Uniform Direct Product Theorems:

Simplified, Optimized, and Derandomized

Russell Impagliazzo (UCSD and IAS)
Ragesh Jaiswal (UCSD->Columbia)
Valentine Kabanets (SFU)
Avi Wigderson (IAS)

Direct Product(DP) Theorem
(the general statement)

- "If a problem is a hard to solve on average, then solving multiple instances of the problem is even harder".

Applications of such Statements

- Average-case Complexity
- Cryptography
- Derandomization
- Error-correcting codes

Formulating DP Theorems

- "If a problem is a hard to solve on average, then solving multiple instances of the problem is even harder".
  - What is the problem?
    (e.g., computing functions, interactive arguments)
  - What is the entity solving the problem?
    (e.g., circuits, randomized algorithms)
  - What does it mean by a problem being hard on average?

A Simple DP Theorem
(boolean functions against circuits)

- Problem: Computing boolean functions
- Computational model: Circuits
- Hardness: A boolean function $f:\{0,1\}^n \rightarrow \{0,1\}$ is called δ-hard for circuits of size s if for any circuit C of size at most s, we have

$$Pr_x[C(x) \neq f(x)] > \delta$$

A Simple DP Theorem
(boolean functions against circuits)

- Let $f:\{0,1\}^n \rightarrow \{0,1\}$ be a boolean function and f^k defined as

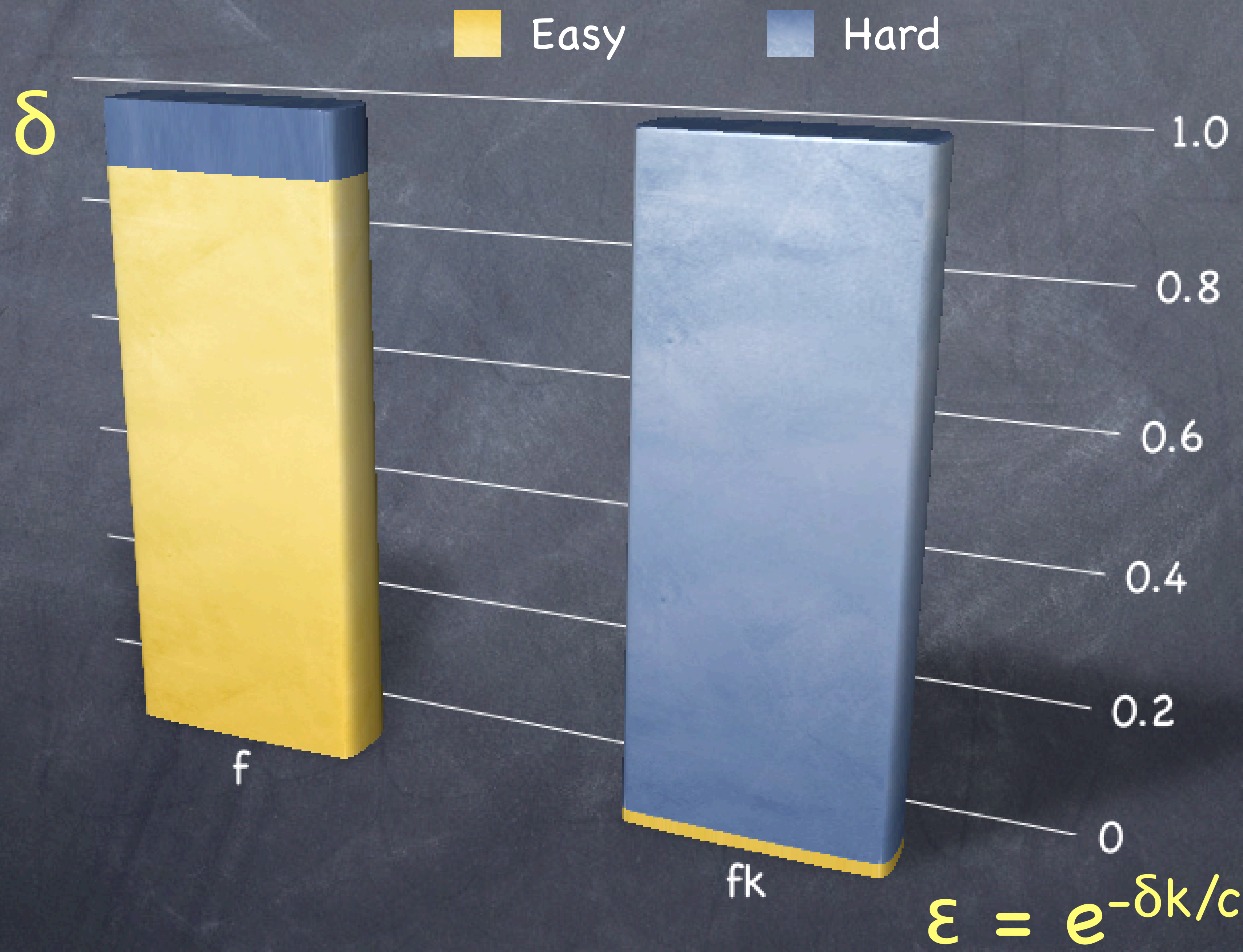
$$f^k(x_1,\dots,x_k) = f(x_1).f(x_2)\dots f(x_k)$$

- If f is δ-hard for circuits of size s, then f^k is $(1-\varepsilon)$-hard for circuits of size s', where

$\delta = \Theta(\log(1/\varepsilon)/k)$ and $s' = s \cdot \mathrm{poly}(\varepsilon,\delta,1/k,1/n)$.

A Simple DP Theorem

(boolean functions against circuits)



A Related XOR Lemma

(boolean functions against circuits)

- Let $f:\{0,1\}^n \to \{0,1\}$ be a boolean function and $f^{\oplus k}$ defined as

$$f^{\oplus k}(x_1,\dots,x_k) = f(x_1)\oplus f(x_2)\oplus\dots\oplus f(x_k)$$

- If f is δ-hard for circuits of size s, then $f^{\oplus k}$ is $(1/2-\varepsilon)$-hard for circuits of size s', where

$\delta = \Theta(\log(1/\varepsilon)/k)$ and $s' = s\cdot \mathrm{poly}(\varepsilon,\delta,1/k,1/n)$.

DP Theorems: A History
(from the perspective of proof idea)

- Levin style Argument [Yao82, Lev87]:
  - Pseudorandom generators
- Impagliazzo's Hard-core set theorem [Imp95]:
  - Hardness of boolean function, Derandomization
- Trust Halving Strategy [IW97, BIN97]:
  - Derandomization, Cryptography

General Proof Strategy

(proof by contradiction)

- Assume: there exists C' such that

$$Pr_{(x1,\ldots,xk)}[C'(x_1,\ldots,x_k) = f^k(x_1,\ldots,x_k)] > \varepsilon$$

- Construct: a circuit C such that

$$Pr_x[C(x) = f(x)] > (1 - \delta)$$

General Proof Strategy

(proof by contradiction)

- Bottleneck: there can possibly exist $f_1,\ldots,f_T$ ($T = 1/\varepsilon$) such that for all $i \in [T]$

$$\Pr_{(x1,\ldots,xk)}[C'(x_1,\ldots,x_k) = f_i^k(x_1,\ldots,x_k)] > \varepsilon$$

General Proof Strategy
(proof by contradiction)

- Assume: there exists C' such that

$$Pr_{(x1,\dots,xk)}[C'(x_1,\dots,x_k) = f^k(x_1,\dots,x_k)] > \varepsilon$$

- Construct: a list of circuit $C_1,\dots,C_T$ such that there exists $i \in [T]$ such that

$$Pr_x[C_i(x) = f(x)] > (1 - \delta)$$

How large could T be?

Nonuniformity in DP Theorems

- A string of length log(T) can be used to point out the correct circuit in the list.
- Generalize the results to general functions f:{0,1}*->{0,1} w.r.t. randomized algorithms with advice (nonuniform model)
- A strong DP Theorem in the uniform model is not possible
- Uniform DP Theorem: A DP theorem with "minimum amount of nonuniformity"

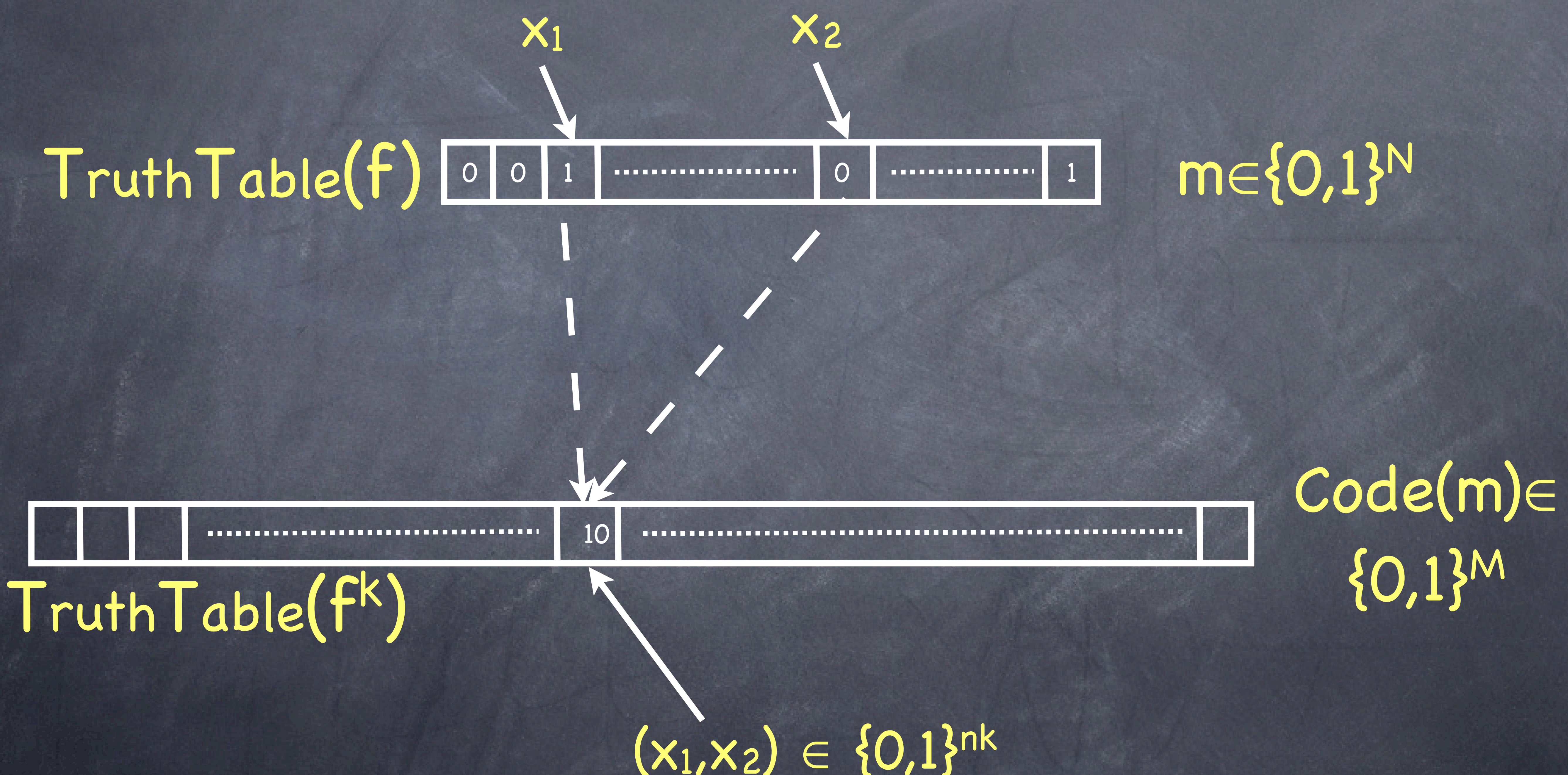
DP Theorem

(a coding theoretic perspective)

- Direct Product code:
  - Let $N = 2^n$, $\Sigma = \{0,1\}^k$, $M = N^k$
  - Message: $m \in \{0,1\}^N$
  - Code: $Code:\{0,1\}^N \rightarrow \Sigma^M$ defined as
    - let each bit of m be indexed by $x \in \{0,1\}^n$ denoted by $m[x]$
    - each alphabet of $Code(m)$ can be indexed by $(x_1,\ldots,x_k)$
    - $Code(m)[(x_1,\ldots,x_k)] = m[x_1].m[x_2]\ldots m[x_k]$

DP Theorem

(a coding theoretic perspective)



Connection with DP Theorem

(a coding theoretic perspective)

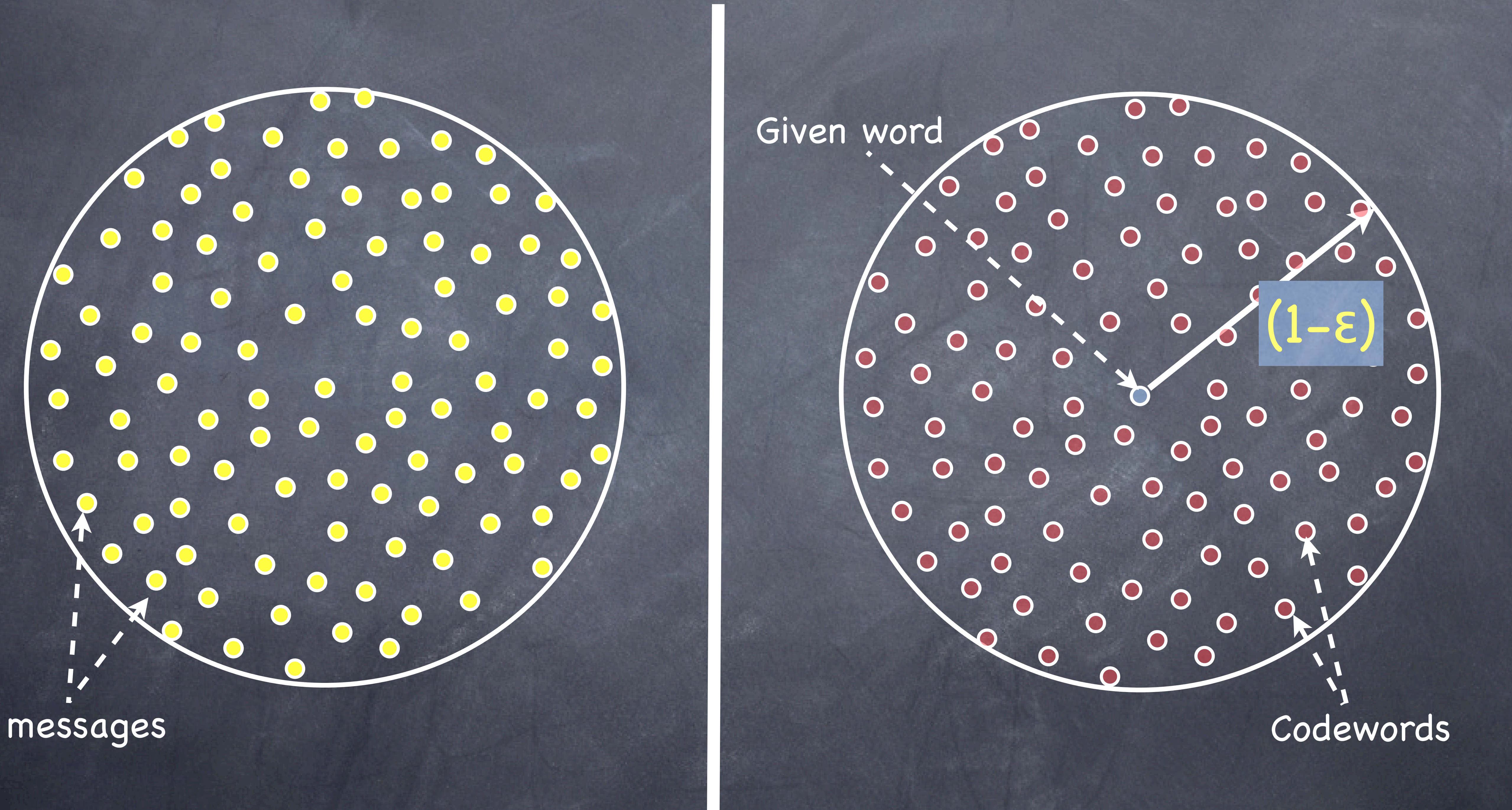
- Any constructive proof of the DP Theorem gives an approximate, local, list decoding algorithm for DP code.

A circuit which computes the corrupted codeword

List of circuits such that at least one of them approximately computes the message

DP Theorem

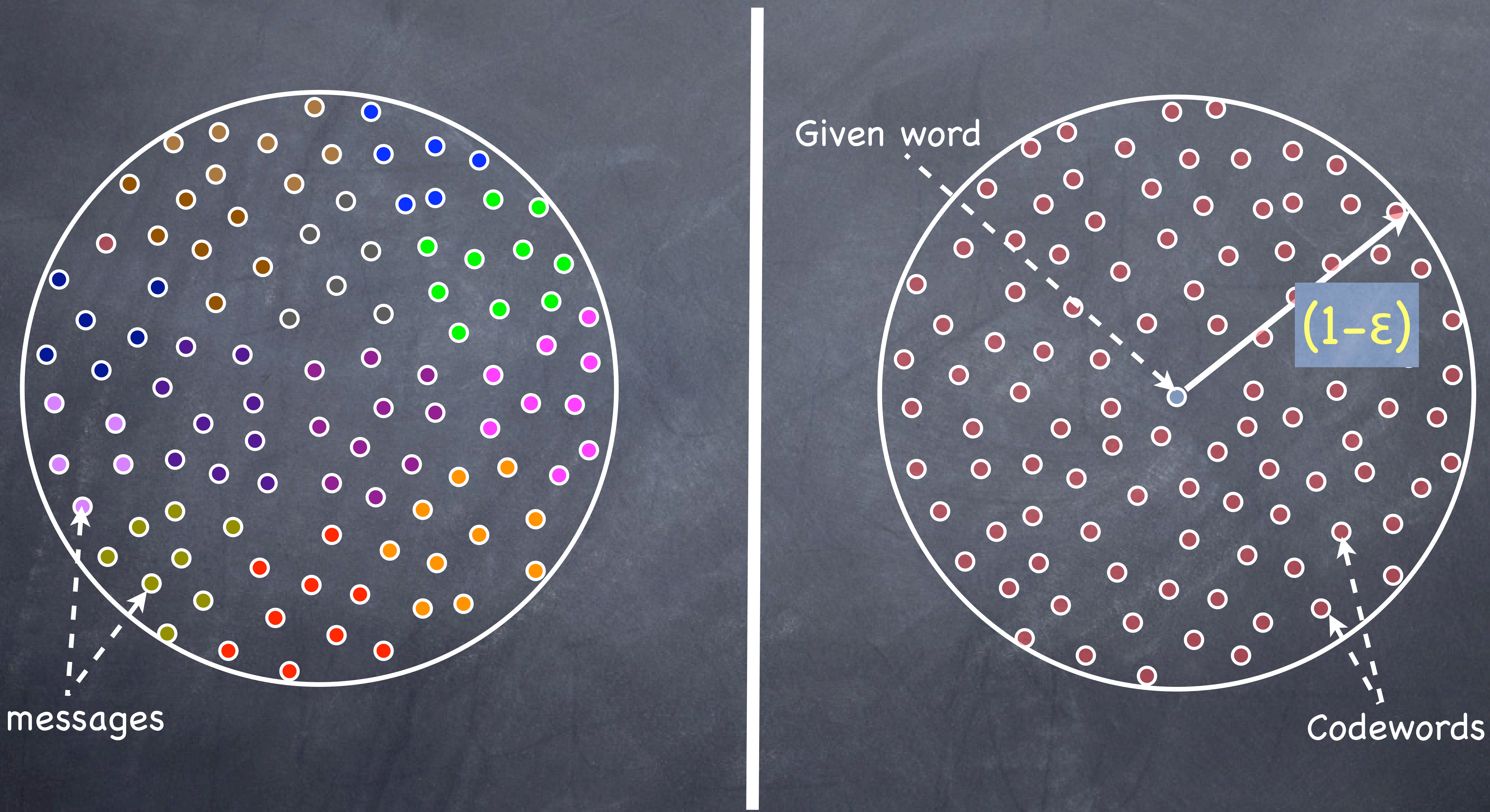
(a coding theoretic perspective)



List Decoding

DP Theorem

(a coding theoretic perspective)



Approximate list decoding

DP Theorem
(a coding theoretic perspective)

- Let $\delta = \Theta(\log(1/\varepsilon)/k)$
- For any message m and its corrupted codeword $w \in \{0,1\}^N$ such that $Ham(Code(m),w) < (1-\varepsilon).M$, then there are $T = \Theta(1/\varepsilon)$ messages $m_1, \ldots, m_T$ such that for at least one m_i, $Ham(m_i,m) < \delta \cdot N$

Bounds for the Related XOR Code

- Let $\delta = \Theta(\log(1/\varepsilon)/k)$
- Given a message m and its corrupted codeword $w \in \{0,1\}^N$ such that $Ham(XOR\text{-}Code(m),w) < (1/2-\varepsilon).M$, then there are $T=\Theta(1/\varepsilon^2)$ messages $m_1,\ldots,m_T$ such that for at least one m_i, $Ham(m_i,m) < \delta \cdot N$

DP Theorem

(a coding theoretic perspective)

- All previous proofs [Lev87, Imp95, IW97...] of the DP theorem gave list size $2^{poly(1/\varepsilon)}$.
- [IJK06, IJKW08]: List decoding algorithm with size $\Theta(1/\varepsilon)$ which is information theoretically optimal.

Uniform DP Theorem

(the first attempt)

- Main Theorem [IJK06]: Let $f:U\to\{0,1\}$ be some function and C' be a circuit such that $\Pr[C' \text{ computes } f^k] > \varepsilon$.
  There is an algorithm which outputs a list of circuits $C_1,\ldots,C_T$ such that
  $\exists i, \Pr[C_i \text{ computes } f] > (1-\delta)$, where
  $\varepsilon=\text{poly}(1/k)$, $\forall i, |C_i|=|C'|\cdot\text{poly}(1/\varepsilon,1/\delta,k)$, $T=\text{poly}(1/\varepsilon)$.
- Drawbacks:
  - Worked for large ε.
  - Complicated algorithm and analysis.

Uniform DP Theorem
(the final attempt)

- Main Theorem [IJKW08]: Let $f:U\to R$ be some function and C' be a circuit such that $\Pr[C' \text{ computes } f^k] > \varepsilon$.
  There is an algorithm which outputs a list of circuits $C_1,\ldots,C_T$ such that $\exists i, \Pr[C_i \text{ computes } f] > (1-\delta)$, where $\delta=\Theta(\log(1/\varepsilon)/k)$, $\forall i, |C_i|=|C'|\cdot \text{poly}(1/\varepsilon,1/\delta,k)$, $T=O(1/\varepsilon)$.

Uniform XOR Lemma

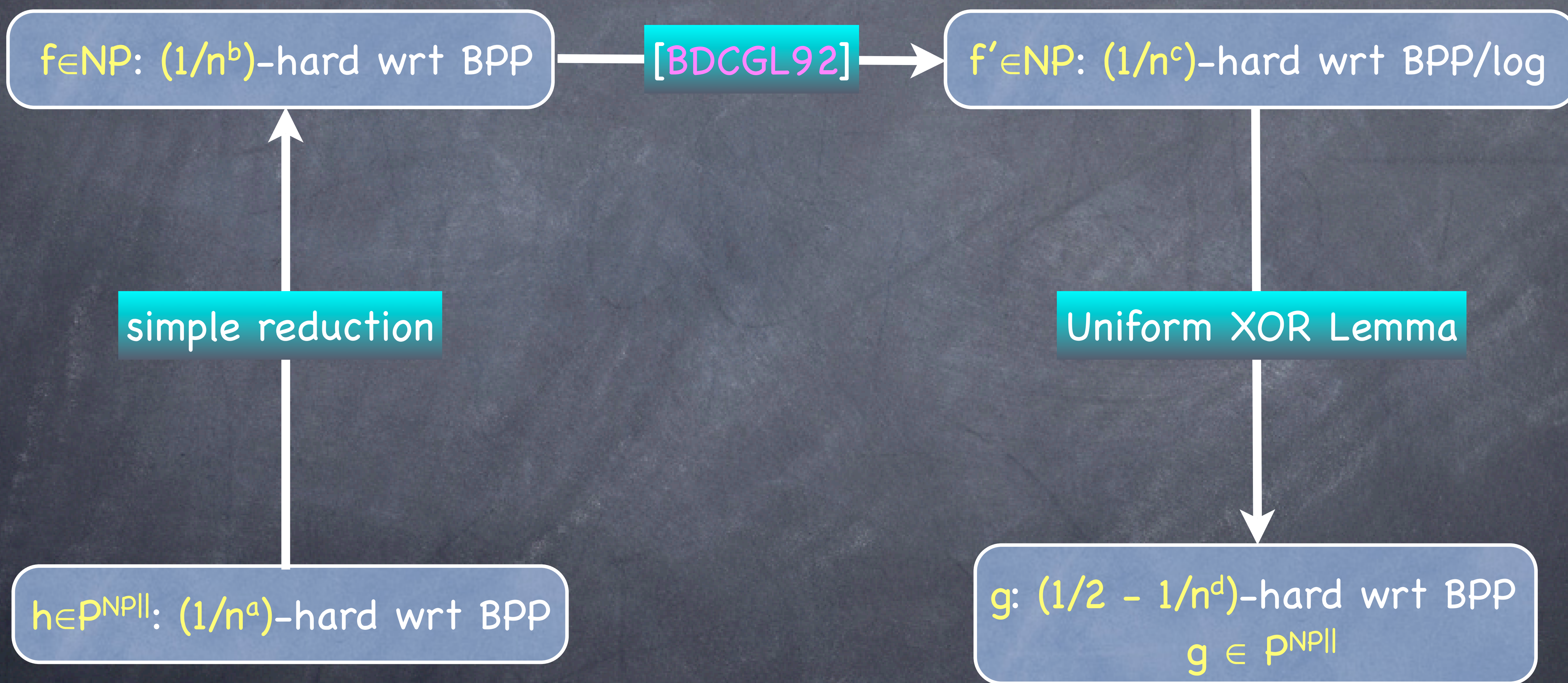
- Theorem [IJKW08]: Let $f:U\to\{0,1\}$ be some function and C' be a circuit such that
$\Pr[C' \text{ computes } f^{\oplus k}] > 1/2 + \varepsilon$.
There is an algorithm which outputs a list of circuits $C_1,\ldots,C_T$ such that
$\exists i, \Pr[C_i \text{ computes } f] > (1-\delta)$, where
$\delta=\Theta(\log(1/\varepsilon)/k)$, $\forall i, |C_i|=|C'|\cdot \mathrm{poly}(1/\varepsilon,1/\delta,k)$, $T=O(1/\varepsilon^2)$.

Uniform Hardness Amplification

- Average-case Complexity: Average-case hardness of problems instead of worst-case.
- Uniform hardness amplification within C: If there is a problem within C which is mildly hard on average for probabilistic polynomial time algorithms, then is there another problem in C which is very hard for probabilistic polynomial time algorithms.

Uniform Hardness Amplification

(Hardness Amplification within $P^{NP||}$)



$P^{NP||}$: polynomial time turing machine which can make polynomial parallel oracle queries to an NP oracle.

Uniform Direct Product Theorem: The Proof

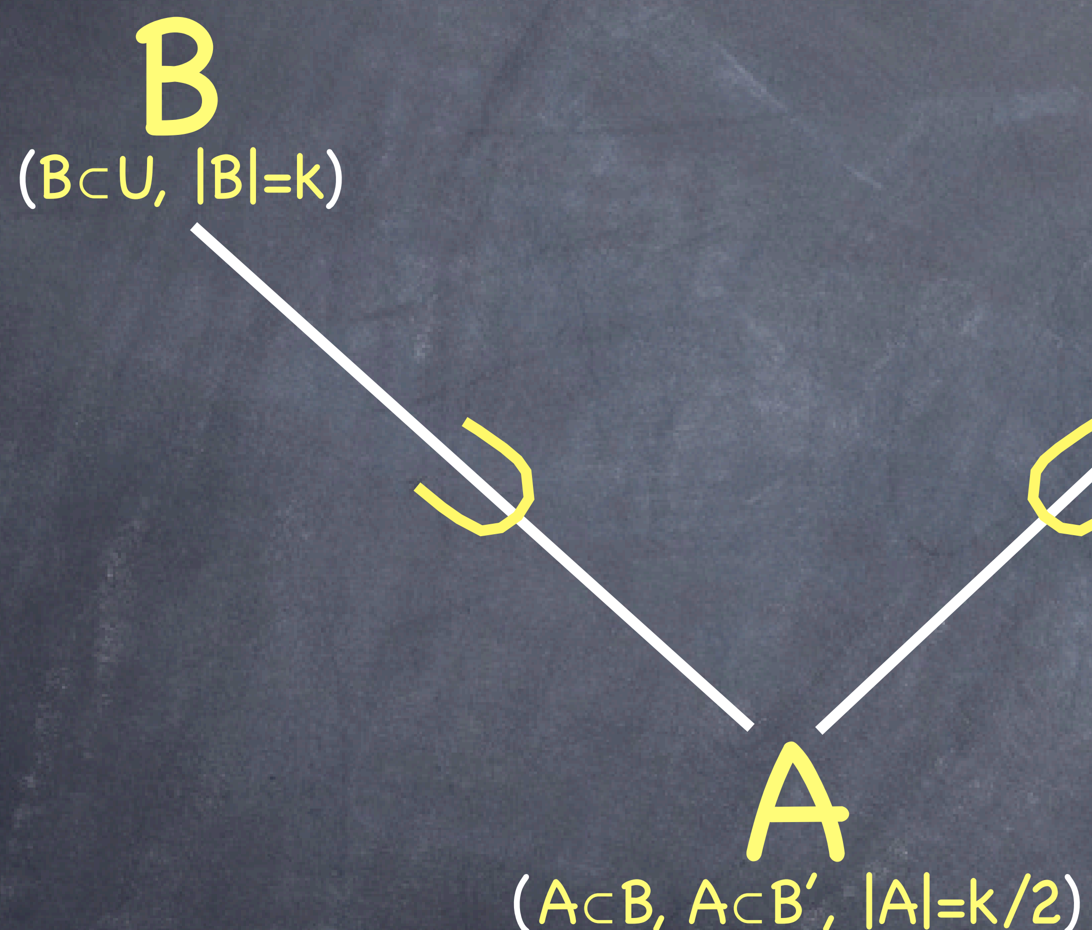
Main Theorem

- Theorem [IJKW08]: Let $f:U\to R$ be some function and C' be a circuit such that
  Pr[C' computes f^k] > ε.
  There is an algorithm which outputs with probability $\Omega(\varepsilon)$ a circuit C such that
  Pr[C computes f] > $(1-\delta)$,
  where $\delta=\Theta(\log(1/\varepsilon)/k)$, $|C|=|C'|\cdot \mathrm{poly}(1/\varepsilon,1/\delta,k)$.

Main Theorem

- Previous Theorem => Uniform DP Theorem
  - Repeat the algorithm $O(1/\varepsilon)$ times to produce a list of circuits.

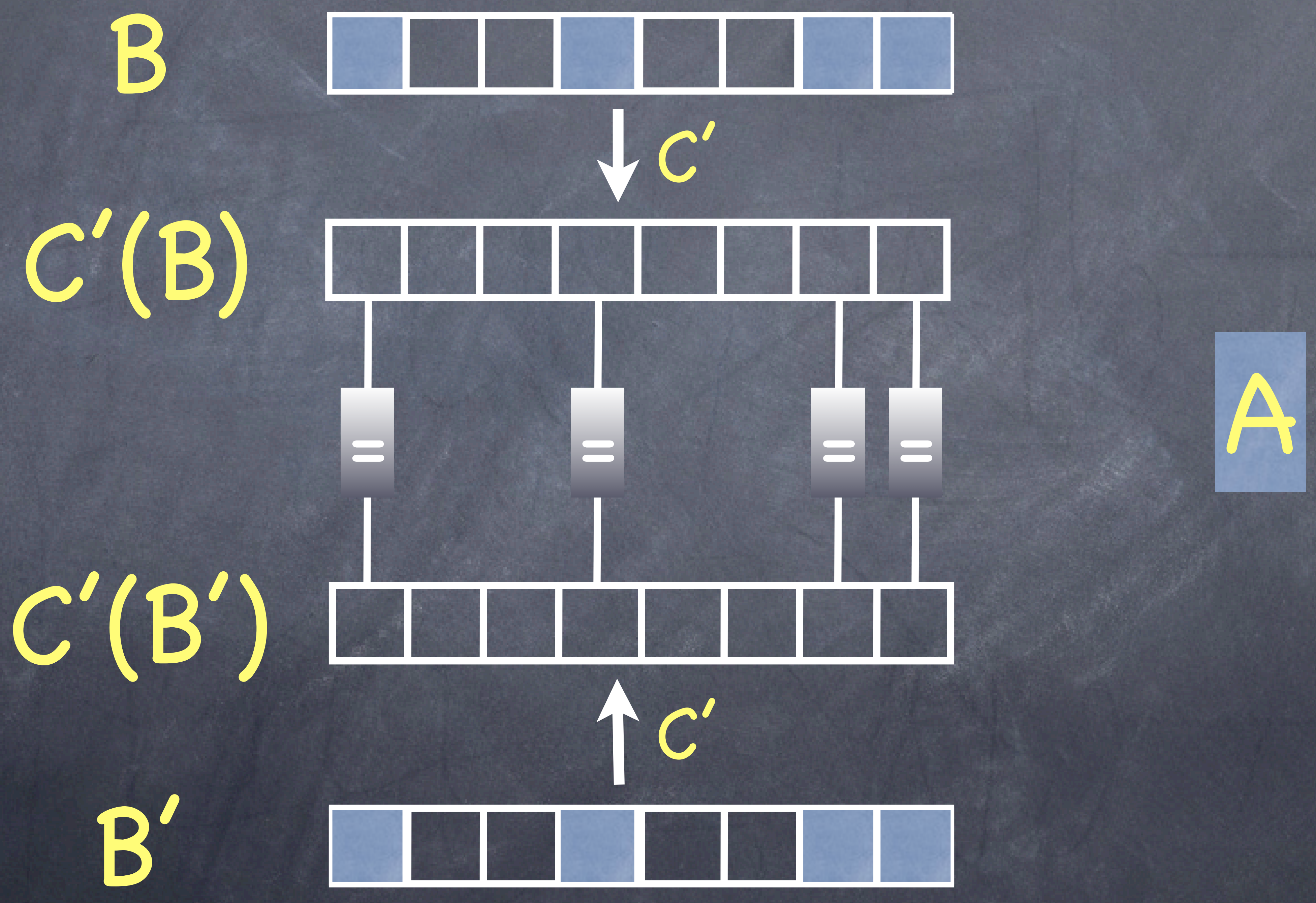
Local Consistency Test



B′ is said to pass the consistency test wrt (A,B) if

$$C'(B)|_A = C'(B')|_A$$

Local Consistency Test



An Idea Based on Local Consistency Test

- Suppose there are sets $A, B \supset A$ such that
  - $C'(B) = f^k(B)$
  - "some other nice properties"
- $C_{A,B}$: Given an input $x \in U$
  - If $x \in B$, then output $C'(B)[x]$
  - Randomly select B', such that $A \subset B'$ and $x \in B'$
  - If B' passes consistency test wrt (A,B), then output $C'(B')[x]$ else repeat

When does $C_{A,B}$ work?

- Under what conditions does $C_{A,B}$ work?
- Under what conditions "local consistency implies correctness"?
- What are the "nice properties" A,B need to satisfy?

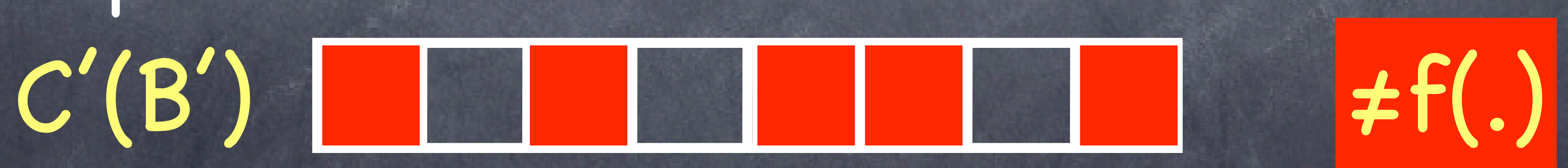
When does $C_{A,B}$ work?

- Under what conditions does $C_{A,B}$ work?

  (1) $C'(B) = f^k(B)$

  (2) There are non-negligible number of $B' \supset A$ s.t. $C'(B')=f^k(B)$ and which pass the consistency test wrt. (A,B)

  (3) "Bad" $B' \supset A$ fail the consistency test w.h.p.



- Let us call such (A,B) "excellent".

Choosing Excellent (A,B)

- Choose $A, B \supset A$ randomly
- <u>Lemma:</u> $\Pr_{A,B\supset A}[(A,B) \text{ is excellent}] = \Omega(\varepsilon)$

Choosing Excellent (A,B)

(Proof: $Pr_{A,B\supset A}[(A,B) \text{ is excellent}] = \Omega(\varepsilon)$)

- Recall (1) $C'(B) = f^k(B)$
- Since $Pr_B[C'(B) = f^k(B)] > \varepsilon$, randomly chosen $A, B\supset A$ satisfies (1) with probability at least ε.
- We will try to show that (2) and (3) almost always follows from (1).

Choosing Excellent (A,B)

(Proof: $\Pr_{A,B\supset A}[(A,B) \text{ is excellent}] = \Omega(\varepsilon)$)

- Recall: (2) There are non-negligible number of $B'\supset A$ s.t. $C'(B')=f^k(B)$ and which pass the consistency test wrt. A,B
- (2) almost always follows from (1):
  - Let $P(A)$ be the event that $\Pr_{B\supset A}[C'(B) = f^k(B)] \leq \varepsilon/2$
  - $\Pr_{A,B\supset A}[C'(B) = f^k(B) \mid P(A)] \leq \varepsilon/2$ => $\Pr_{A,B\supset A}[C'(B) = f^k(B) \;\&\; P(A)] \leq \varepsilon/2$

Choosing Excellent (A,B)

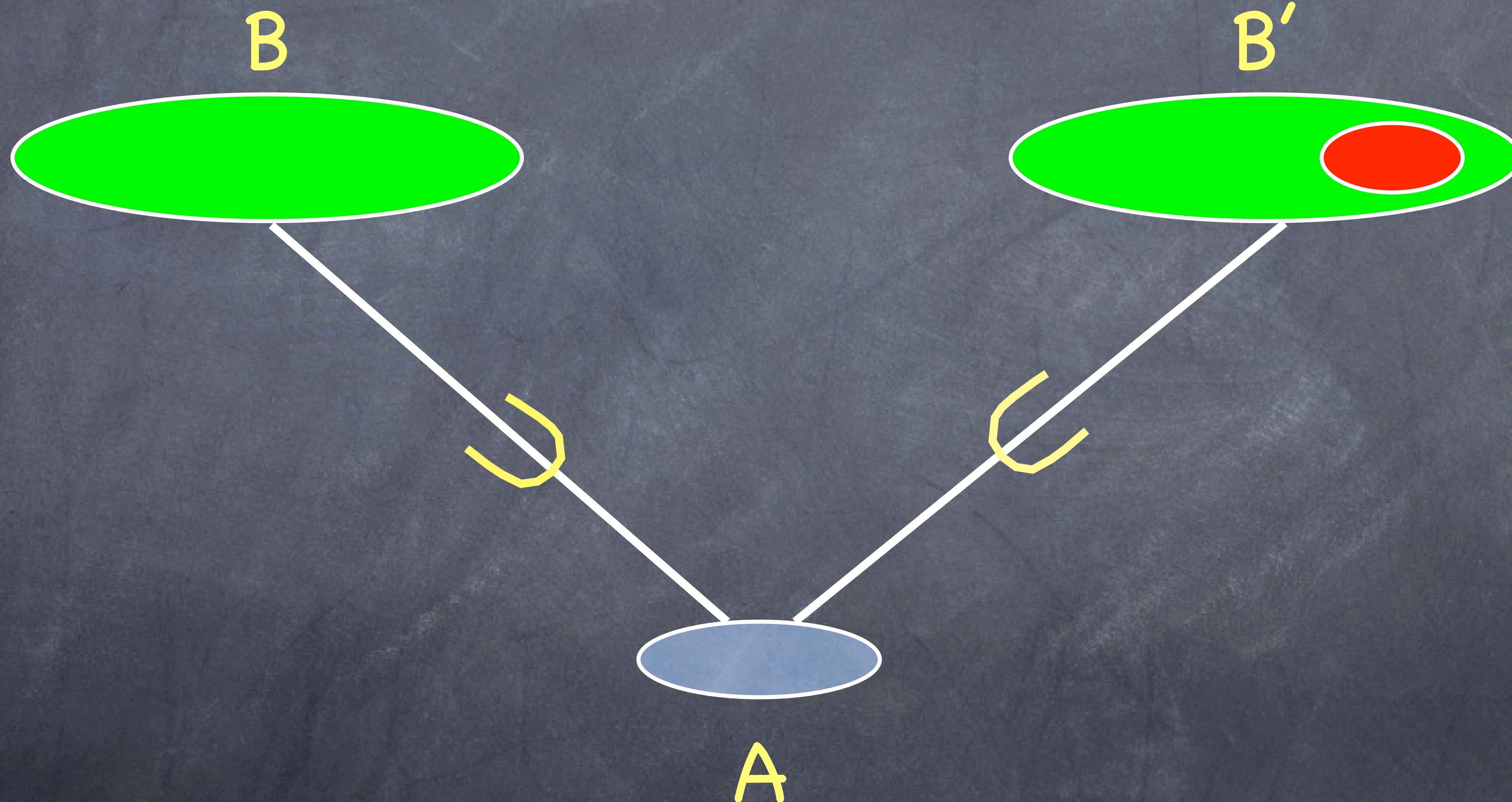
(Proof: $\Pr_{A,B\supset A}[(A,B) \text{ is excellent}] = \Omega(\varepsilon)$)

- Recall: (3) "Bad" $B'\supset A$ fail the consistency test w.h.p.
- (3) almost always follows from (1):
  - We want to show:

    $\Pr_{A,B\supset A,B'\supset A}[C'(B)=f^k(B)$ & B' is "bad" & B' passes consistency test wrt $(A,B)]$ is very small (say $< \varepsilon^3$)

Choosing Excellent (A,B)

(Proof: $\Pr_{A,B\supset A}[(A,B) \text{ is excellent}] = \Omega(\varepsilon)$)

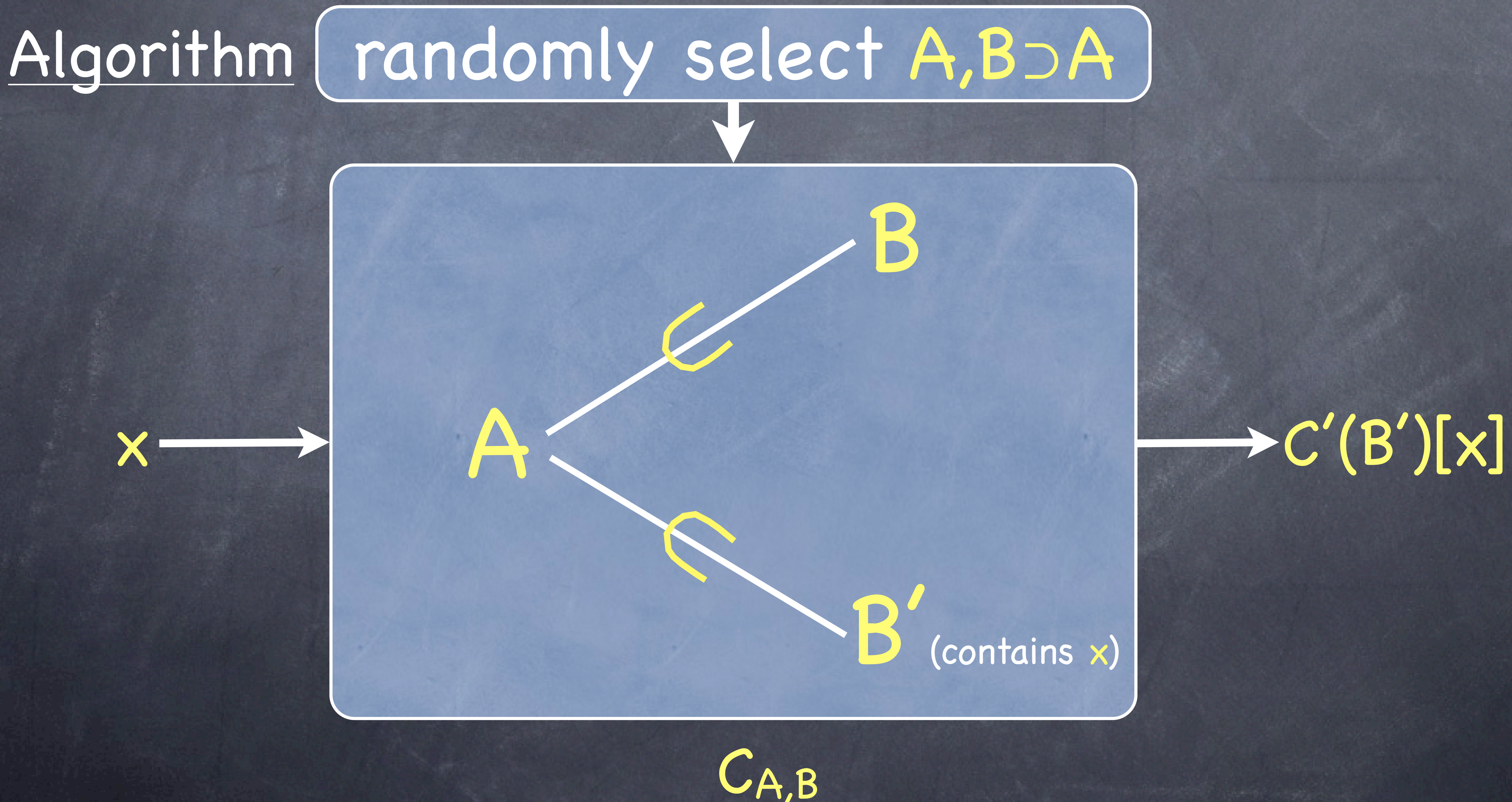


w.h.p A contains a “bad” element of B′

Where we are in the proof

- What we have shown:
  - Lemma: $\Pr_{A,B\supset A}[(A,B)$ is excellent$] = \Omega(\varepsilon)$
- What we need to show:
  - Lemma: For any excellent (A,B), $C_{A,B}$ computes f with probability at least $(1-\delta)$

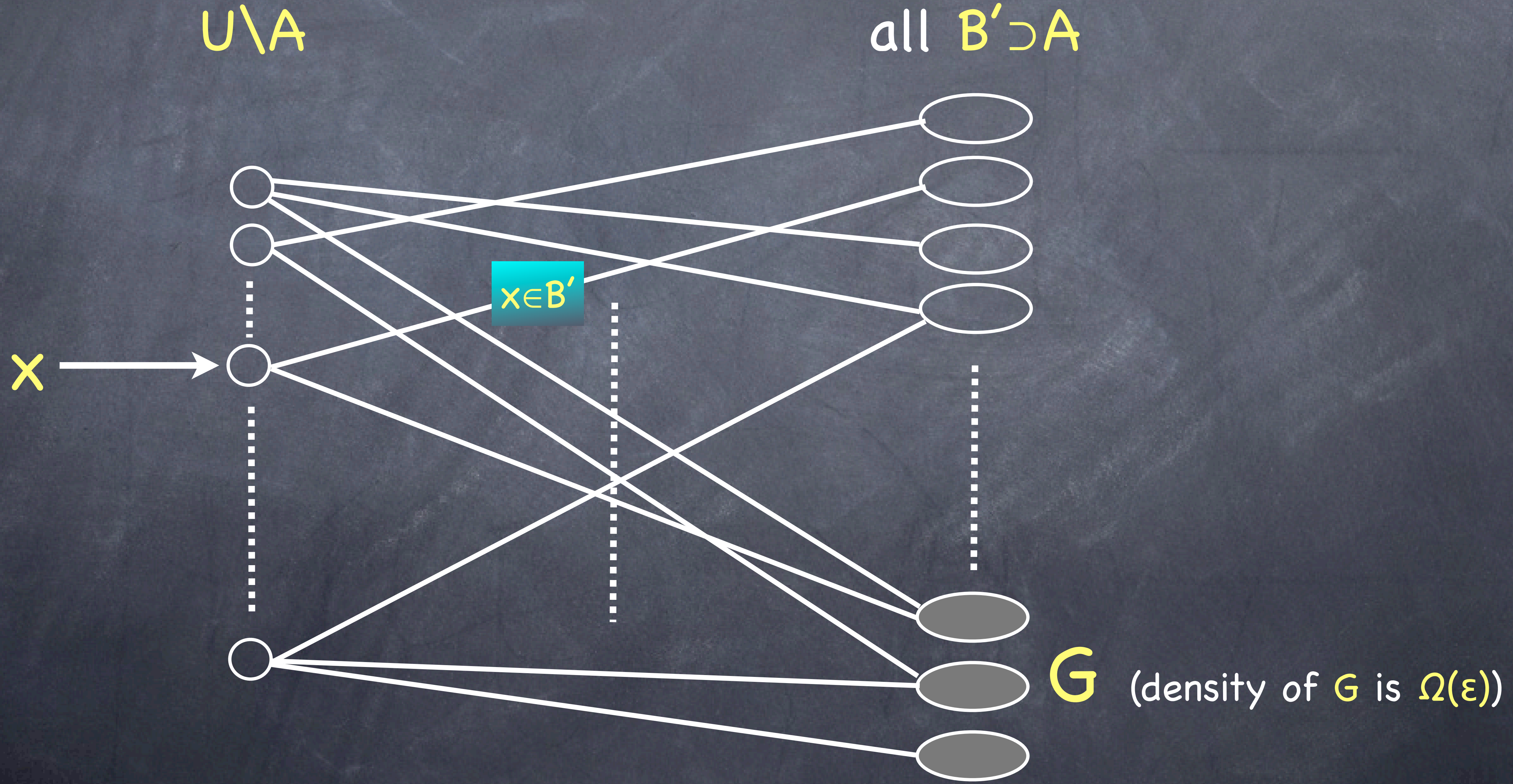
Analyzing $C_{A,B}$ given excellent (A,B)



Analyzing $C_{A,B}$ given excellent (A,B)

- Pr[$C_{A,B}$ fails] ≤ Pr[$C_{A,B}$ does not output an answer] + Pr[$C_{A,B}$ outputs an <u>incorrect</u> answer | $C_{A,B}$ outputs an answer]

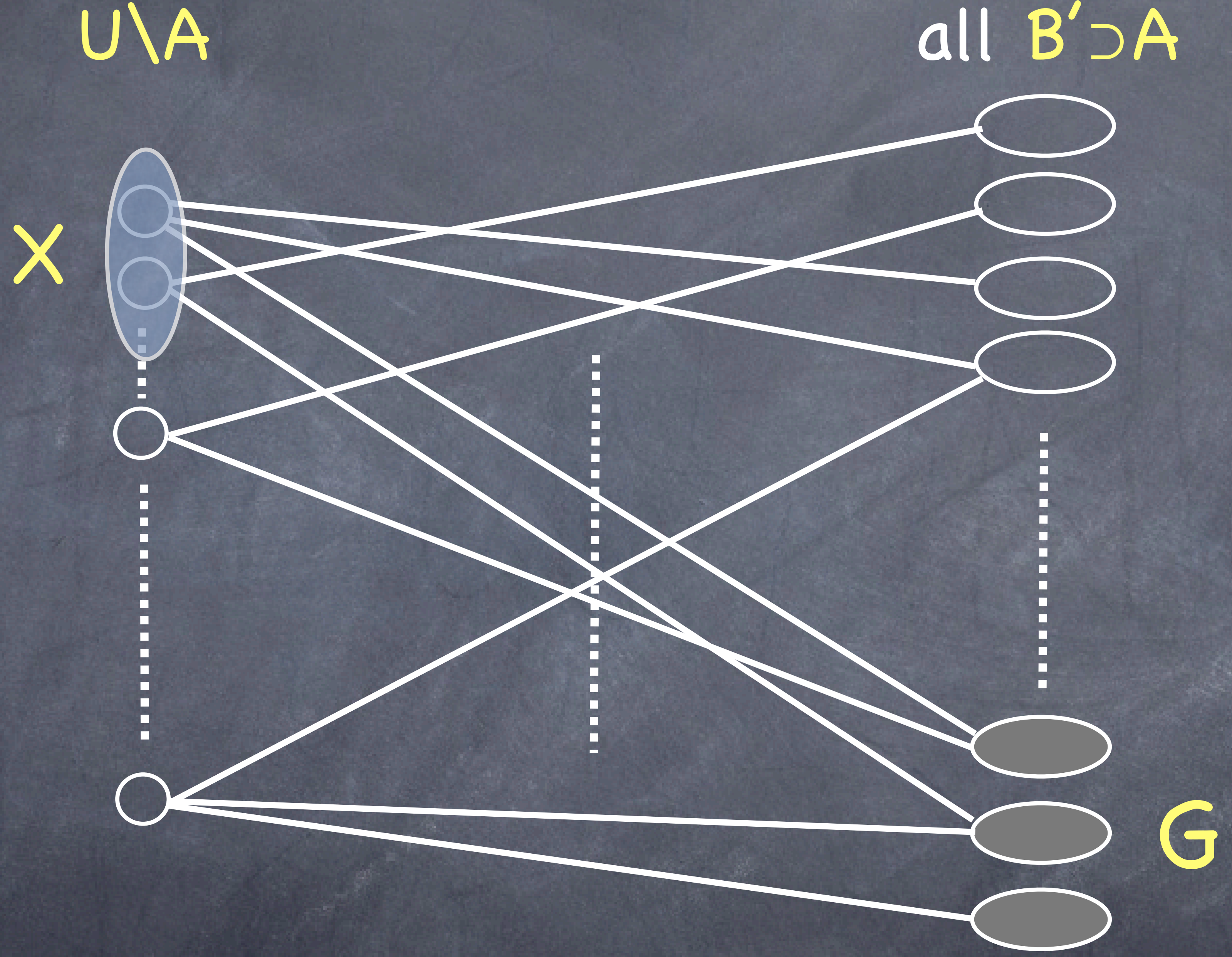
Analyzing $C_{A,B}$ given excellent (A,B)



Analyzing $C_{A,B}$ given excellent (A,B)

- Pr[$C_{A,B}$ fails] ≤ **Pr[$C_{A,B}$ does not output an answer]** + Pr[$C_{A,B}$ outputs an incorrect answer | $C_{A,B}$ outputs an answer]

Analyzing $C_{A,B}$ given excellent (A,B)



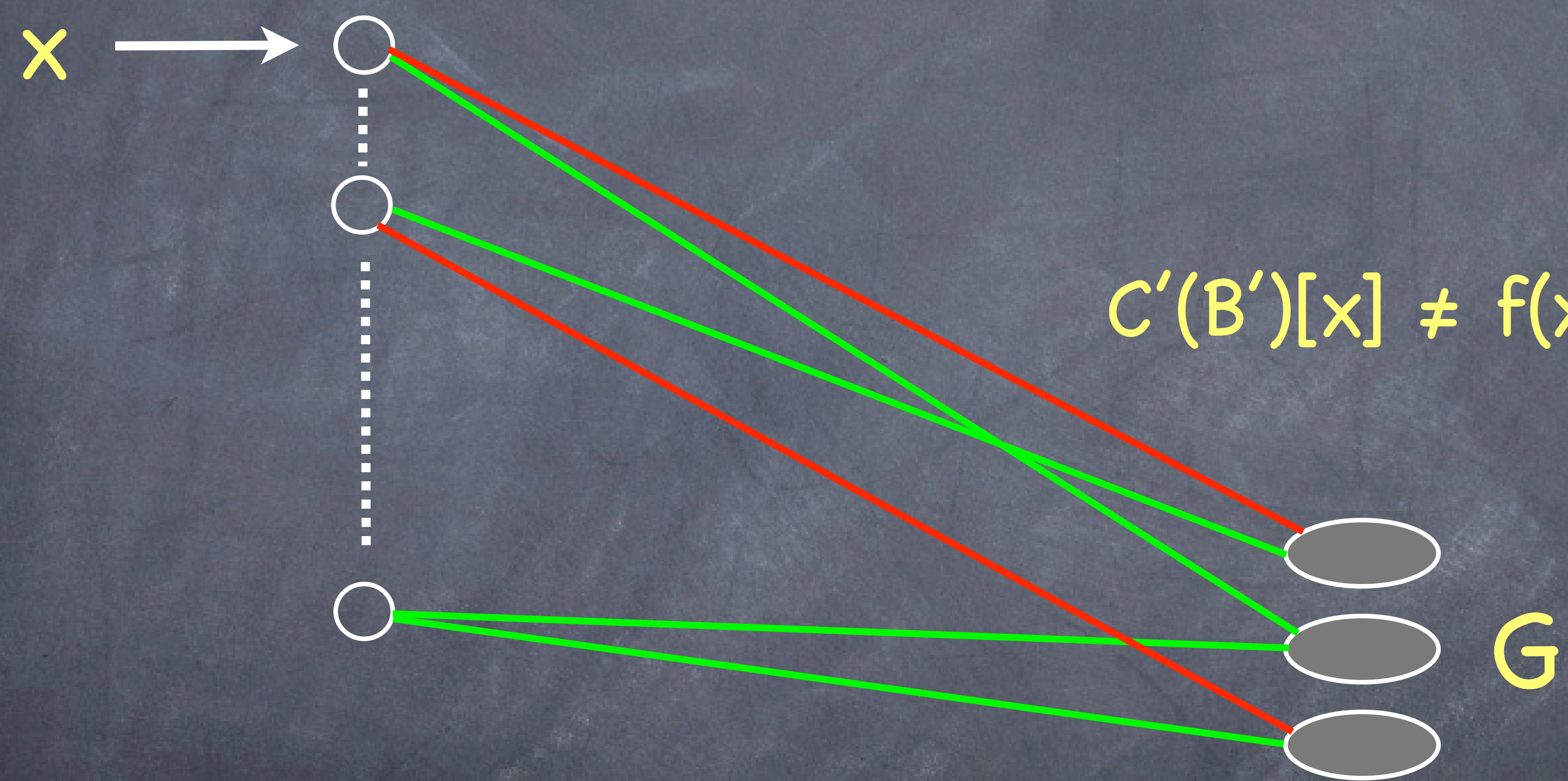
Sampler: For any $X \subset U \setminus A$ of density at least β almost all vertices in the right have at least $\beta/2$ fraction of edges into X.

Analyzing $C_{A,B}$ given excellent (A,B)

- Pr[$C_{A,B}$ fails] ≤ Pr[$C_{A,B}$ does not output an answer] + **Pr[$C_{A,B}$ outputs an incorrect answer | $C_{A,B}$ outputs an answer]**

Analyzing $C_{A,B}$ given excellent (A,B)

$R_z = \frac{\#\text{red incident edges}}{\text{degree}}$



Want to bound $E_x[R_x]$

We know that $E_y[R_y]$ is small

Following holds for Samplers: $E_x[R_x] \approx E_y[R_y]$

Derandomized DP Theorem

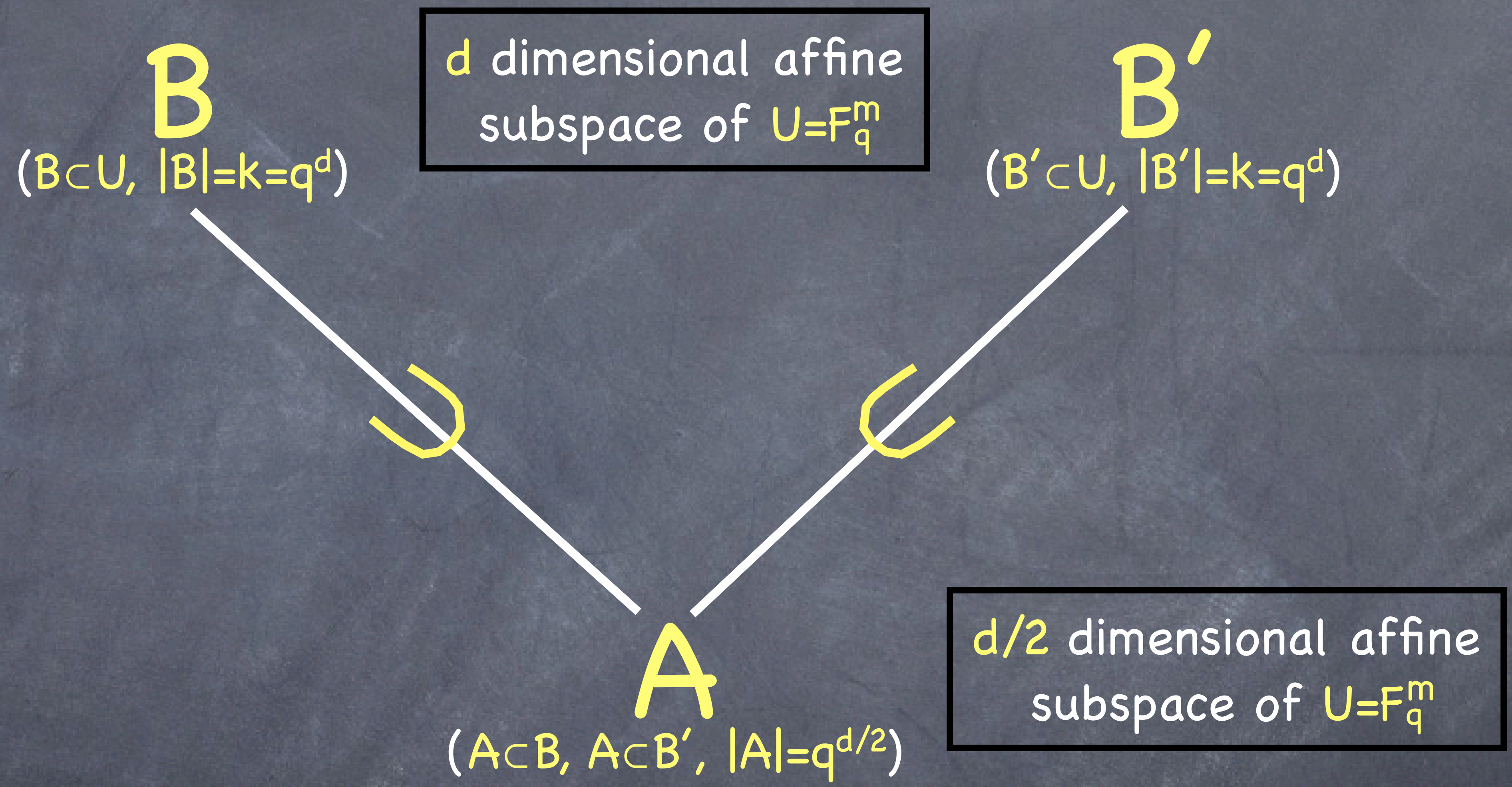
Derandomized DP Theorem

- DP Theorem: Given a hard $f:U\to R$, f^k is harder to compute on independently chosen subsets $B\subset U$, $|B|=k$
- Issue: The size of the inputs grows linearly with k

Derandomized DP Theorem

- Derandomized DP Theorem: Can we show that f^k is harder to compute on subsets $B \subset U$, $|B|=k$, even when these subsets have some limited independence
- [Imp95,IW97]: Derandomized DP Theorem in the nonuniform setting
- $U = F_q^m$, and consider f^k over low dimensional affine subspaces of U

Local Consistency Test

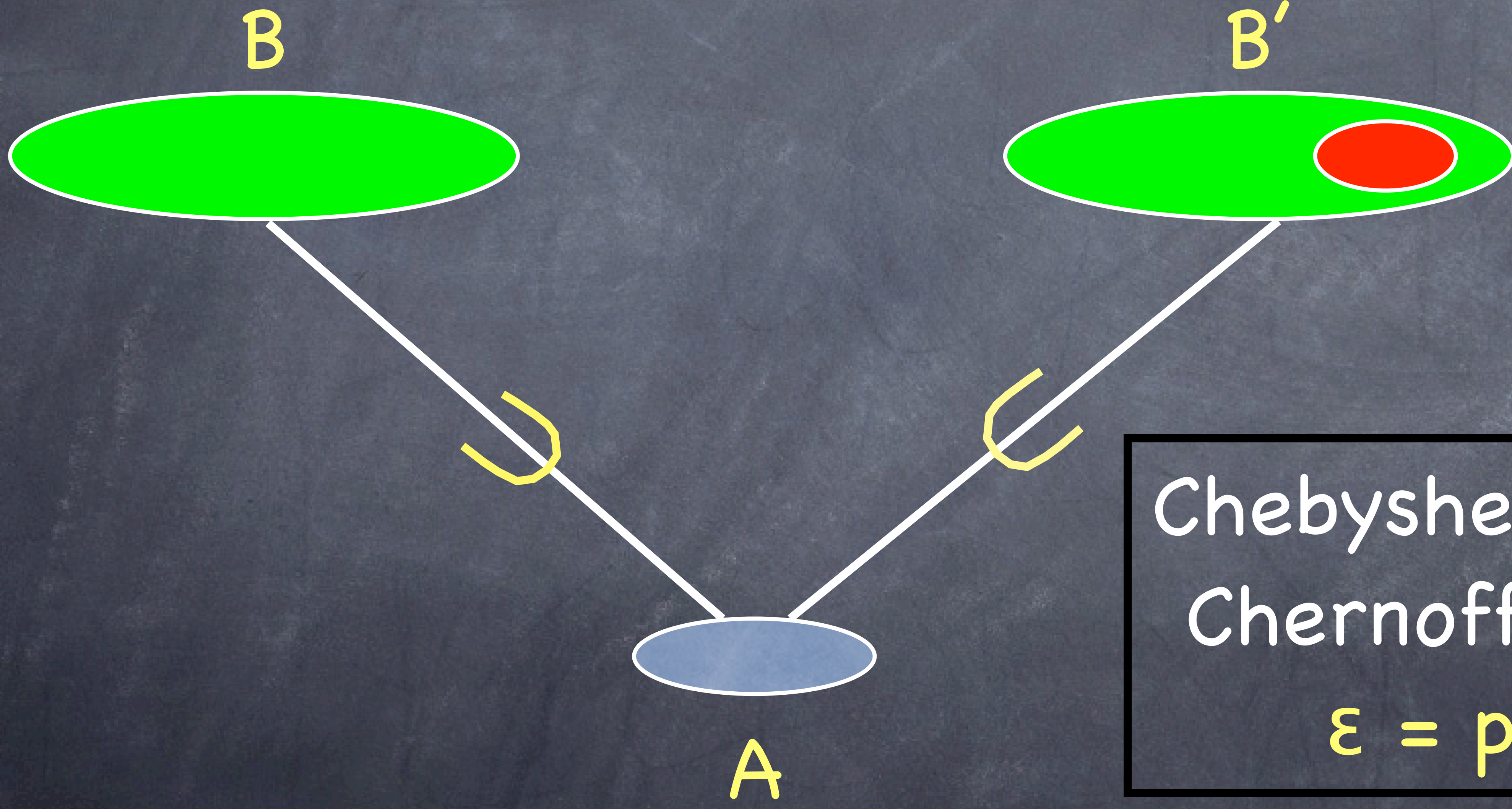


B′ is said to pass the consistency test wrt (A,B) if

$$C'(B)|_A = C'(B')|_A$$

Derandomized DP Theorem

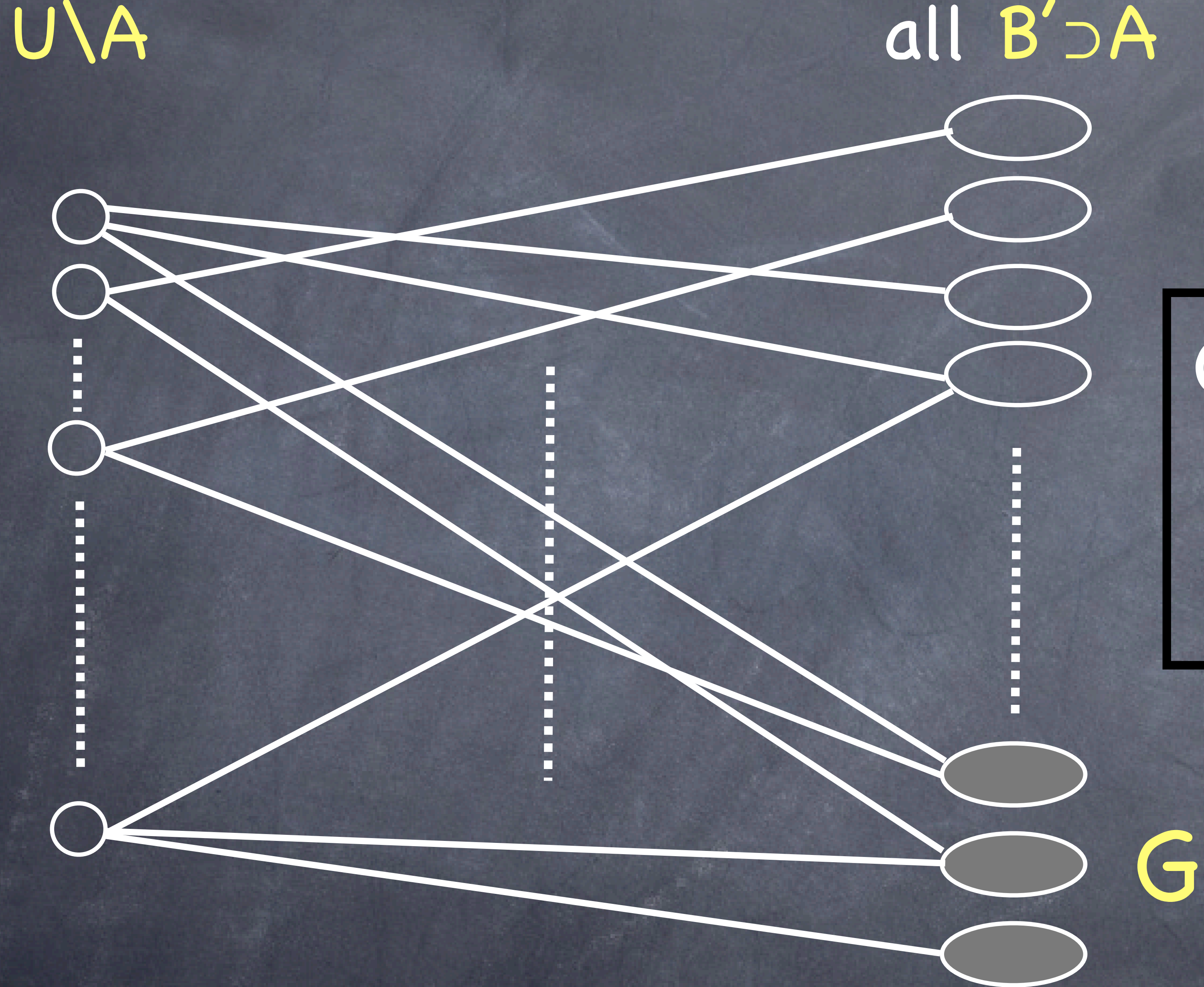
(Proof: $Pr_{A,B\supset A}[(A,B) \text{ is excellent}] = \Omega(\varepsilon)$)



Chebyshev instead of
Chernoff-Hoeffding
$\varepsilon = poly(1/k)$

w.h.p A contains a “bad” element of B′

Derandomized DP Theorem



Chebyshev instead of Chernoff-Hoeffding
ε = poly(1/k)

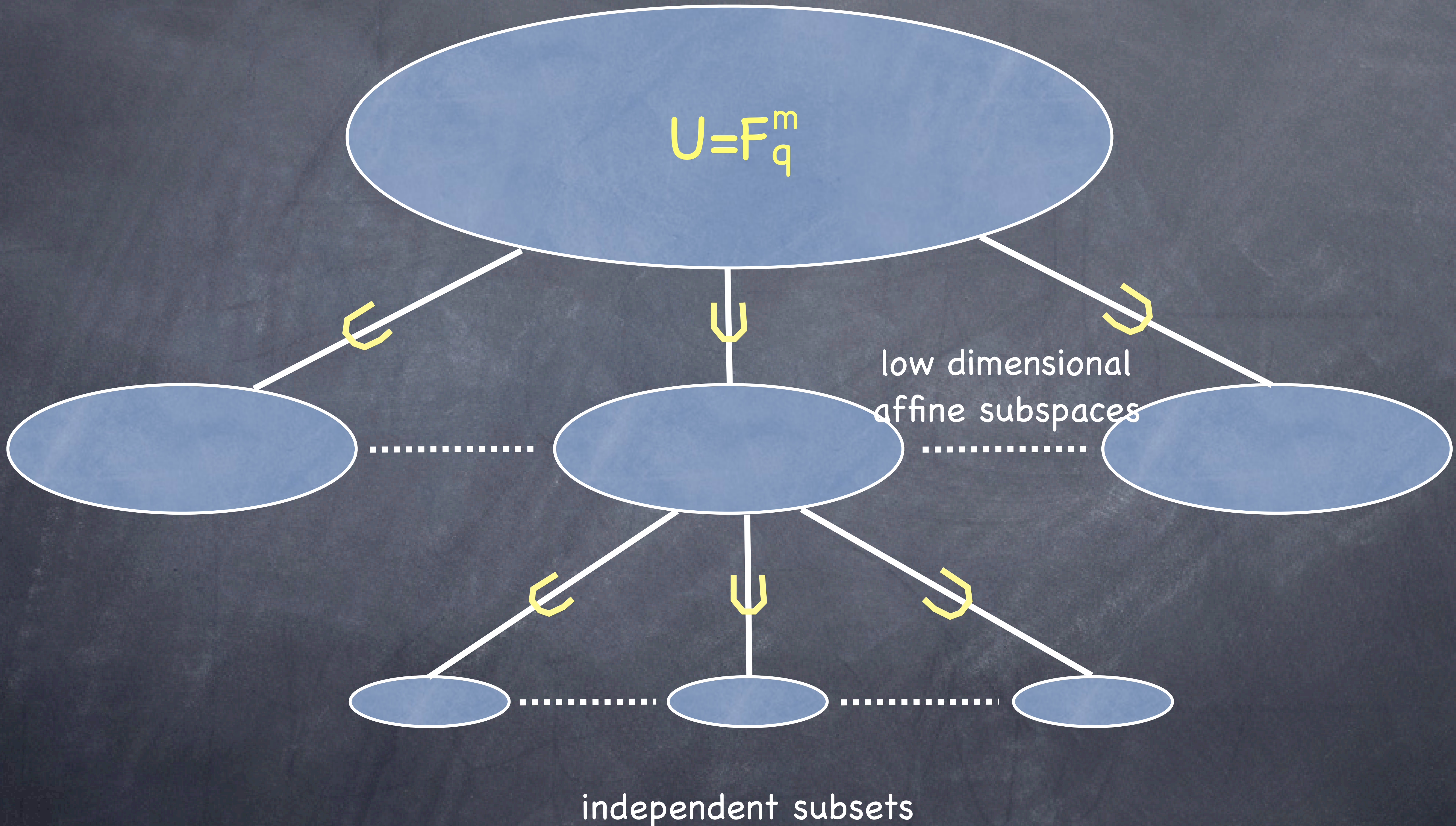
Sampler: For any $X \subset U \setminus A$ of density at least β almost all vertices in the right have at least $\beta/2$ fraction of edges into X.

Derandomized DP Theorem

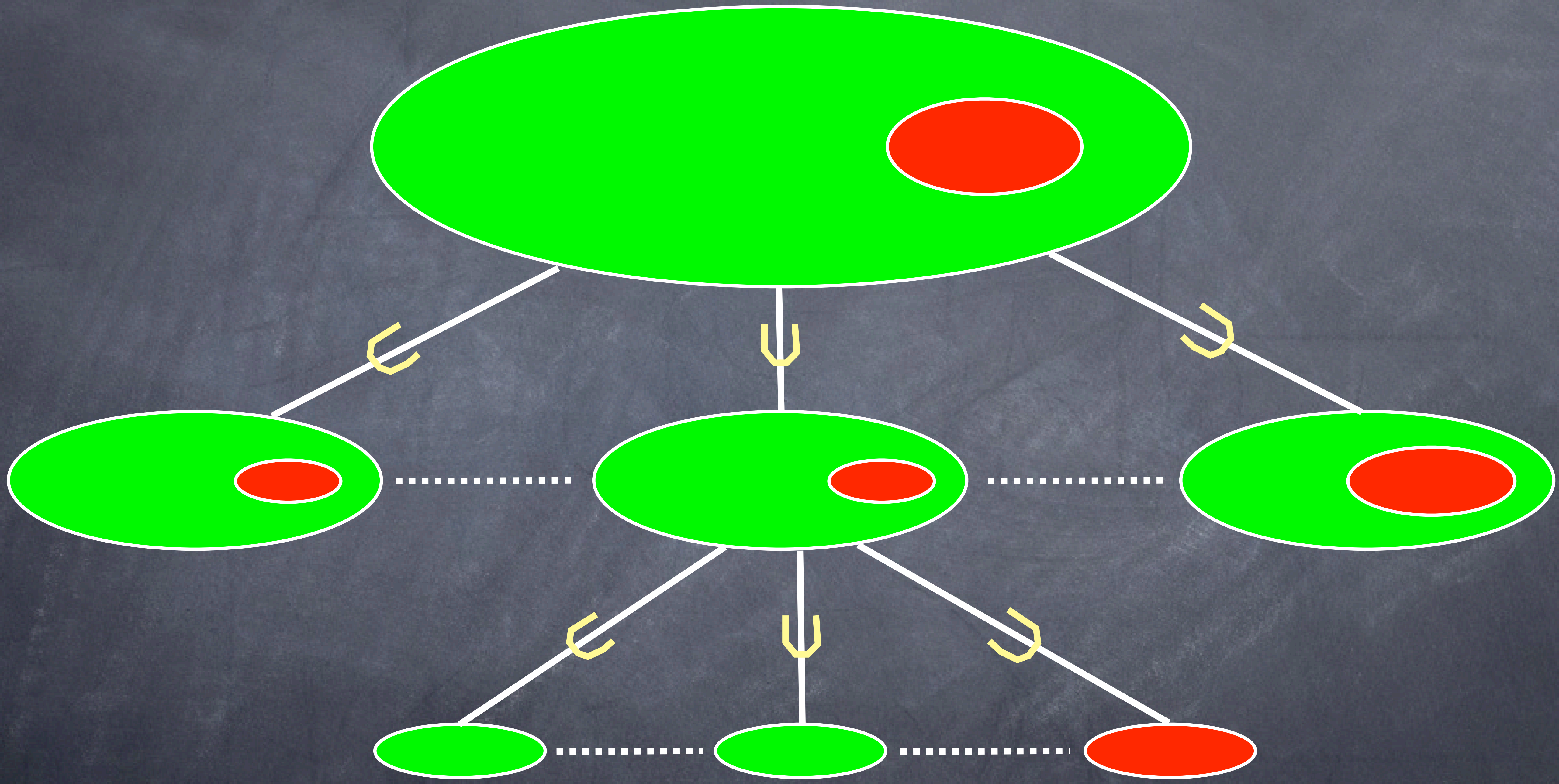
- Theorem [IJKW08]: Let $f:U\to R$ be some function and C' be a circuit such that $\Pr_{\text{affine subspace } B\subset U}[C' \text{ computes } f^k(B)] > \varepsilon$.
  There is an algorithm which outputs with probability $\Omega(\varepsilon)$ a circuit C such that $\Pr[C \text{ computes } f] > (1-\delta)$,
  where $\varepsilon=\text{poly}(1/k)$, $|C|=|C'|\cdot\text{poly}(1/\varepsilon,1/\delta,k)$.

- Note: description length of the input for f^k is $d\cdot\log(|U|)$

Derandomized DP Theorem



Derandomized DP Theorem



Approximate version of Derandomized DP Theorem

Derandomized DP Theorem

- Theorem [IJKW08]: Let $f:U\to R$ be some function and C' be a circuit such that $Pr_{\text{independent } B\subset T, \text{ low dim affine subspace } T\subset U}[C' \text{ computes } f^n(B)] > \varepsilon$. There is an algorithm which outputs with probability $poly(\varepsilon)$ a circuit C such that $Pr[C \text{ computes } f] > (1-\delta)$, where $\varepsilon=e^{-\Omega(\sqrt{n})}$, $|C|=|C'|\cdot poly(1/\varepsilon,1/\delta,n)$.
- Note: description length of input for f^k is $O(n)$ (given $\log(|U|=n)$)
- Open Problem: Bring down ε to $e^{-\Omega(n)}$

Open Problems

- Uniform "Chernoff-type" Direct Product Theorem in the spirit of [IJK07]
- Direct Product Testing
  - Given a circuit C as an oracle, using at most q queries to the oracle distinguish between the following two cases
    - C computes f^k for some f
    - C computes f^k on only some small ε fraction of inputs

Thank You