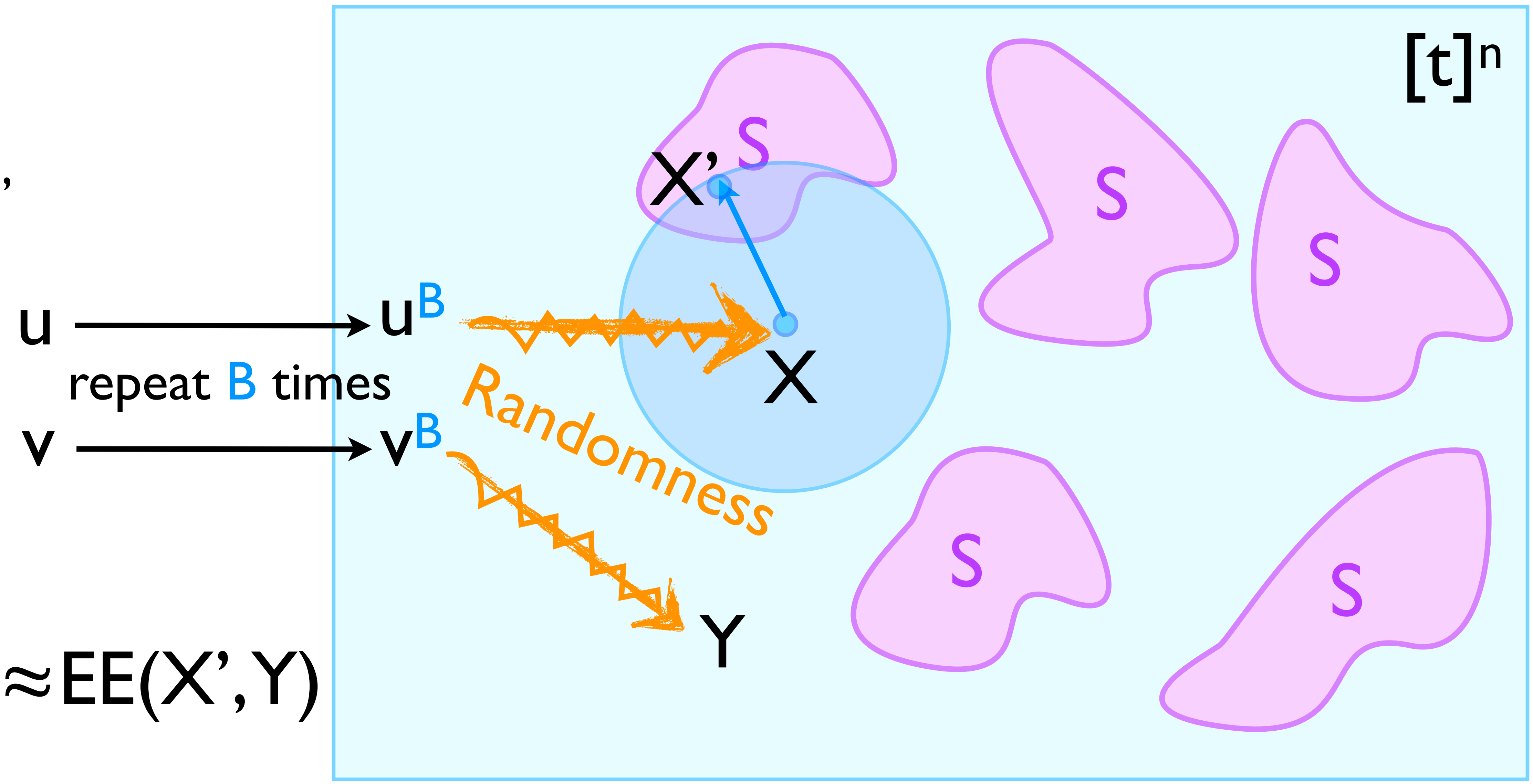


r-round protocol for EE^t_n ⇒ (r-1)-round protocol for t'

$\mathbf{B} = 2^{C/n}$

$n' = n/B$

$u,v \in [t']^{n'}$



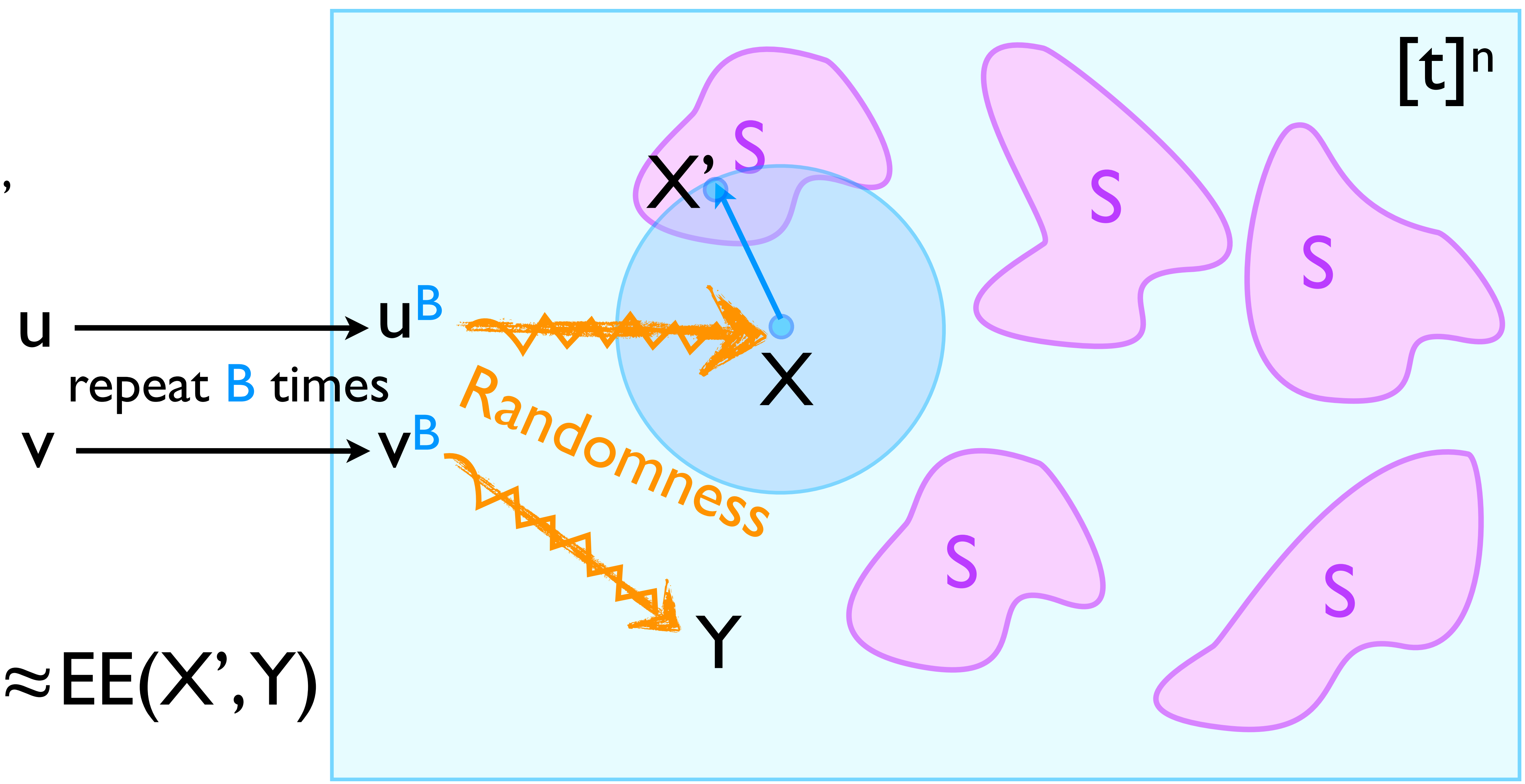
① $EE(u,v) \approx EE(X',Y)$

r-round protocol for EE_n^t ⇒ (r-1)-round protocol for $^{t'}$

$\mathbf{B} = 2^{C/n}$

$n' = n/B$

$u,v \in [t']^{n'}$



① $EE(u,v) \approx EE(X',Y)$

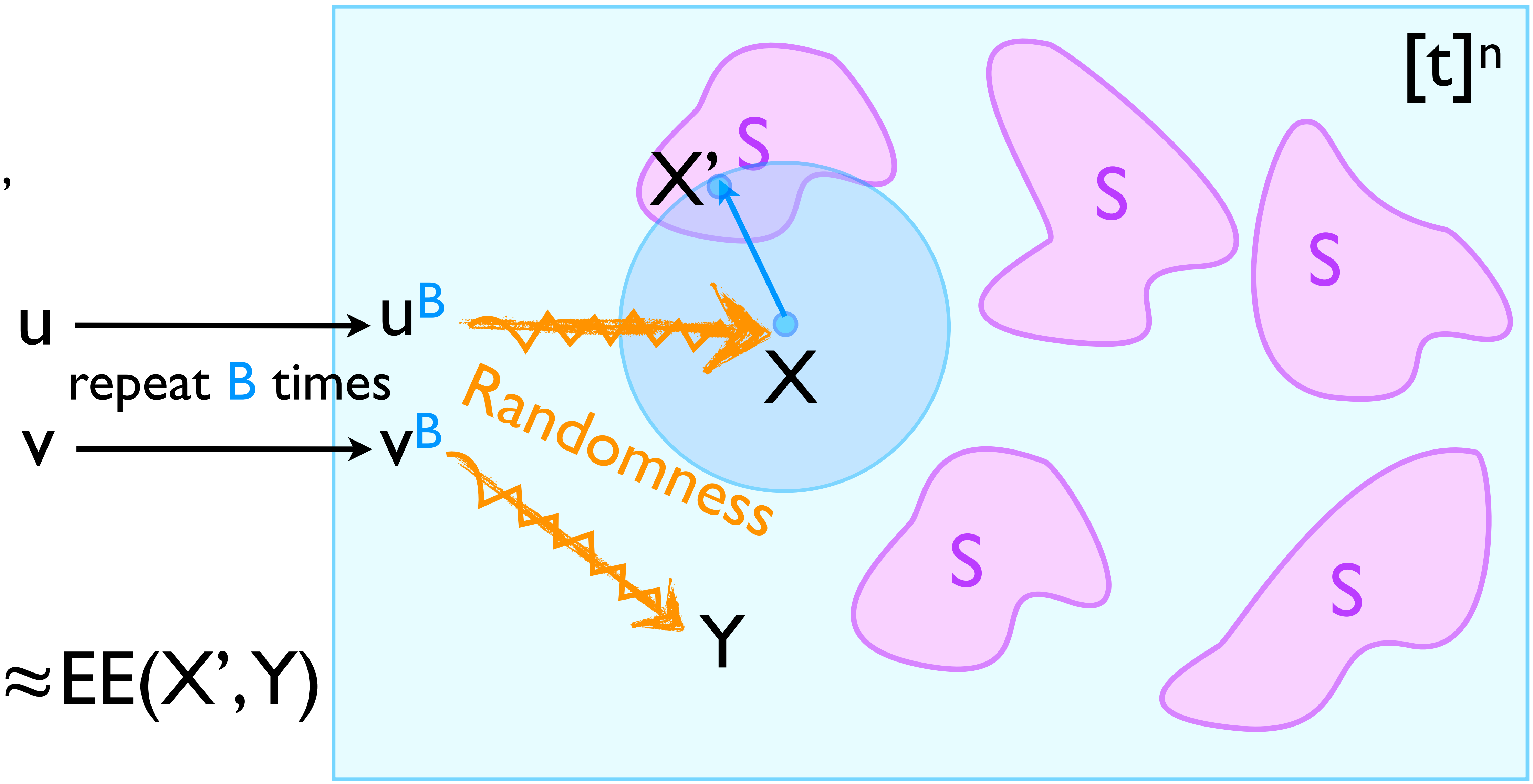
② $Y \mid X'$ is $\approx$uniform

r-round protocol for EE^t_n ⇒ (r-1)-round protocol for t'

$B = 2^{C/n}$

$n' = n/B$

$u,v \in [t']^{n'}$



① $EE(u,v) \approx EE(X',Y)$

② $Y \mid X'$ is ≈uniform

③ $X' \in S$ ⇒ first message fixed

Protocol for $EE^{t'}_{n'}$:

Protocol for $EE_{n'}^{t'}$:

Given $u,v \in [t']^{n'}$

Protocol for $EE_{n'}^{t'}$:

Given $u,v \in [t']^{n'}$

u:	2	3	1	4	2	1
v:	4	2	3	4	1	3

Protocol for $EE_{n'}^{t'}$:

Recall $n' = n/B$

$B = 2^{C/n}$

Given $u,v \in [t']^{n'}$

u:	2	3	1	4	2	1
v:	4	2	3	4	1	3

Protocol for $EE^{t'}_{n'}$:

Recall $n' = n/B$

$\boxed{B = 2^{C/n}}$

Given $u, v \in [t']^{n'}$

Repeat each coordinate B times

u:	2	3	1	4	2	1
v:	4	2	3	4	1	3

Protocol for $EE_{n'}^{t'}$:

Recall $n' = n/B$

$\boxed{B=2^{C/n}}$

Given $u,v \in [t']^{n'}$

Repeat each coordinate B times

u:	2	3	1	4	2	1
v:	4	2	3	4	1	3

u^B:	2	2	2	2	3	3	3	3	1	1	1	1	4	4	4	4	2	2	2	2	1	1	1	1
v^B:	4	4	4	4	2	2	2	2	3	3	3	3	4	4	4	4	1	1	1	1	3	3	3	3

Protocol for $EE^{t'}_{n'}$:

Recall $n' = n/B$

$\boxed{B=2^{C/n}}$

Given $u,v \in [t']^{n'}$

Repeat each coordinate B times

u: 2 3 1 4 2 1

v: 4 2 3 4 1 3

u^B: 2 2 2 2 3 3 3 3 1 1 1 1 4 4 4 4 2 2 2 2 1 1 1 1

v^B: 4 4 4 4 2 2 2 2 3 3 3 3 4 4 4 4 1 1 1 1 3 3 3 3

Pick a random function f_i: $[t'] \mapsto [t]$ for each $i \in [n]$

Protocol for $EE_{n'}^{t'}$:

Recall $n' = n/B$

$B = 2^{C/n}$

Given $u, v \in [t']^{n'}$

Repeat each coordinate B times

u:	2	3	1	4	2	1
v:	4	2	3	4	1	3

u^B:	2	2	2	2	3	3	3	3	1	1	1	1	4	4	4	4	2	2	2	2	1	1	1	1
v^B:	4	4	4	4	2	2	2	2	3	3	3	3	4	4	4	4	1	1	1	1	3	3	3	3

4	2	7	9	2	7	3	6	2	5	2	6	9	3	1	7	6	7	8	4	2	6	9	3
6	3	1	3	8	7	1	5	4	7	8	9	9	3	1	7	2	1	3	6	9	2	2	1

Pick a random function $f_i: [t'] \mapsto [t]$ for each $i \in [n]$

• : Phantom match

Protocol for $EE^{t'}_{n'}$:

Recall $n' = n/B$

$\boxed{B = 2^{C/n}}$

Given $u, v \in [t']^{n'}$

Repeat each coordinate B times

u: 2 3 1 4 2 1

v: 4 2 3 4 1 3

u^B: 2 2 2 2 3 3 3 3 1 1 1 1 4 4 4 4 2 2 2 2 1 1 1 1

v^B: 4 4 4 4 2 2 2 2 3 3 3 3 4 4 4 4 1 1 1 1 3 3 3 3

Pick a random function f_i: $[t'] \mapsto [t]$ for each $i \in [n]$

4 2 7 9 2 7 3 6 2 5 2 6 9 3 1 7 6 7 8 4 2 6 9 3

6 3 1 3 8 7 1 5 4 7 8 9 9 3 1 7 2 1 3 6 9 2 2 1

• : Phantom match

Permute coordinates randomly

Protocol for $EE_{n'}^{t'}$:

Recall $n' = n/B$

$B = 2^{C/n}$

Given $u,v \in [t']^{n'}$

Repeat each coordinate B times

u:	2	3	1	4	2	1
v:	4	2	3	4	1	3

u^B:	2	2	2	2	3	3	3	3	1	1	1	1	4	4	4	4	2	2	2	2	1	1	1	1
v^B:	4	4	4	4	2	2	2	2	3	3	3	3	4	4	4	4	1	1	1	1	3	3	3	3

Pick a random function f_i: $[t'] \mapsto [t]$ for each $i \in [n]$

4	2	7	9	2	7	3	6	2	5	2	6	9	3	1	7	6	7	8	4	2	6	9	3
6	3	1	3	8	7	1	5	4	7	8	9	9	3	1	7	2	1	3	6	9	2	2	1

• : Phantom match

Permute coordinates randomly

3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

Protocol for $EE_{n'}^{t'}$:

Recall $n' = n/B$

$\boxed{B=2^{C/n}}$

Given $u,v \in [t']^{n'}$

Repeat each coordinate B times

u:	2	3	1	4	2	1
v:	4	2	3	4	1	3

u^B:	2	2	2	2	3	3	3	3	1	1	1	1	4	4	4	4	2	2	2	2	1	1	1	1
v^B:	4	4	4	4	2	2	2	2	3	3	3	3	4	4	4	4	1	1	1	1	3	3	3	3

Pick a random function f_i: $[t'] \mapsto [t]$ for each $i \in [n]$

4	2	7	9	2	7	3	6	2	5	2	6	9	3	1	7	6	7	8	4	2	6	9	3
6	3	1	3	8	7	1	5	4	7	8	9	9	3	1	7	2	1	3	6	9	2	2	1

• : Phantom match

Permute coordinates randomly

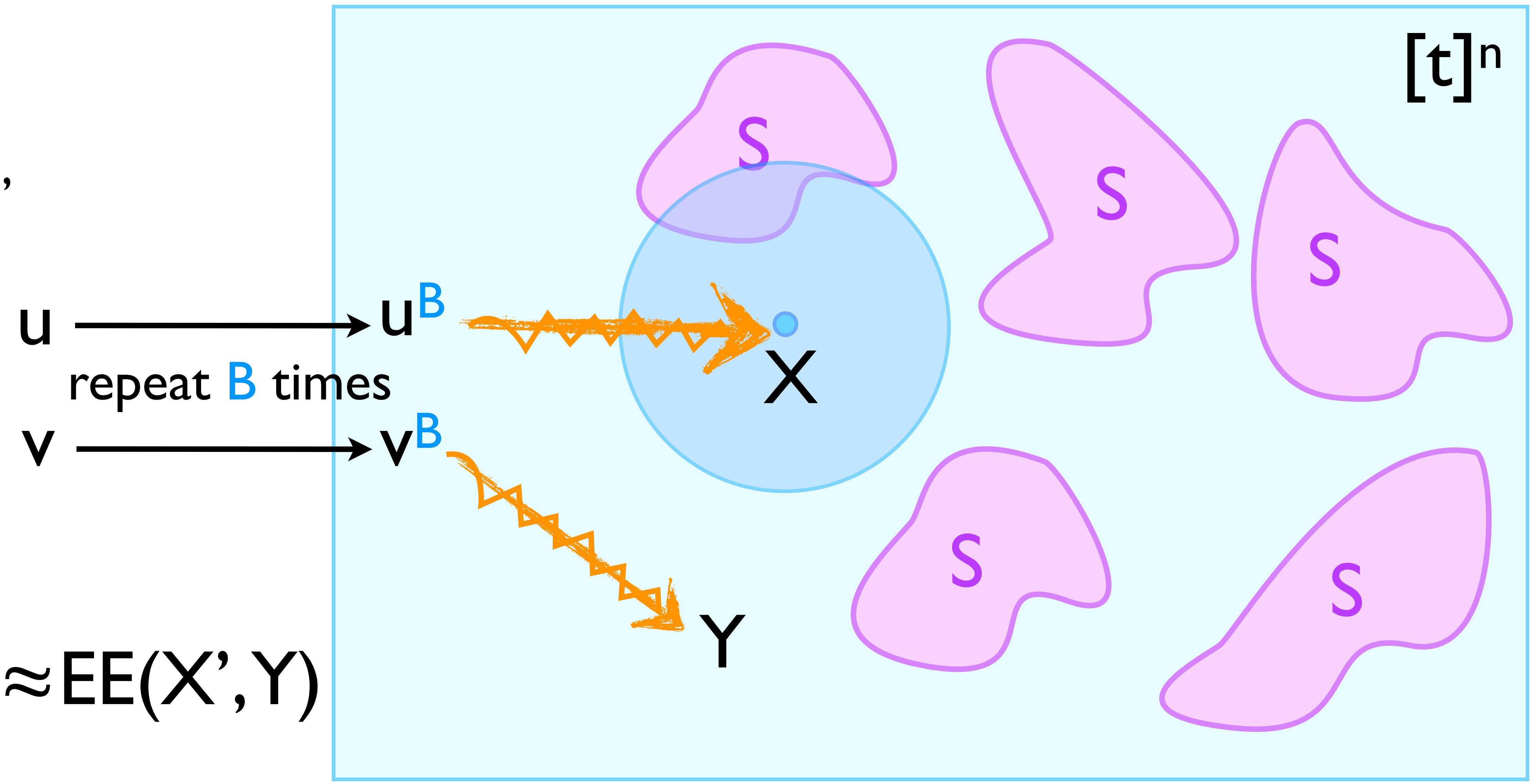
X:	3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

Step 2: Rounding to S

$\mathbf{B} = 2^{C/n}$

$n' = n/B$

$u,v \in [t']^{n'}$



① $EE(u,v) \approx EE(X',Y)$

② $Y \mid X'$ is ≈uniform

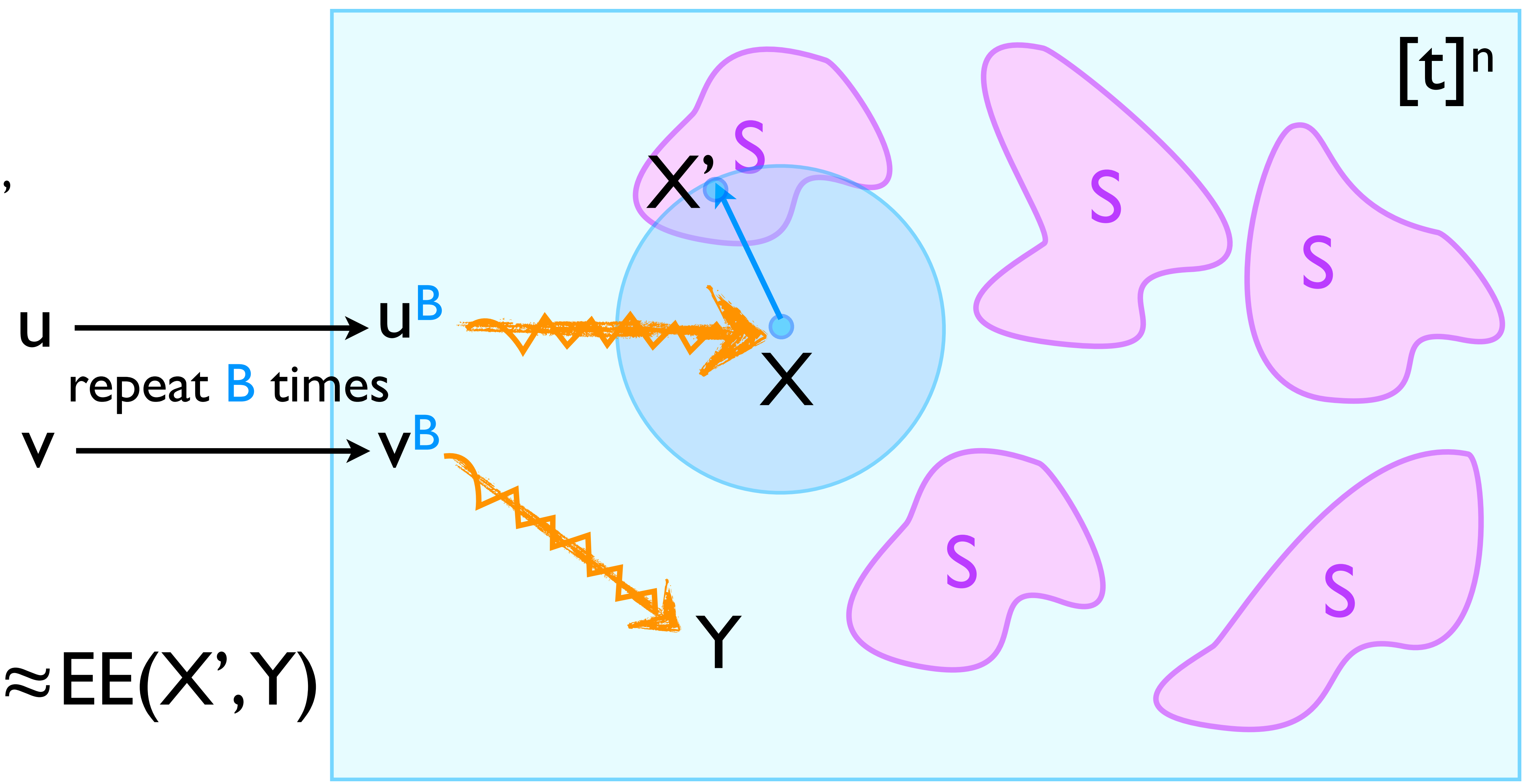
③ $X' \in S \Rightarrow$ first message fixed

Step 2: Rounding to S

$\mathbf{B} = 2^{C/n}$

$n' = n/B$

$u,v \in [t']^{n'}$

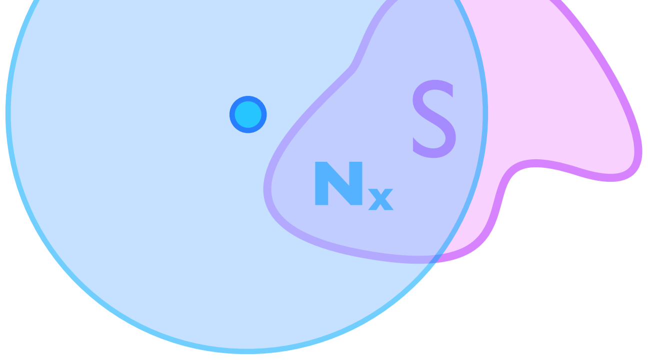


① $EE(u,v) \approx EE(X',Y)$

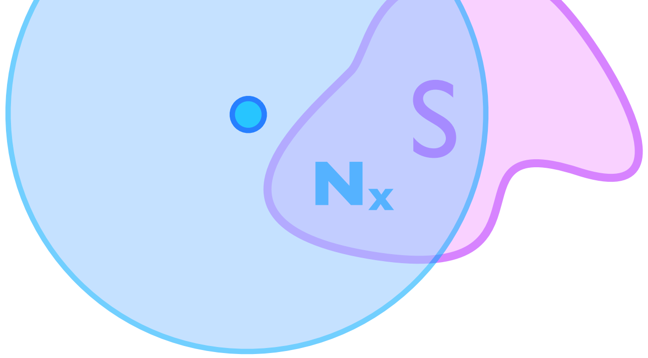
② $Y \mid X'$ is $\approx$uniform

③ $X' \in S \Rightarrow$ first message fixed

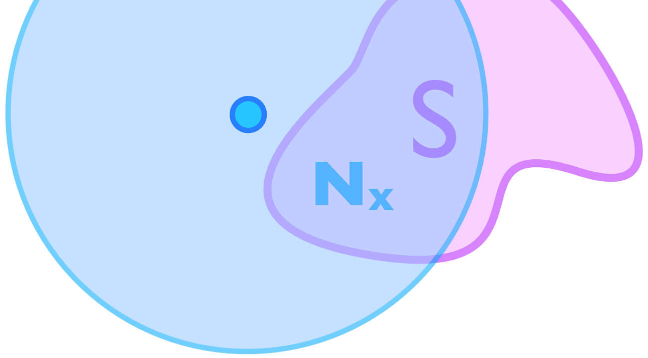
- $N_x = S \cap \mathrm{Ball}(X, n(1-1/B))$
- X': uniform in N_x



- $N_x = S \cap \mathrm{Ball}(X, n(1-1/B))$

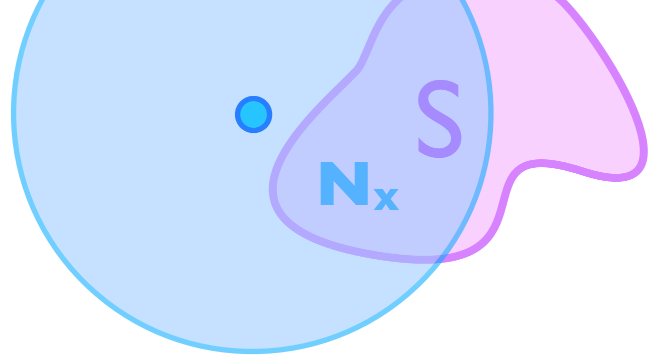


- X': uniform in N_X ③ $X' \in S \Rightarrow$ first message fixed



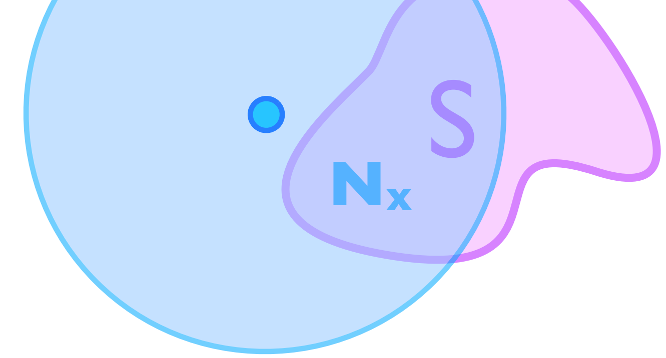
- $N_x = S \cap \mathrm{Ball}(X, n(1-1/B))$
- X': uniform in N_X ③ X'$\in$S $\Rightarrow$ first message fixed

- $N_x = S \cap \mathrm{Ball}(X, n(1-1/B))$
- X': uniform in N_x ③ ✓ $X' \in S \Rightarrow$ first message fixed

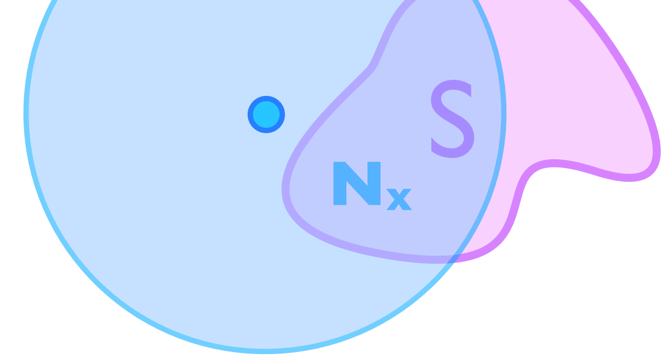


X:	3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

- $N_x = S \cap \mathrm{Ball}(X, n(1-1/B))$
- X': uniform in N_x 3✓ $X' \in S \Rightarrow$ first message fixed



X:	3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3



- $N_x = S \cap \mathrm{Ball}(X, n(1-1/B))$
- X': uniform in N_x ③✓ X'∈S ⇒ first message fixed

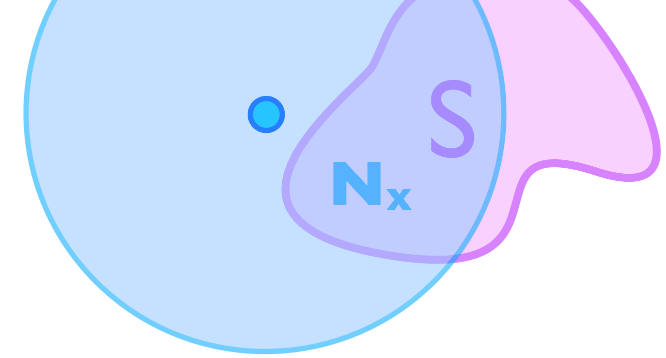
X: 3 2 8 1 5 9 2 3 3 7 6 3 9 7 6 9 7 7 4 2 5 9 5 8

Y: 1 8 3 1 7 3 9 3 1 1 2 1 2 7 5 3 1 7 6 3 7 9 7 3

X': Correlated randomness from N_x 5 9 5 8

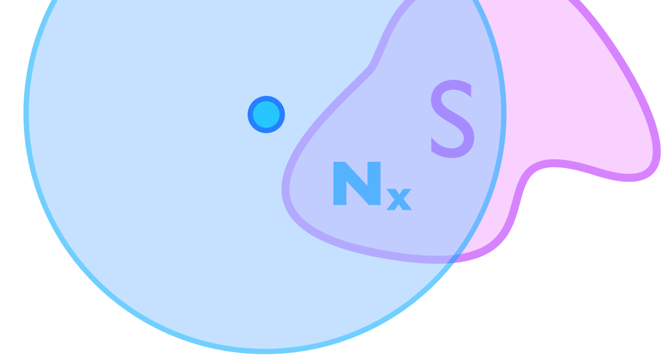
Y: 1 8 3 1 7 3 9 3 1 1 2 1 2 7 5 3 1 7 6 3 7 9 7 3

- $N_x = S \cap \mathrm{Ball}(X, n(1-1/B))$
- X': uniform in N_x ③✓ X'∈S ⇒ first message fixed



X:	3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3
X':	2	4	7	2	5	1	4	6	3	7	2	1	8	3	3	8	2	9	6	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

- $N_x = S \cap \mathrm{Ball}(X, n(1-1/B))$
- X’: uniform in N_x ③✓ X’∈S ⇒ first message fixed

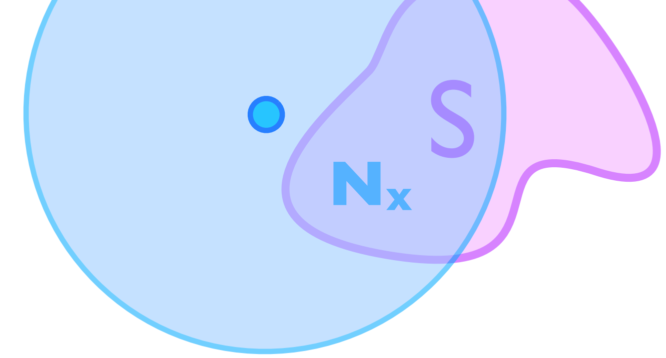


X:	3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

X’:	2	4	7	2	5	1	4	6	3	7	2	1	8	3	3	8	2	9	6	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

About one match survives, so EE(u,v)=1 ⇒ EE(X’,Y)=1

- $N_x = S \cap \text{Ball}(X, n(1-1/B))$



- X’: uniform in N_x ③ X’∈S ⇒ first message fixed

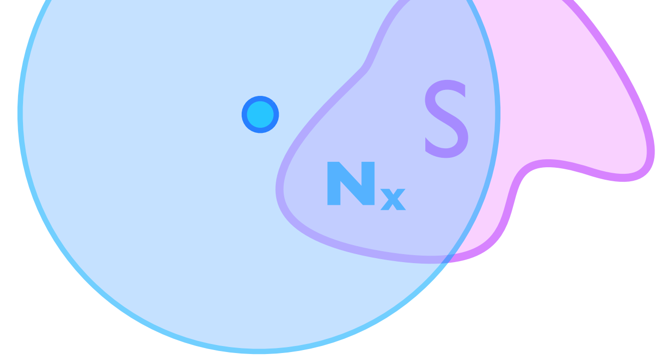
X:	3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

X’:	2	4	7	2	5	1	4	6	3	7	2		8	3	3	8	2	9	6	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

About one match survives, so EE(u,v)=1 ⇒ EE(X’,Y)=1

Since t=4n, Pr[Phantom] < 1/4, so EE(u,v)=0 ⇒ EE(X’,Y)=0

- N_x = S ∩ Ball(X, n(1-1/B))
- X': uniform in N_x ③ ✓ X'∈S ⇒ first message fixed



X:	3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

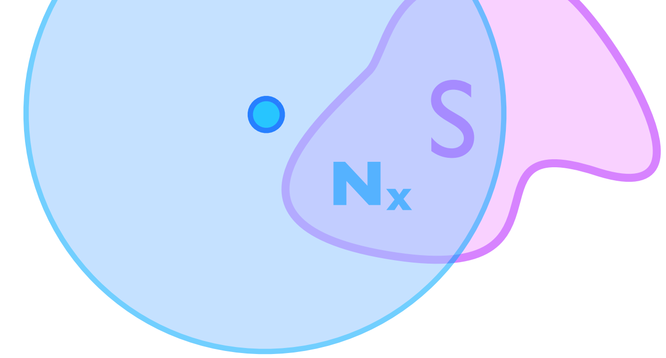
X':	2	4	7	2	5	1	4	6	3	7	2		8	3	3	8	2	9	6	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

About one match survives, so EE(u,v)=1 ⇒ EE(X',Y)=1

Since t=4n, Pr[Phantom] < 1/4, so EE(u,v)=0 ⇒ EE(X',Y)=0

① EE(u,v)≈EE(X',Y)

- $N_x = S \cap \mathrm{Ball}(X, n(1-1/B))$
- X': uniform in N_x ③ X'∈S ⇒ first message fixed



X:	3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

X':	2	4	7	2	5	1	4	6	3	7	2		8	3	3	8	2	9	6	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

About one match survives, so EE(u,v)=1 ⇒ EE(X',Y)=1

Since t=4n, Pr[Phantom] < 1/4, so EE(u,v)=0 ⇒ EE(X',Y)=0

① EE(u,v)≈EE(X',Y)

Show: 2 $Y \mid X'$ is $\approx$uniform

Show: 2 $Y \mid X'$ is $\approx$uniform

- This is needed as $\Pr[P(X',Y) \neq EE(X,Y)] \leq 2\delta$ only if $Y \mid X'$ is uniform

Show: 2 $Y \mid X'$ is $\approx$uniform

- This is needed as $\Pr[P(X',Y) \neq EE(X,Y)] \leq 2\delta$ only if $Y \mid X'$ is uniform

We show $H(Y \mid X') = n \log t - O(1)$

X:	3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

L: Set of $\boxed{\cdot}$, intentional matches of X, Y

Entropy loss comes from coordinates in L

$$\text{Entropy loss} = |L|\log t - H(X_L \mid L, X')$$

$$\leq |L|\log t - (|L|/n)H(X \mid X')$$

X:	3	2	8	1	5	9	2	3	3	7	6	3	9	7	6	9	7	7	4	2	5	9	5	8
Y:	1	8	3	1	7	3	9	3	1	1	2	1	2	7	5	3	1	7	6	3	7	9	7	3

L: Set of $\boxed{\cdot}$, intentional matches of X, Y

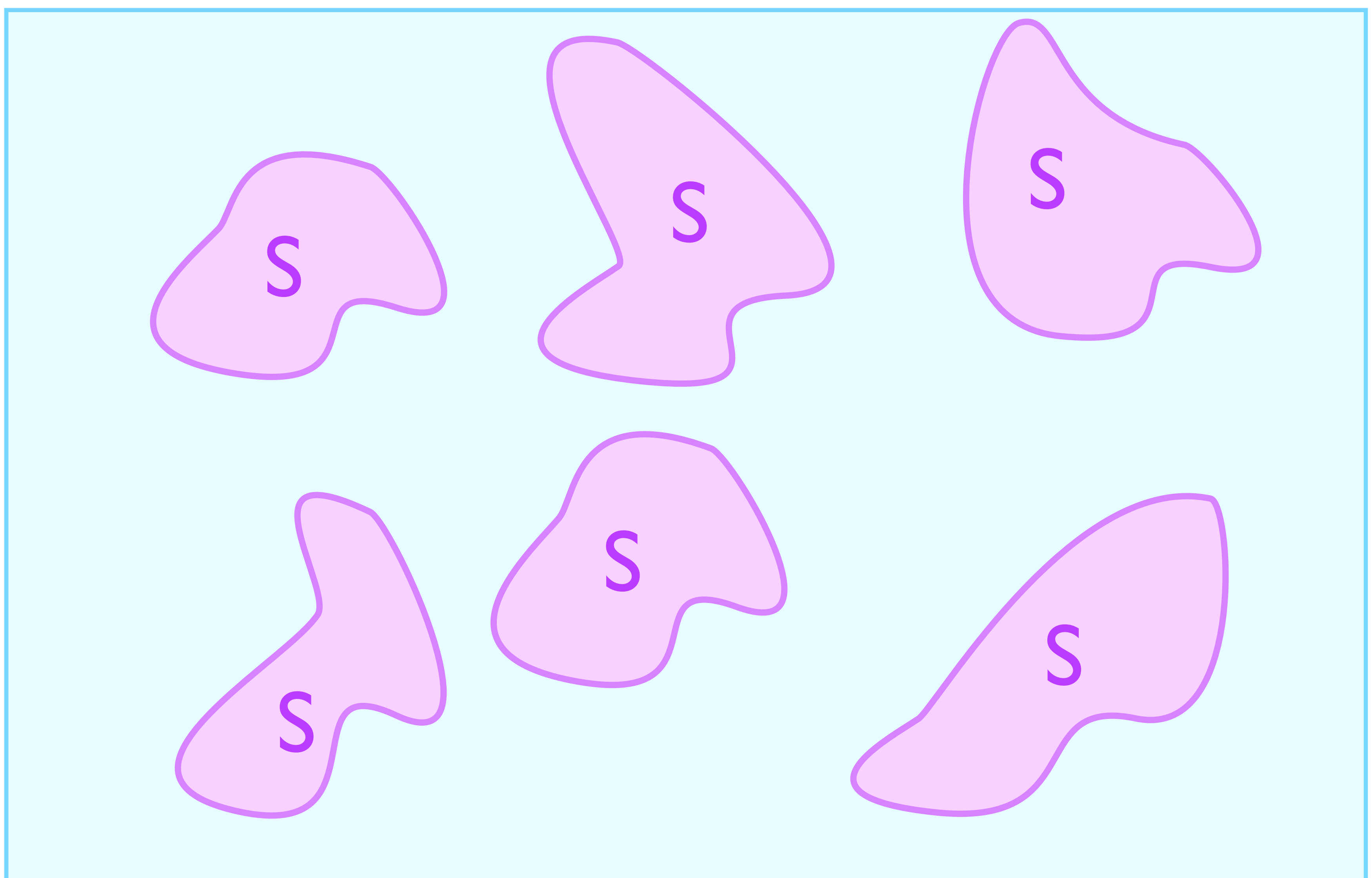
Entropy loss comes from coordinates in L

$$\text{Entropy loss} = |L|\log t - H(X_L \mid L, X')$$

$$\leq |L|\log t - (|L|/n)H(X \mid X')$$

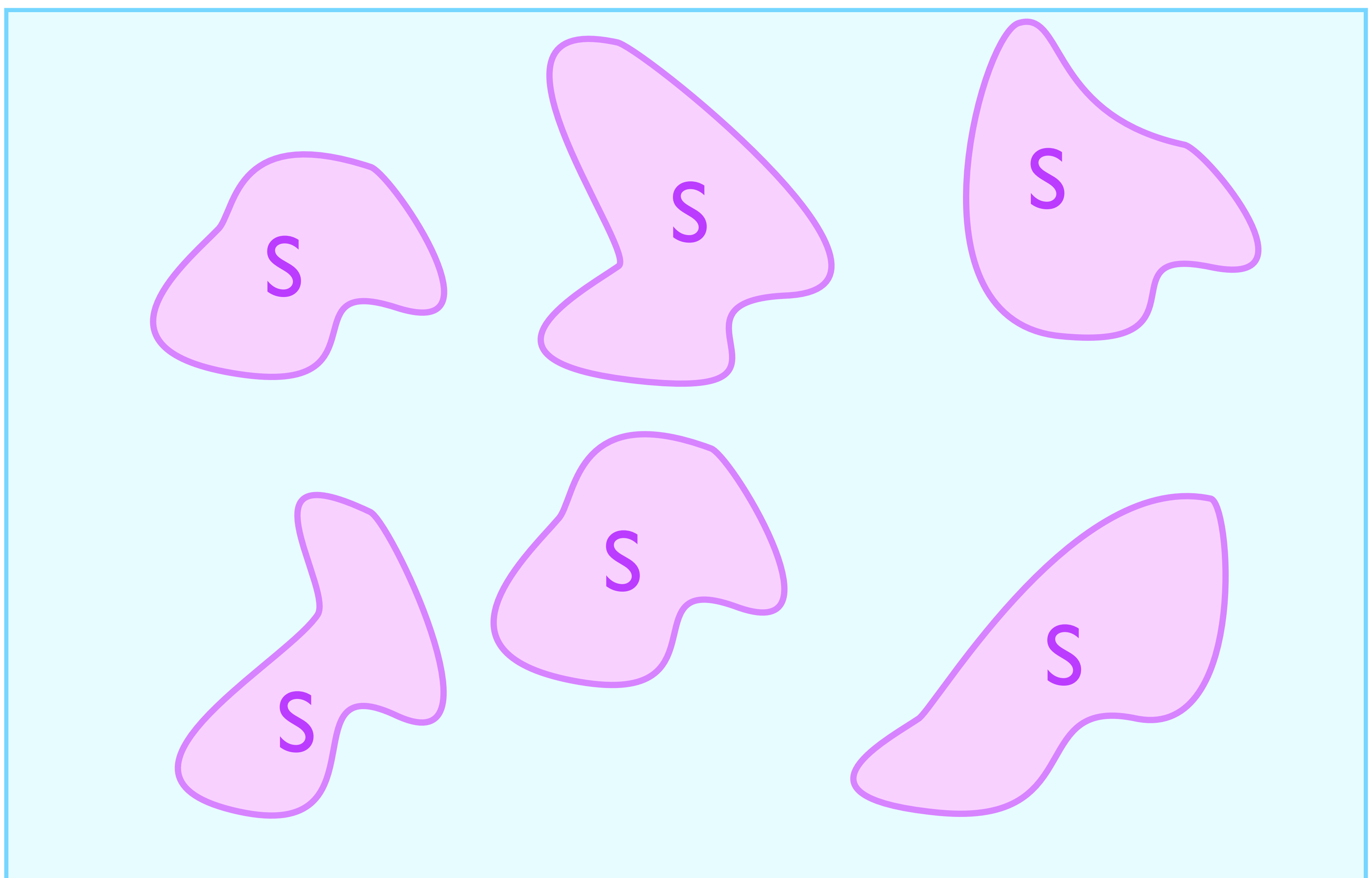
(Han-Shearer)

Want to lower bound **H(X' | X)**



Want to lower bound **H(X' | X)**

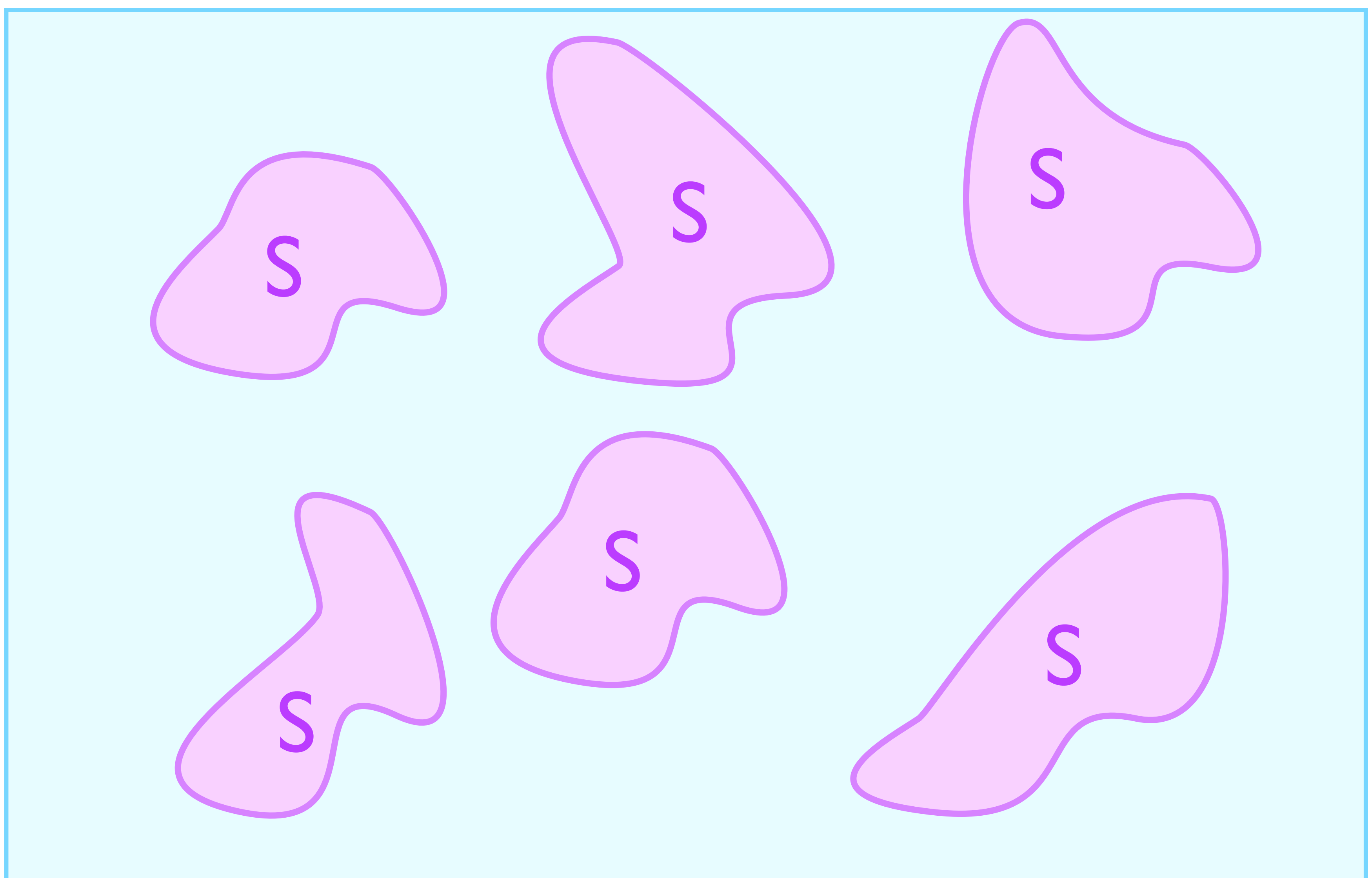
i.e., log |$\mathbf{N_x}$| for uniform random x



Want to lower bound **$H(X' \mid X)$**

i.e., log $|\mathbf{N_x}|$ for uniform random x

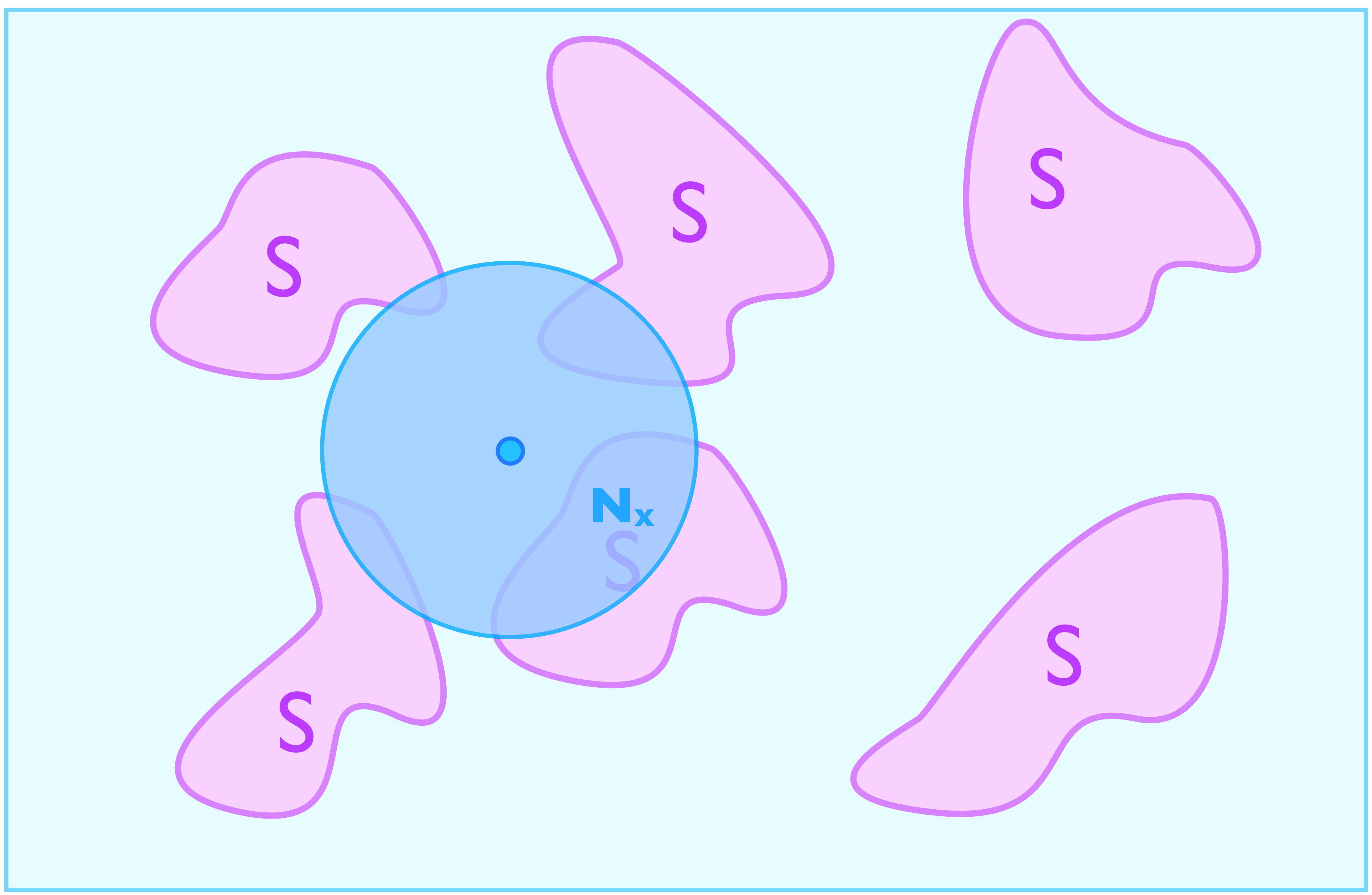
Recall $\mathbf{N_x} = S \cap \mathrm{Ball}(X, n(1-1/B))$



Want to lower bound **H(X' | X)**

i.e., log $|N_x|$ for uniform random x

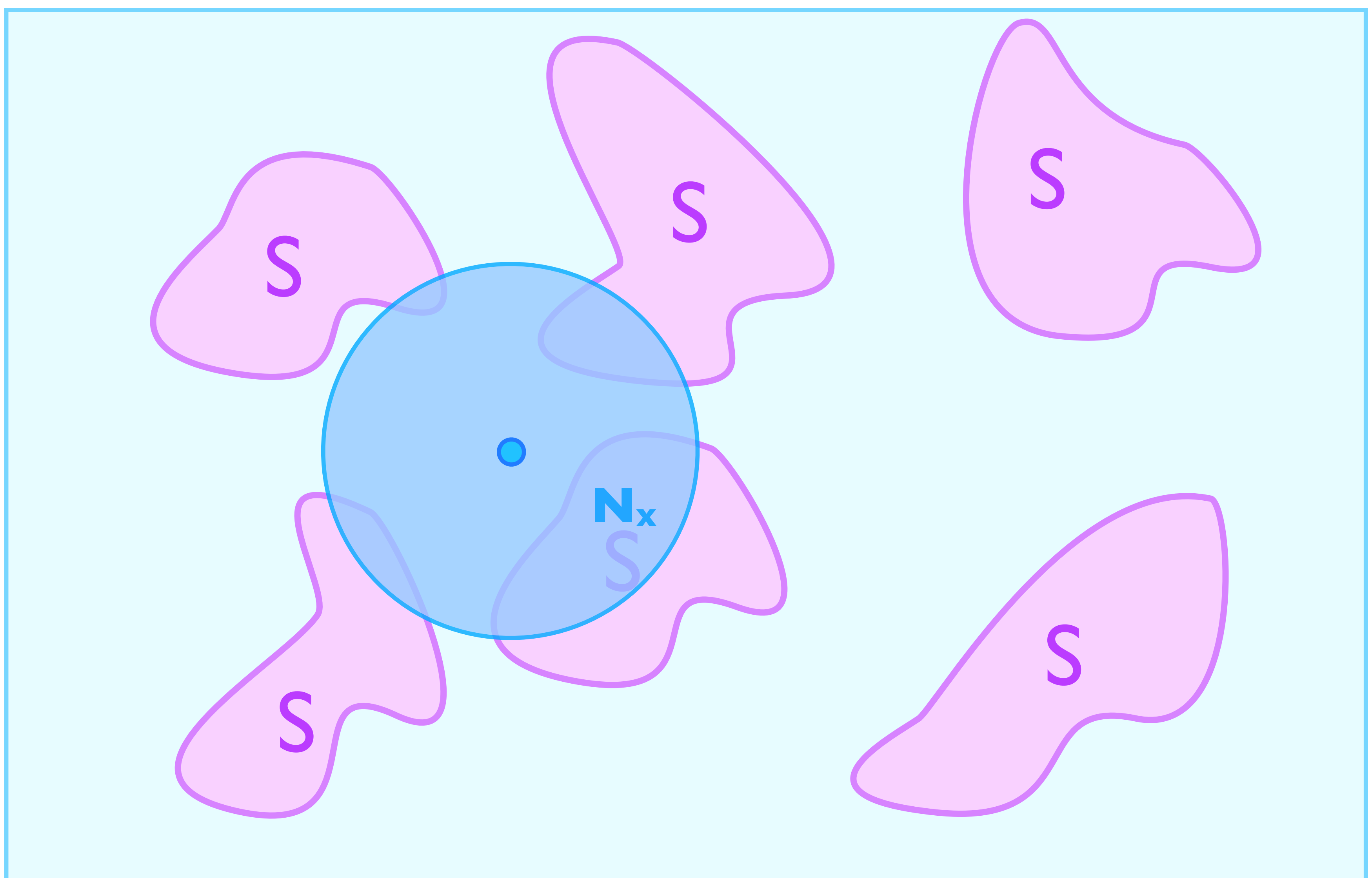
Recall $N_x = S \cap \mathrm{Ball}(X, n(1-1/B))$



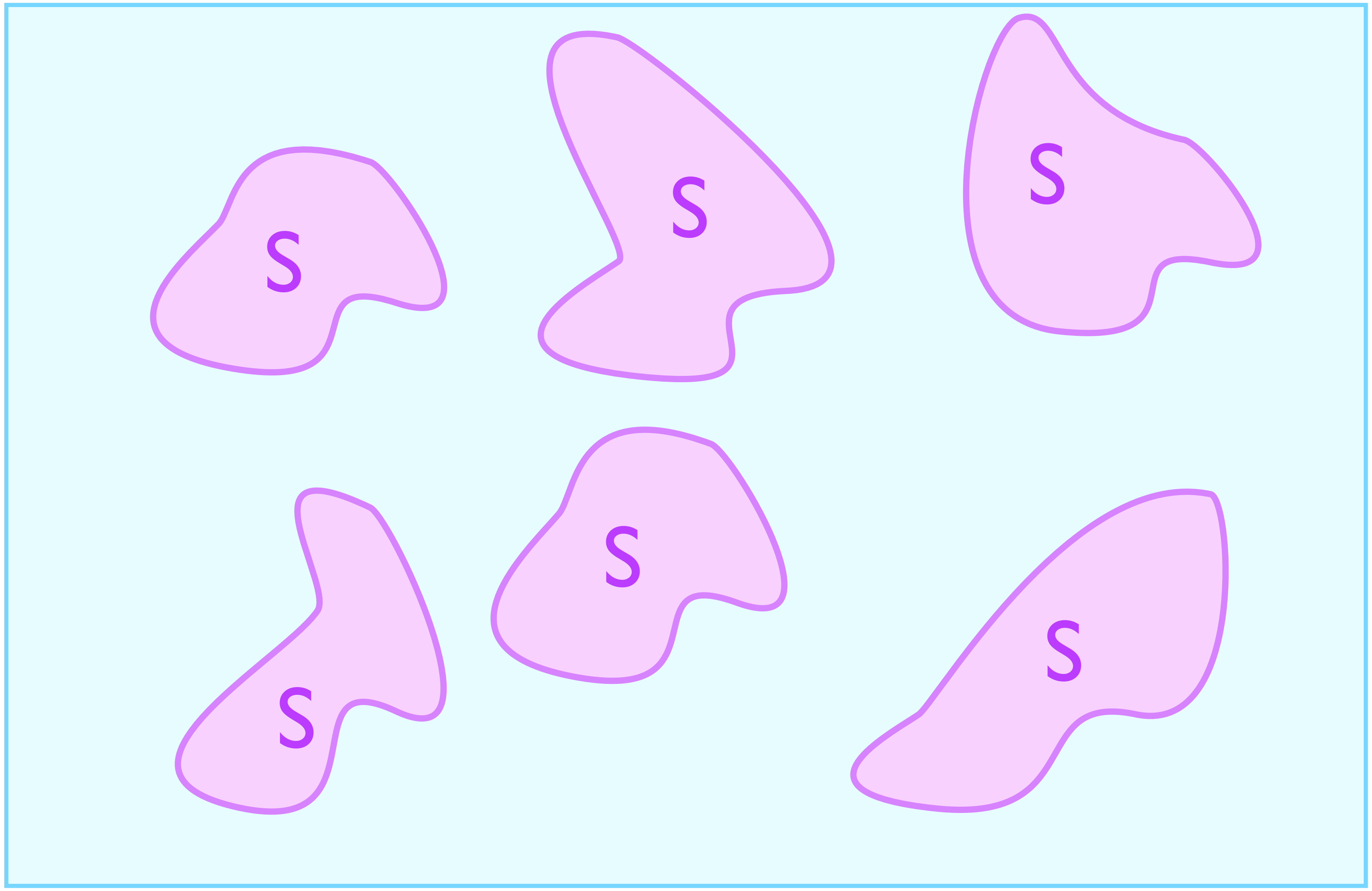
Want to lower bound **H(X' | X)**

i.e., log |$\mathbf{N_x}$| for uniform random x

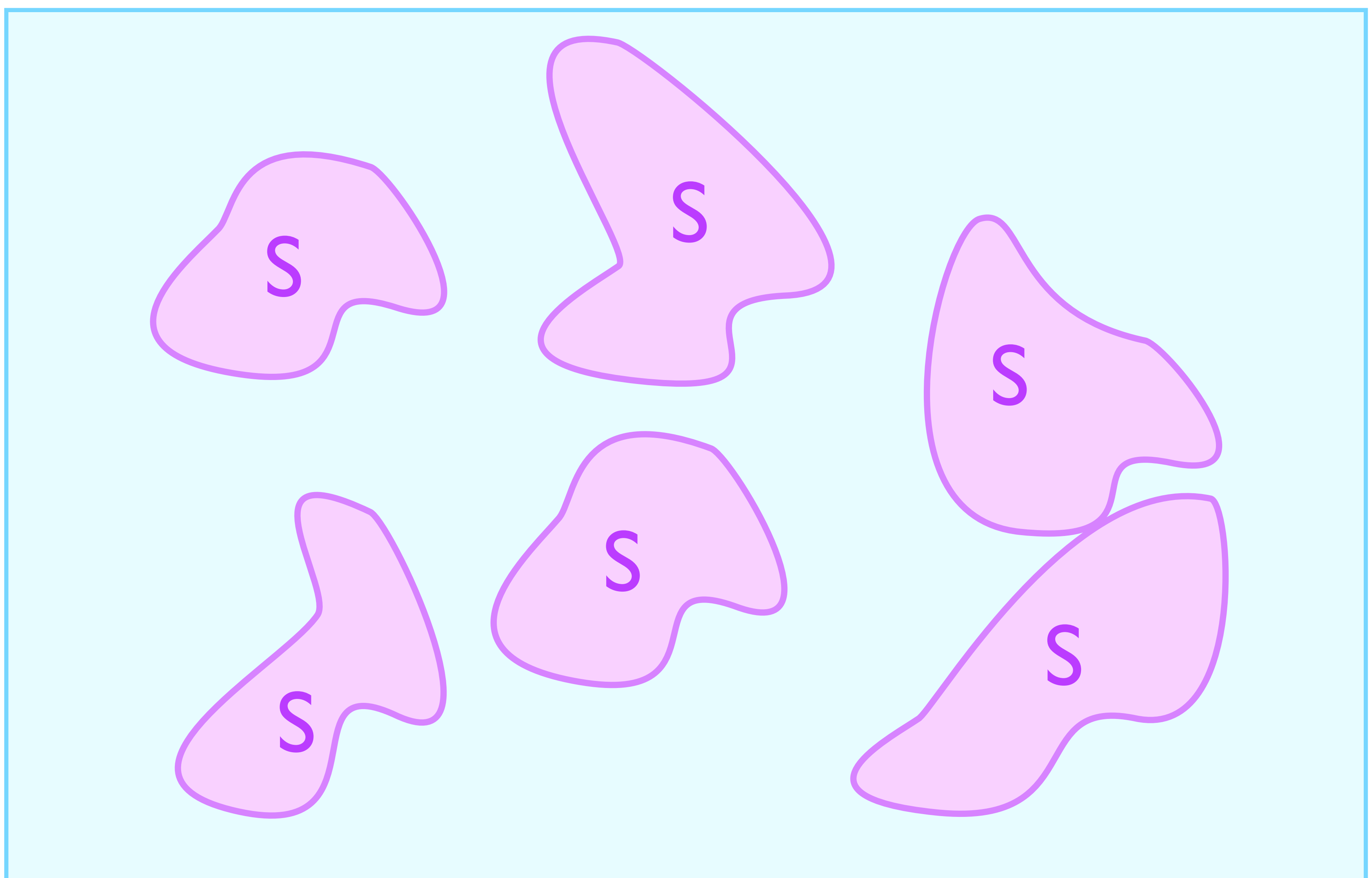
Recall $\mathbf{N_x}$ = S ∩ Ball(X, n(1-1/B))



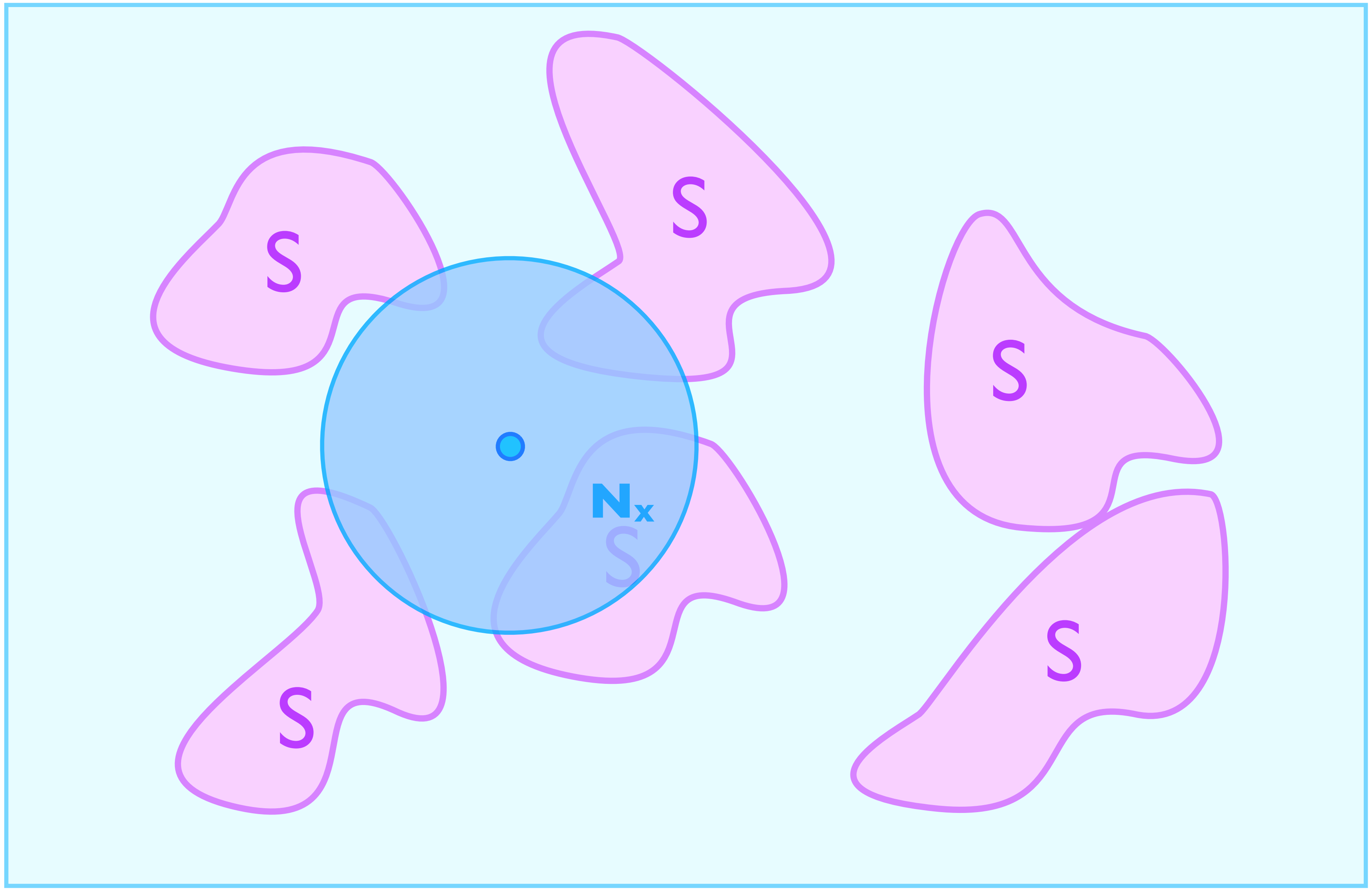
What is the worst case S?



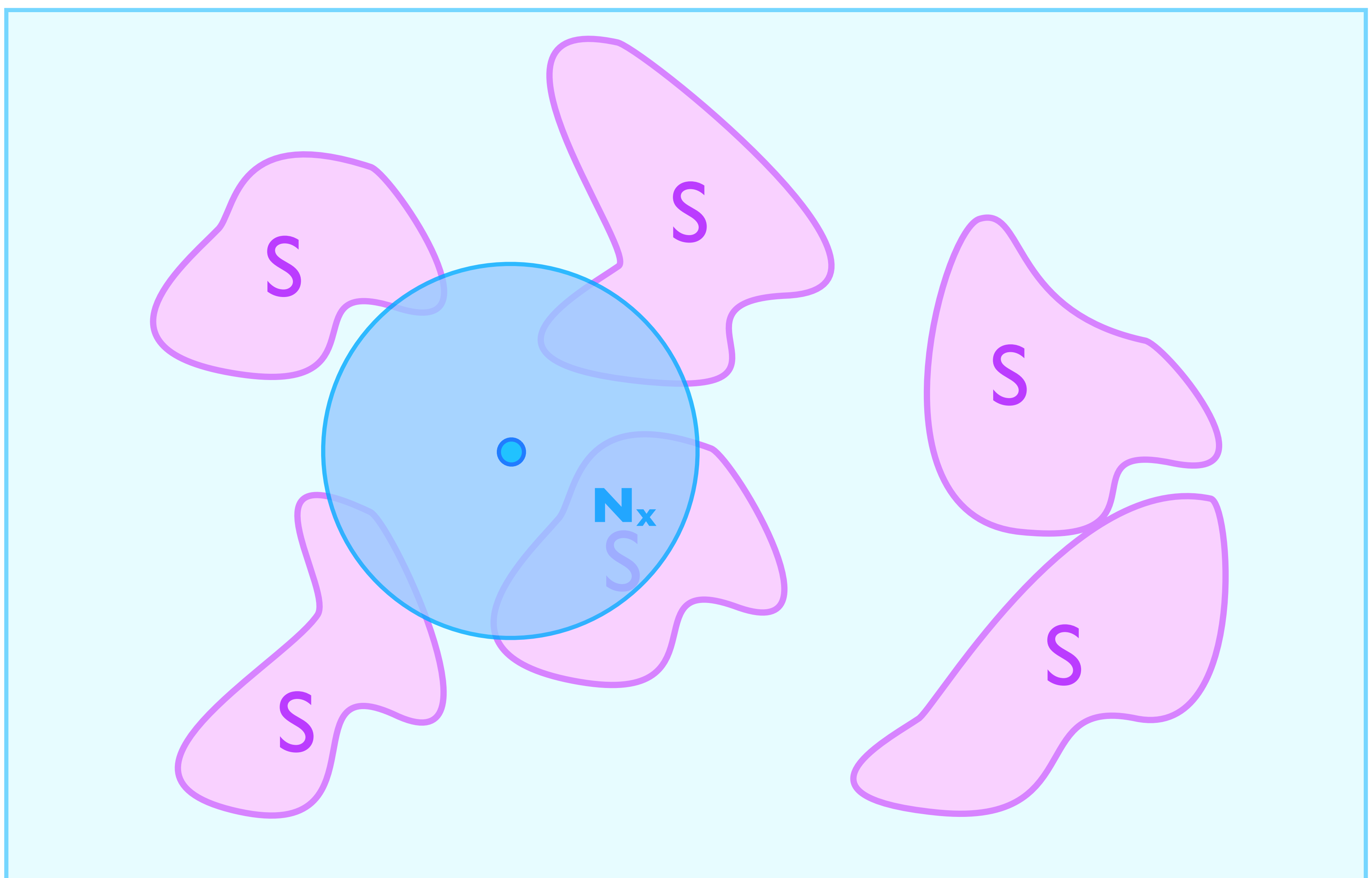
What is the worst case S?



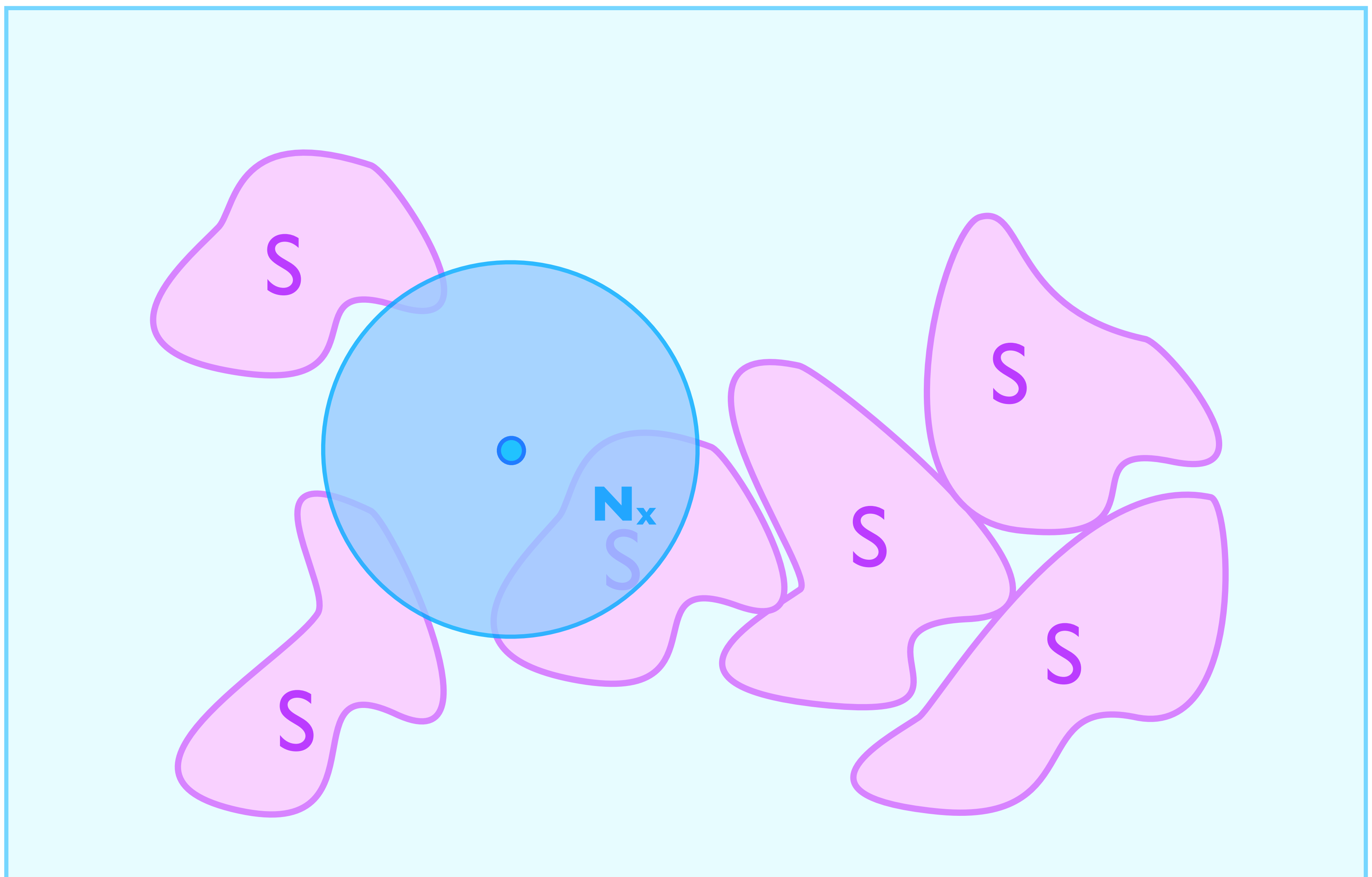
What is the worst case S?



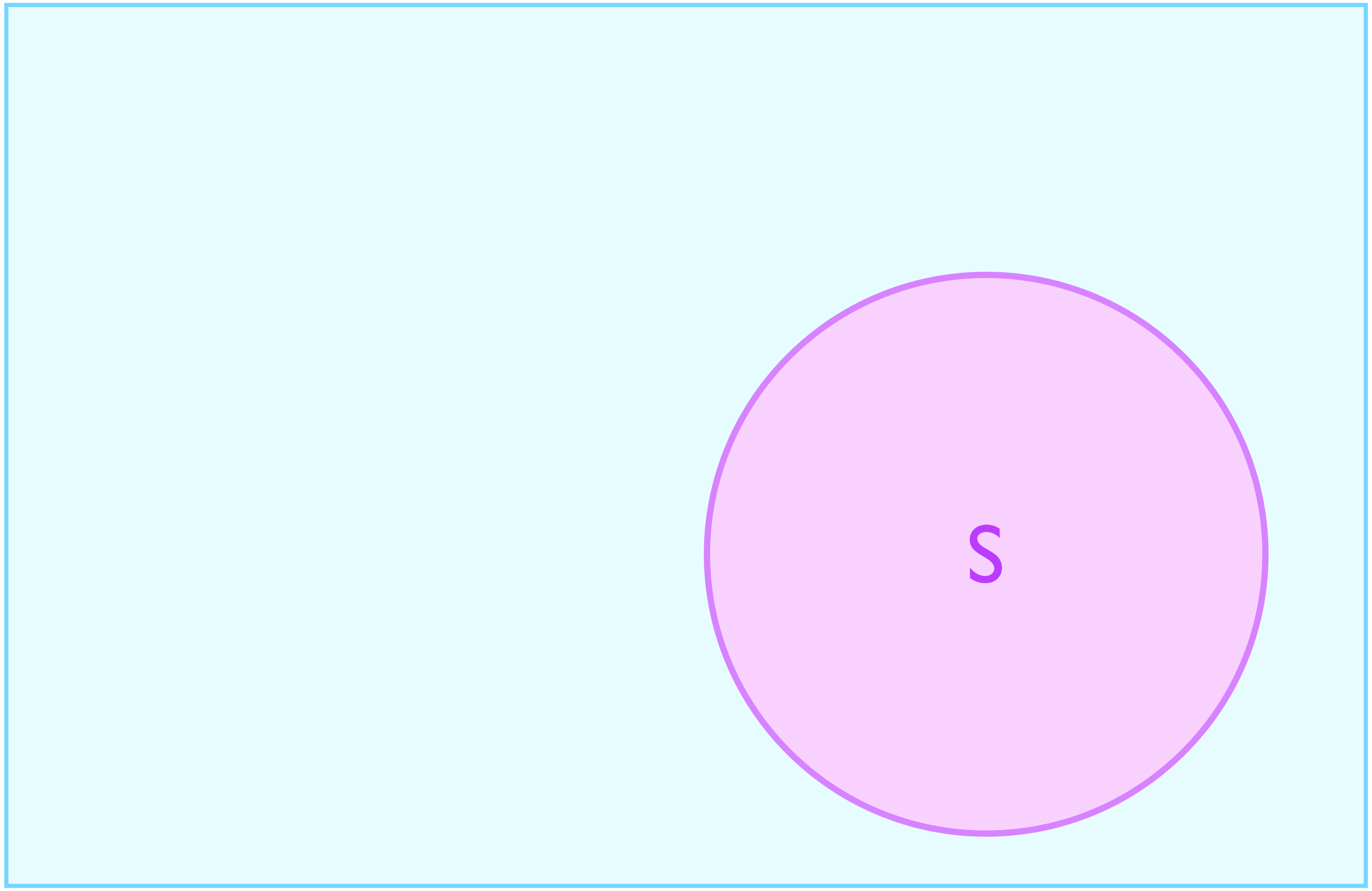
What is the worst case S?



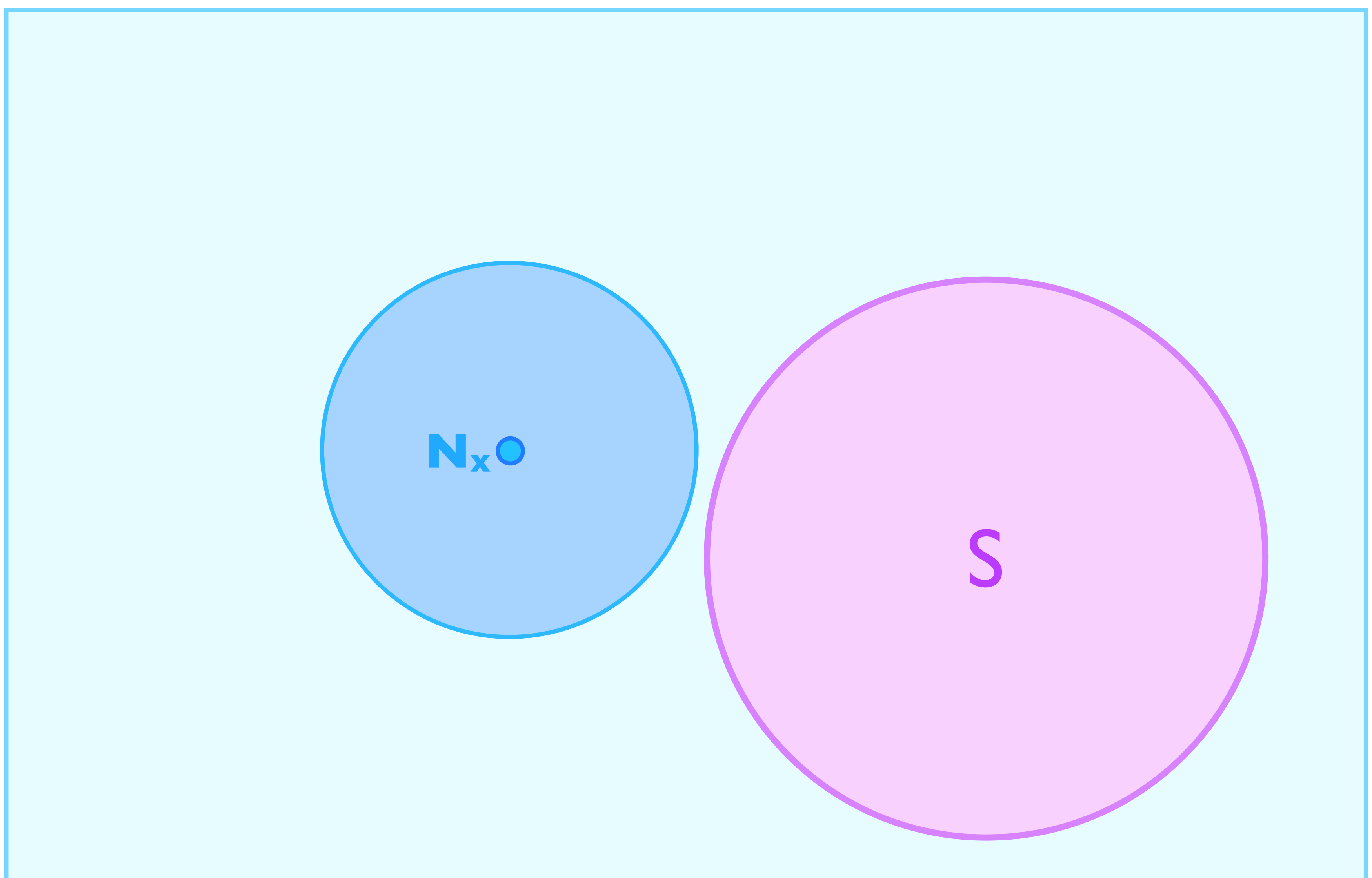
What is the worst case S?



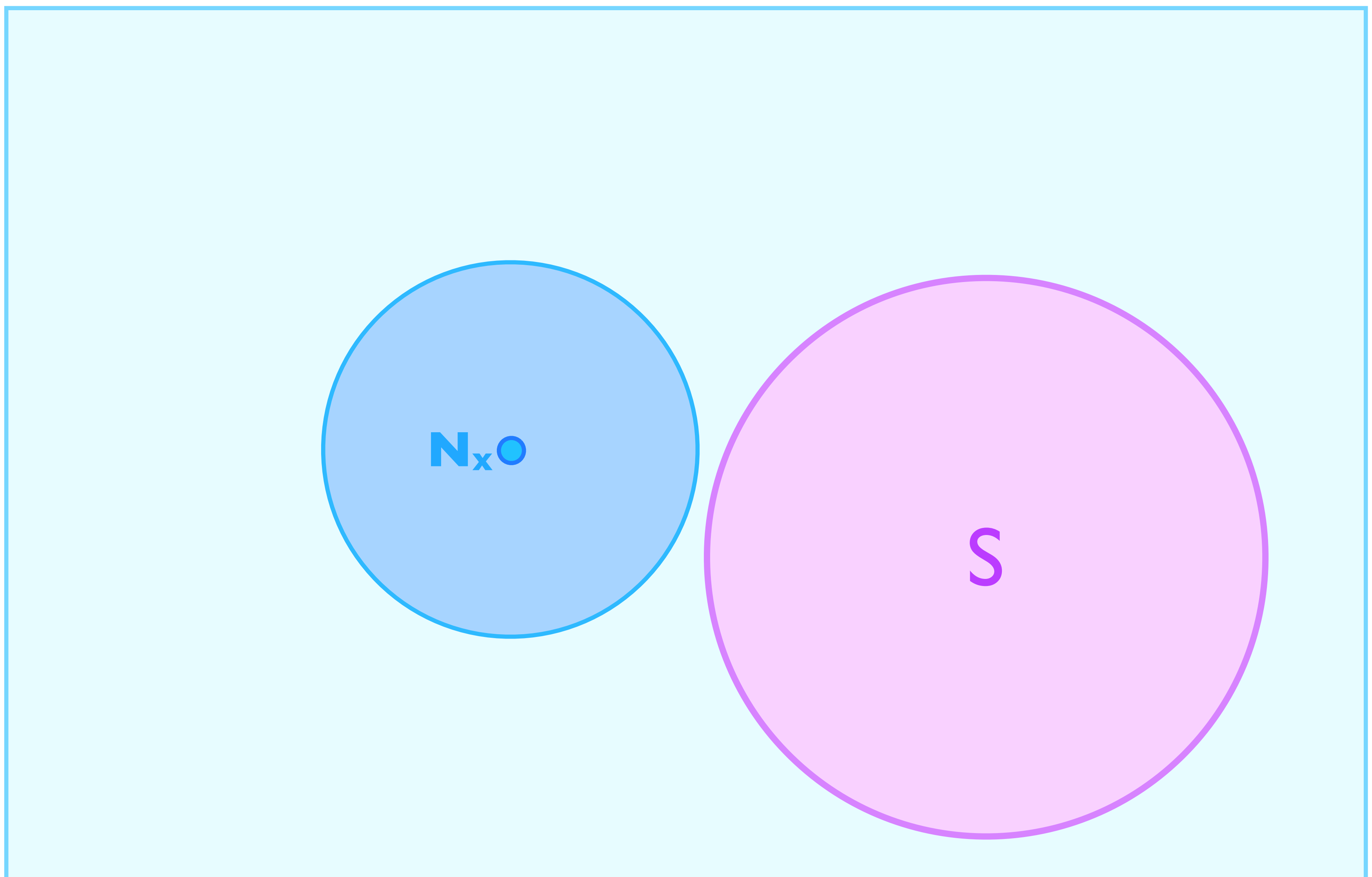
What is the worst case S?



What is the worst case S?

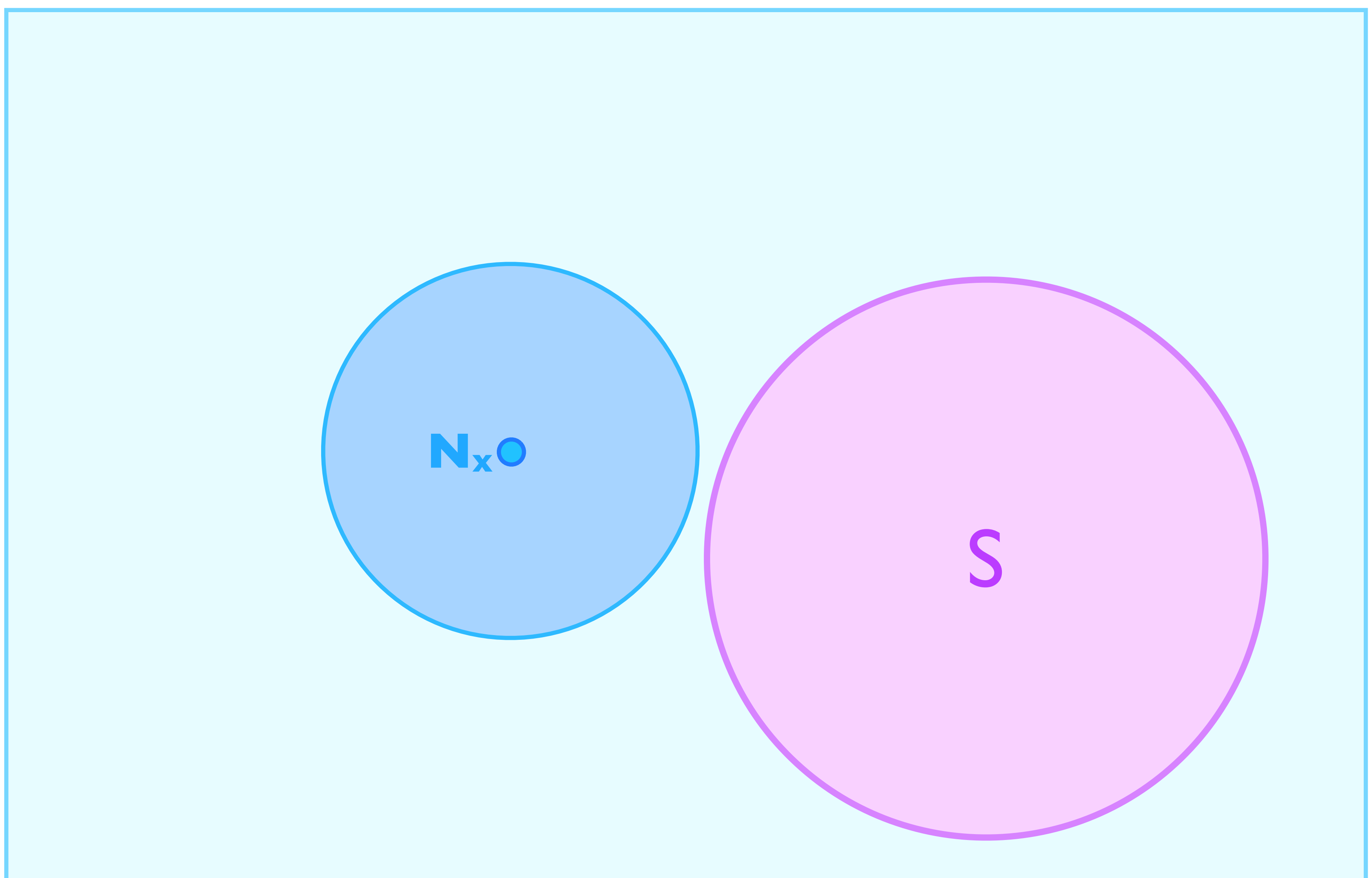


What is the worst case S?



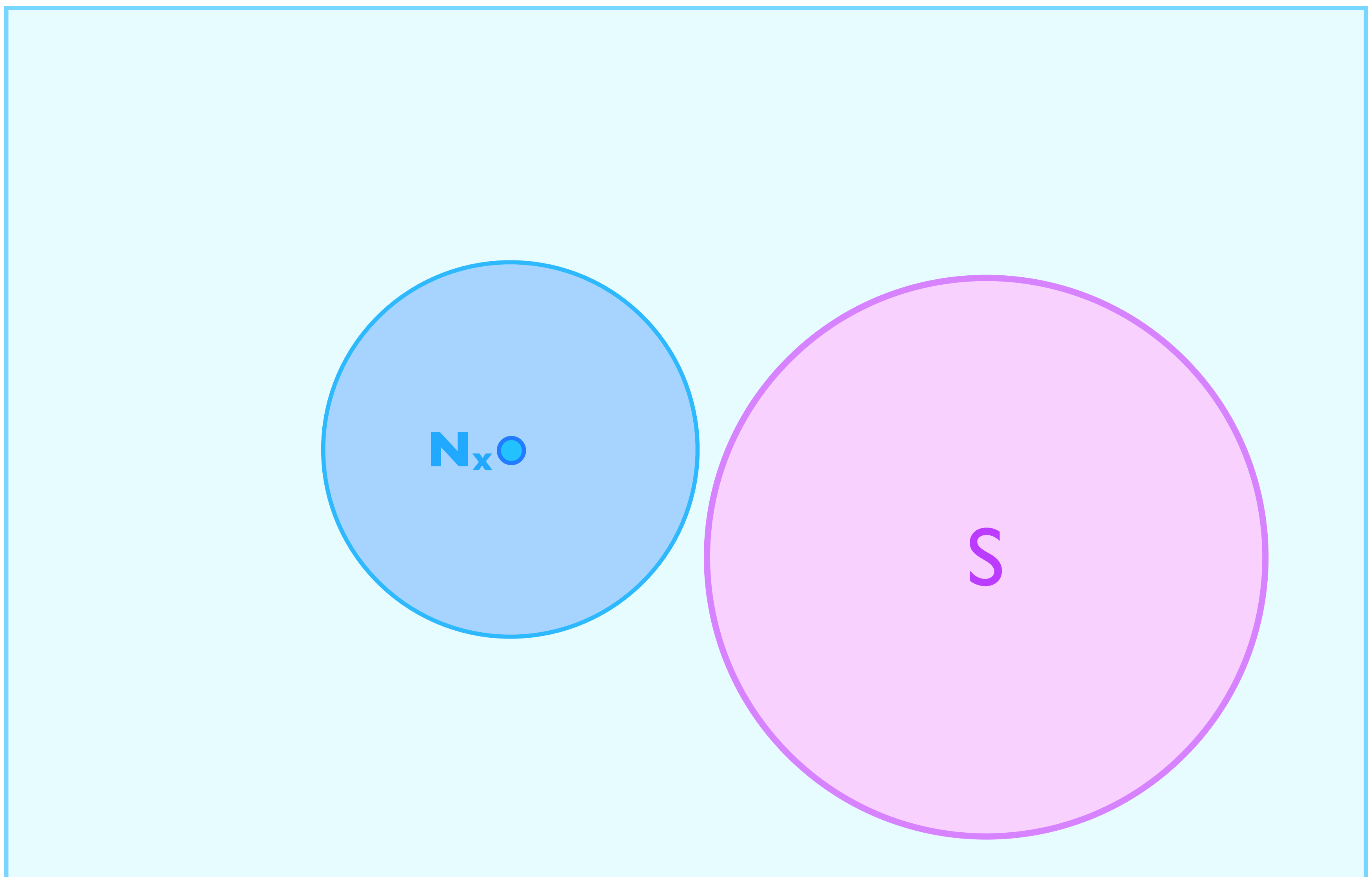
What is the worst case S?

- This intuition is correct even in high dimensions



What is the worst case S?

- This intuition is correct even in high dimensions
- Even for the Hamming distance



An isoperimetric inequality on $[t]^n$

Conjecture: Let $S \subseteq [t]^n$, $|S| = k^n$ $(k < t)$. Then

$$E[\log |B(x, d) \cap S|] \geq E[\log |B(x, d) \cap [k]^n|]$$

where:

x: uniform random

B(x,d): radius-d Hamming ball around x

log 0: -1.

An isoperimetric inequality on $[t]^n$

Conjecture: Let $S \subseteq [t]^n$, $|S| = k^n$ $(k < t)$. Then

$$E[\log |B(x, d) \cap S|] \geq E[\log |B(x, d) \cap [k]^n|]$$

where:

x: uniform random

B(x,d): radius-d Hamming ball around x

log 0: -1.

Theorem (informal): For any $S \subseteq [t]^n$, $\exists\, I \subset [n]$, $|I| = n/5$, the conjecture is true in the projected space

Review

Review

- Tight round / communication tradeoff for DISJ

Review

- Tight round / communication tradeoff for DISJ
- Super-sum for OR of equality: $R(EE_n) = \omega(n) \cdot R(EQ)$

Review

- Tight round / communication tradeoff for DISJ
- Super-sum for OR of equality: $R(EE_n) = \omega(n) . R(EQ)$
- New perspective for direct sum: Isoperimetric considerations

Conjecture: Let $S \subseteq [t]^n$, $|S| = k^n$ $(k < t)$. Then

$$E[\log |B(x,d) \cap S|] \geq E[\log |B(x,d) \cap [k]^n|]$$

where:

- x: uniform random
- $B(x,d)$: radius-d Hamming ball around x
- $\log 0$: -1.

Thank You!