Approximating functions with DNFs

Eric Blais

Joint work with Li-Yang Tan

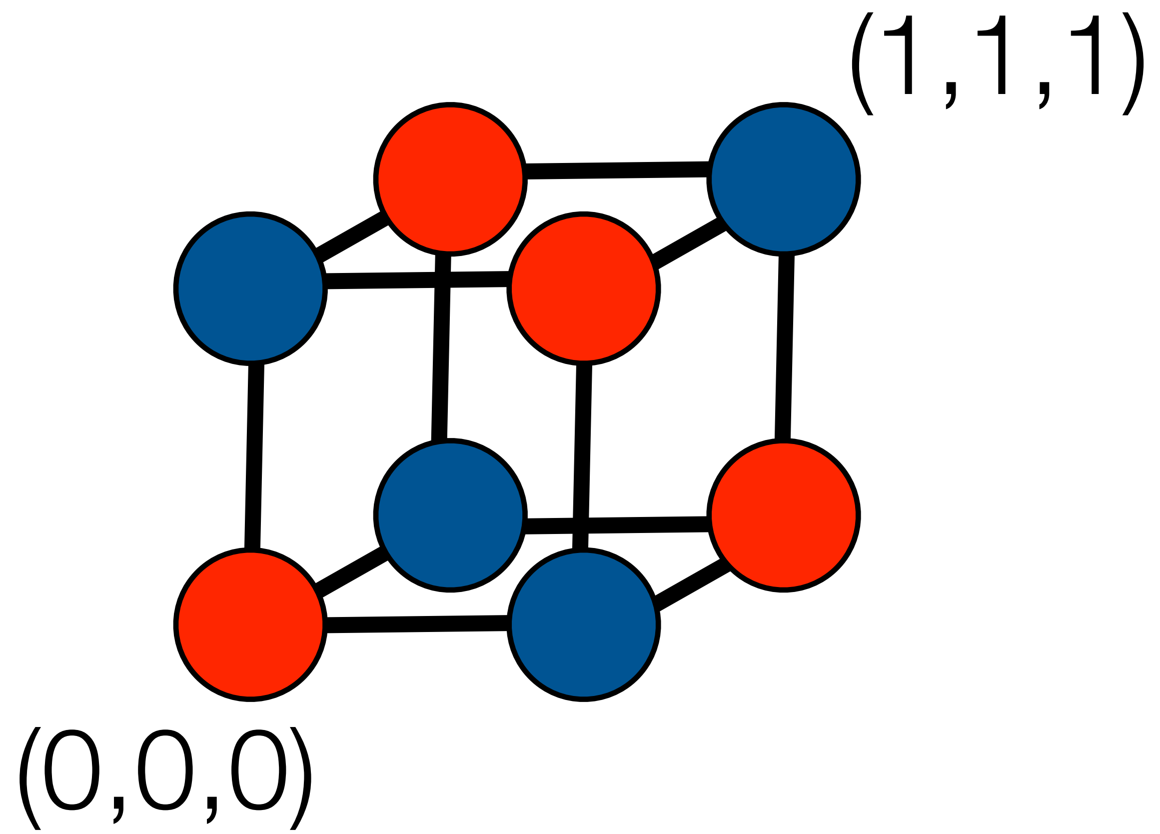
Boolean functions

$$f : \{0,1\}^n \rightarrow \{0,1\}$$

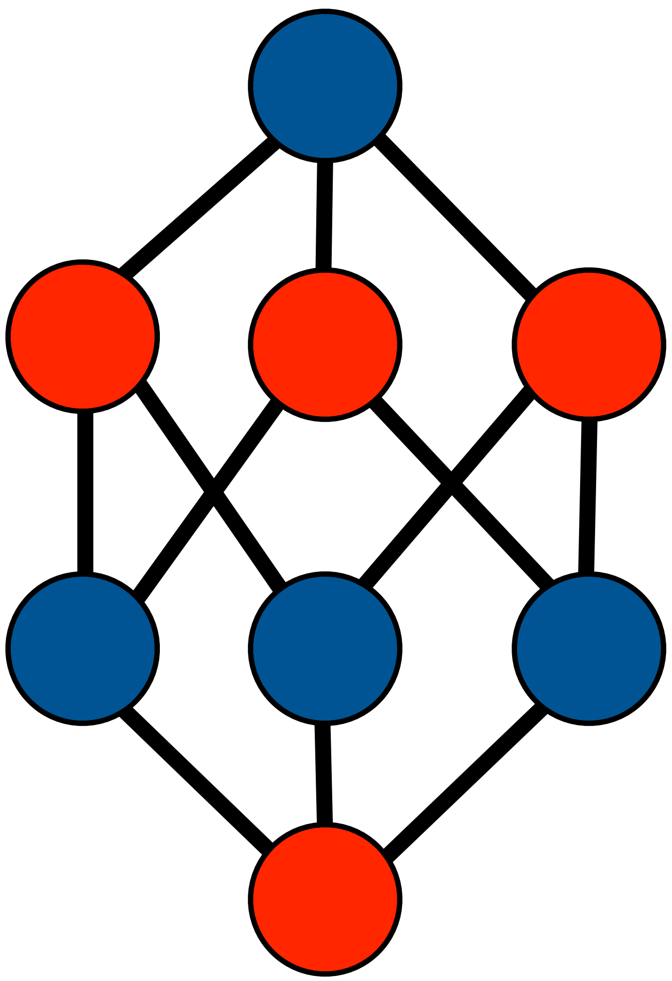
Boolean functions

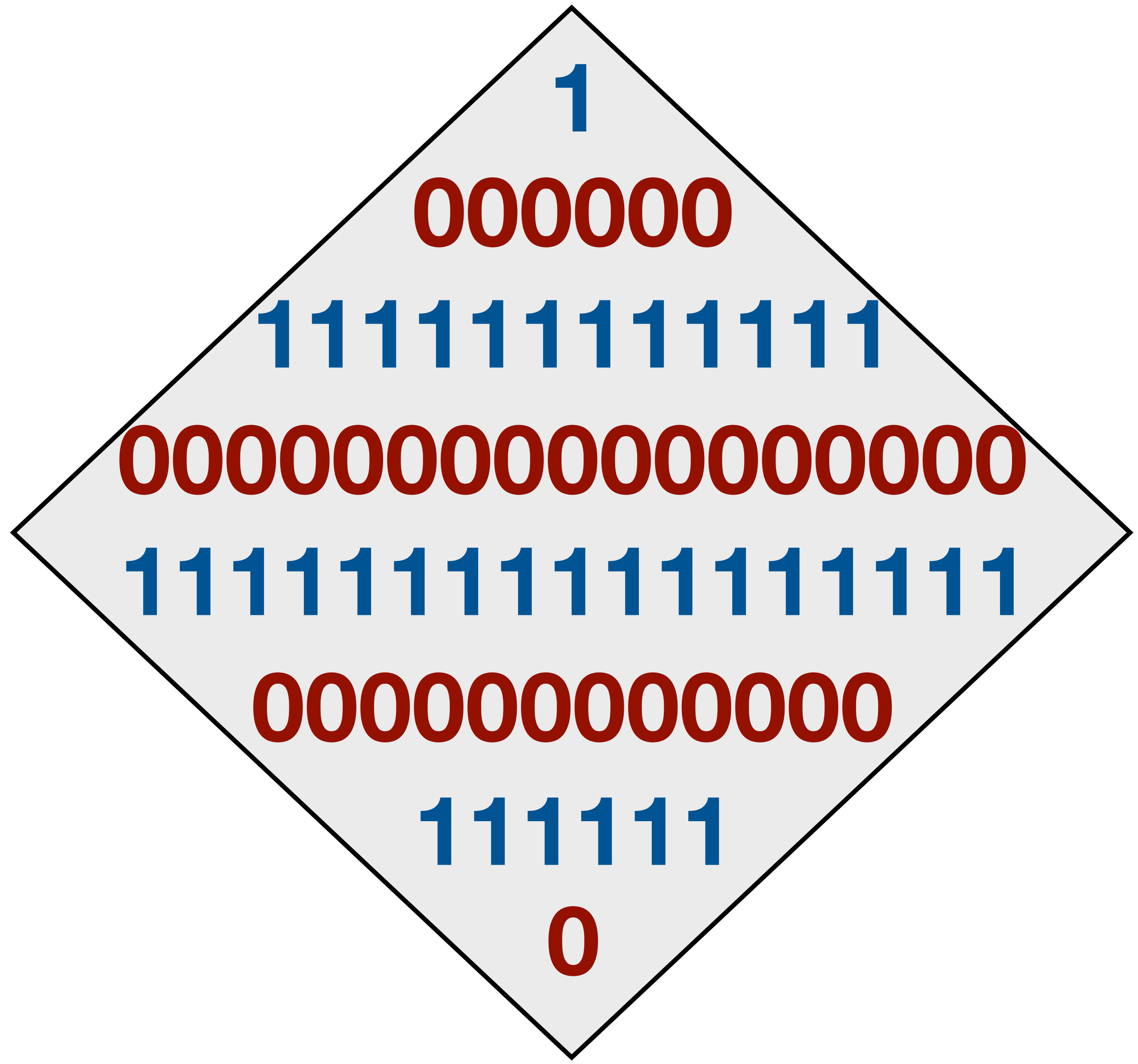
x_1	x_2	x_3	$f(x)$
0	0	0	**0**
0	0	1	**1**
0	1	0	**1**
0	1	1	**0**
1	0	0	**1**
1	0	1	**0**
1	1	0	**0**
1	1	1	**1**

Boolean functions

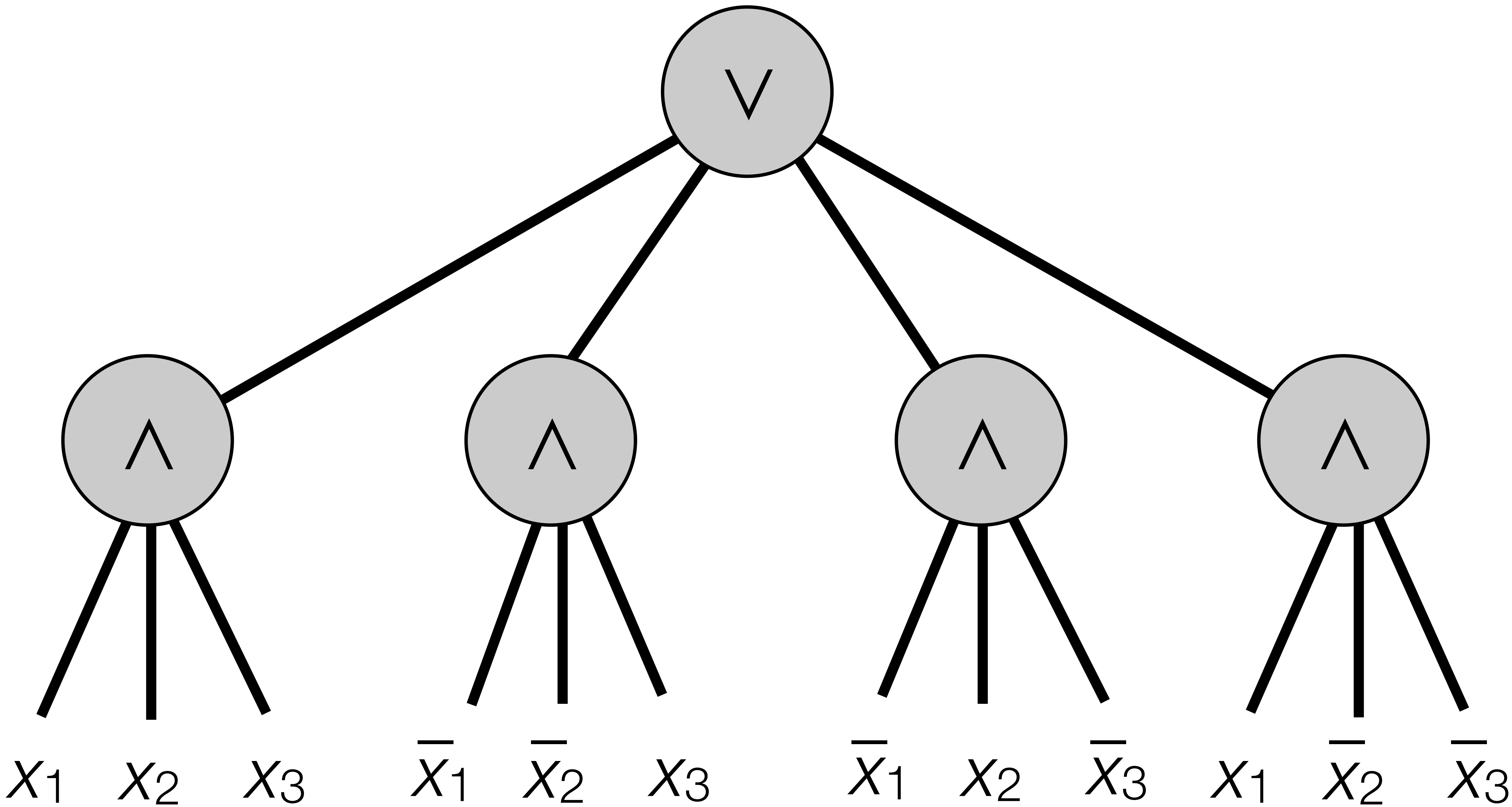


Boolean functions

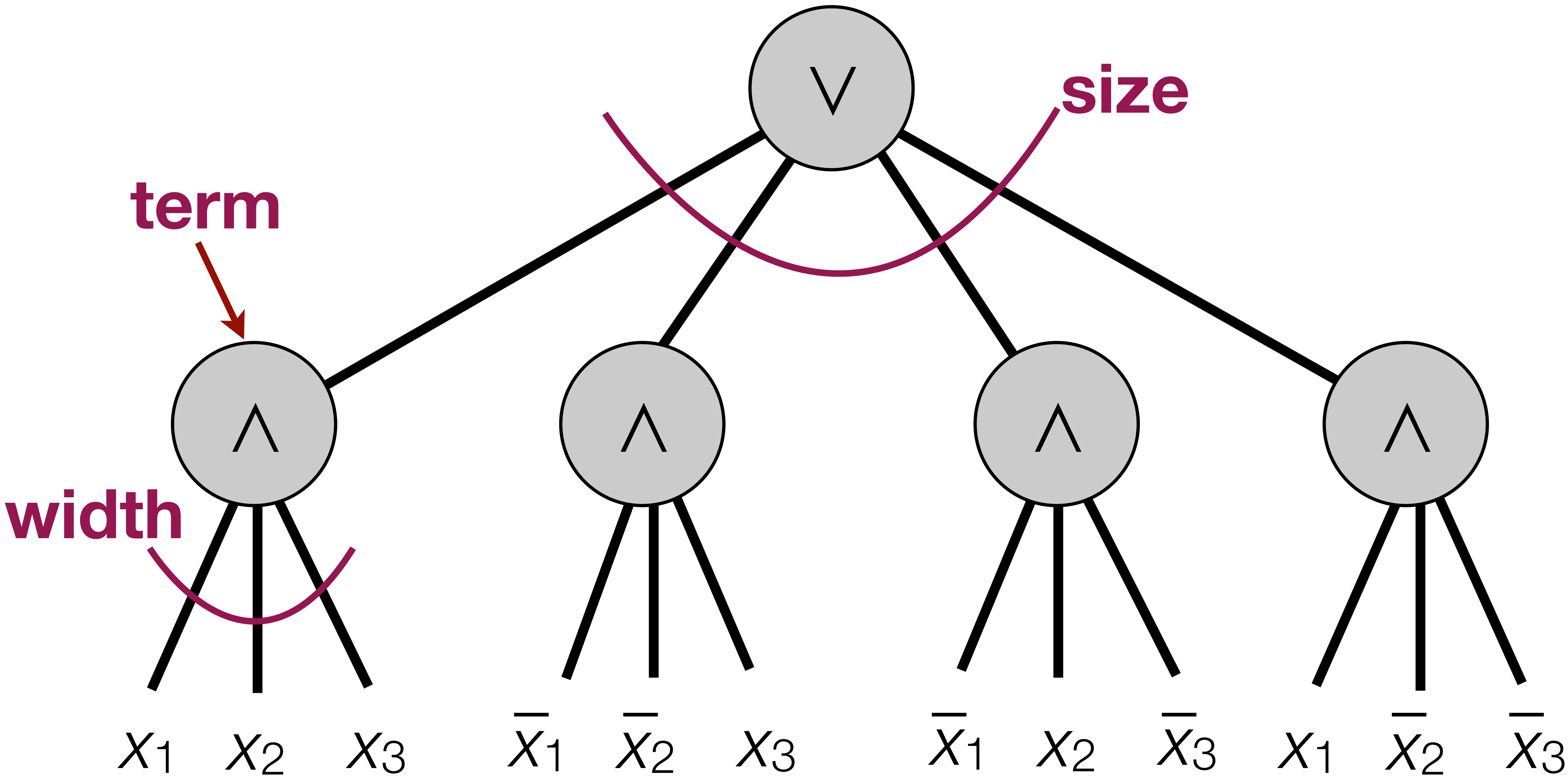




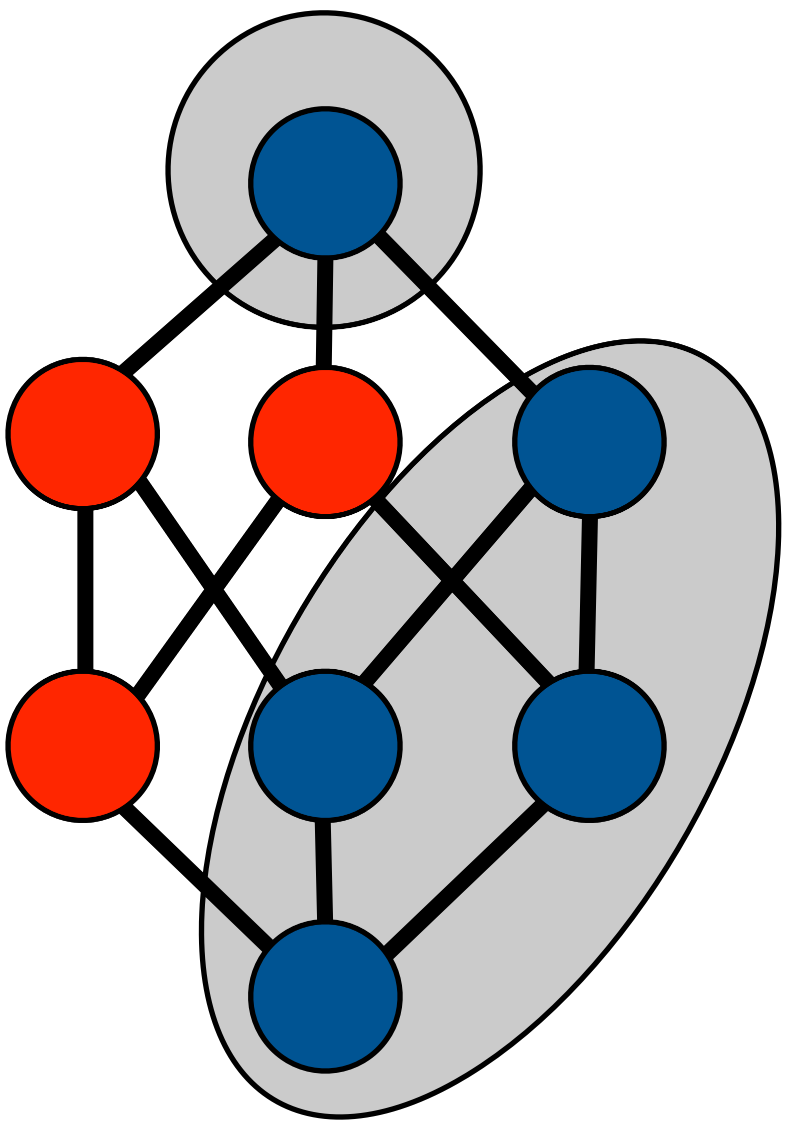
Boolean functions as DNFs

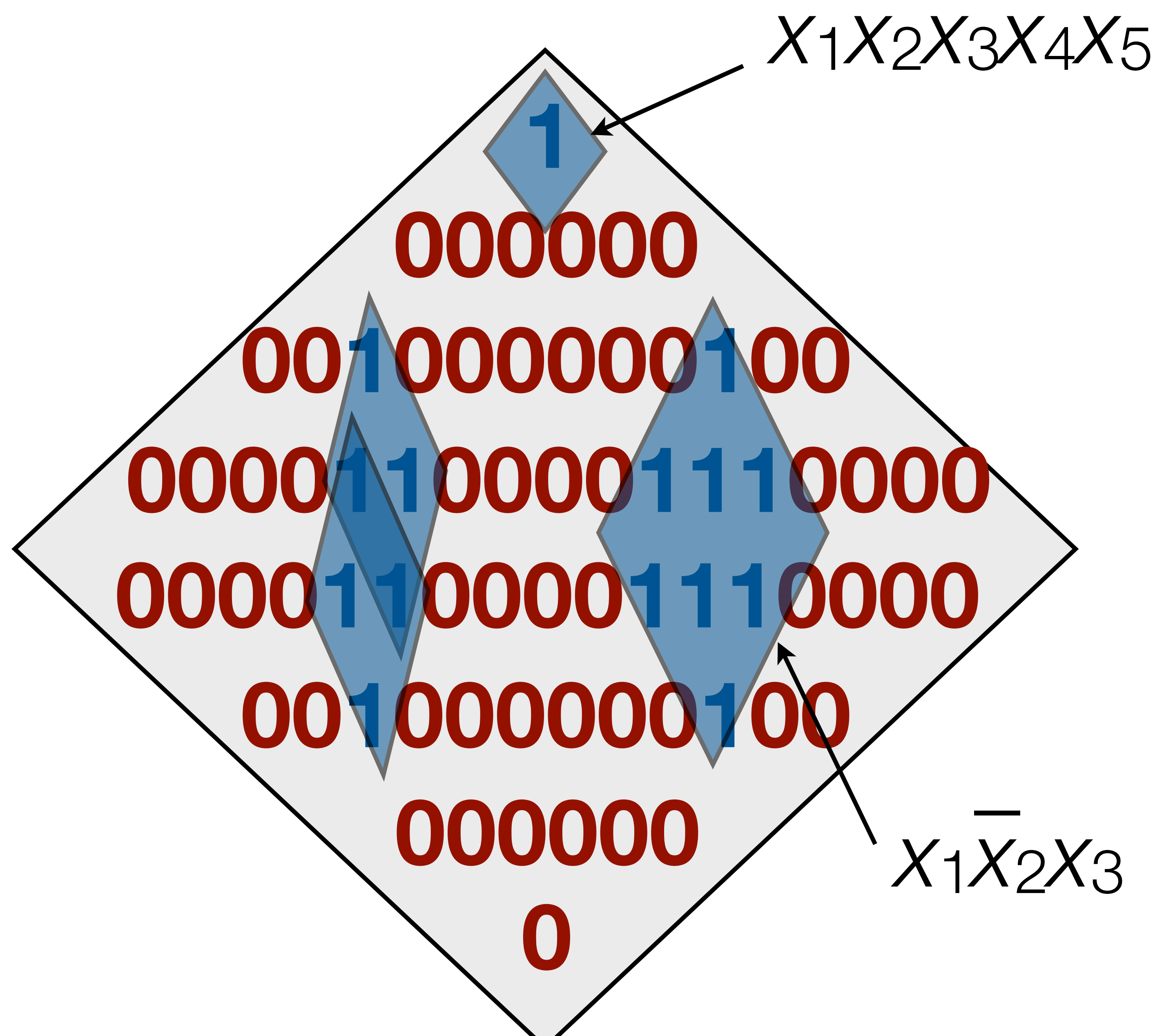


Boolean functions as DNFs



Boolean functions





DNFs are very well understood

Theorem (Lupanov '61). Every function f:$\{0,1\}^n \rightarrow \{0,1\}$ can be computed by a DNF of size 2^n and width n.

Proof.

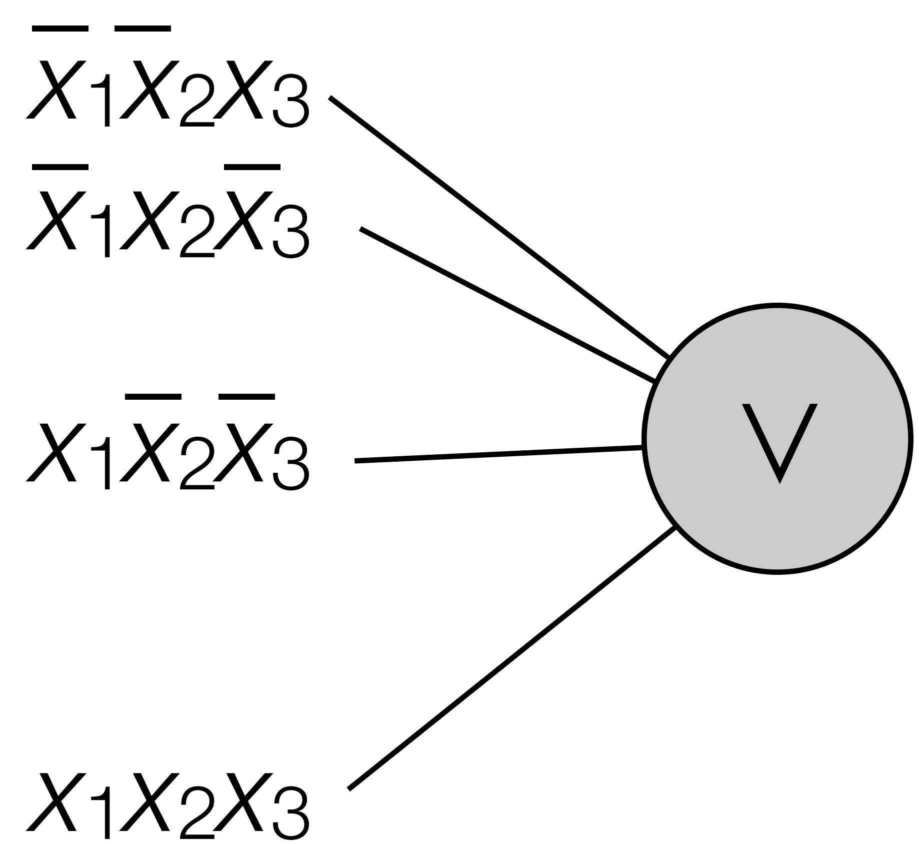
x_1	x_2	x_3	$f(x)$
0	0	0	**0**
0	0	1	**1**
0	1	0	**1**
0	1	1	**0**
1	0	0	**1**
1	0	1	**0**
1	1	0	**0**
1	1	1	**1**

DNFs are very well understood

Theorem (Lupanov '61). Every function f:$\{0,1\}^n \rightarrow \{0,1\}$ can be computed by a DNF of size 2^n and width n.

Proof.

x_1	x_2	x_3	$f(x)$
0	0	0	**0**
0	0	1	**1**
0	1	0	**1**
0	1	1	**0**
1	0	0	**1**
1	0	1	**0**
1	1	0	**0**
1	1	1	**1**

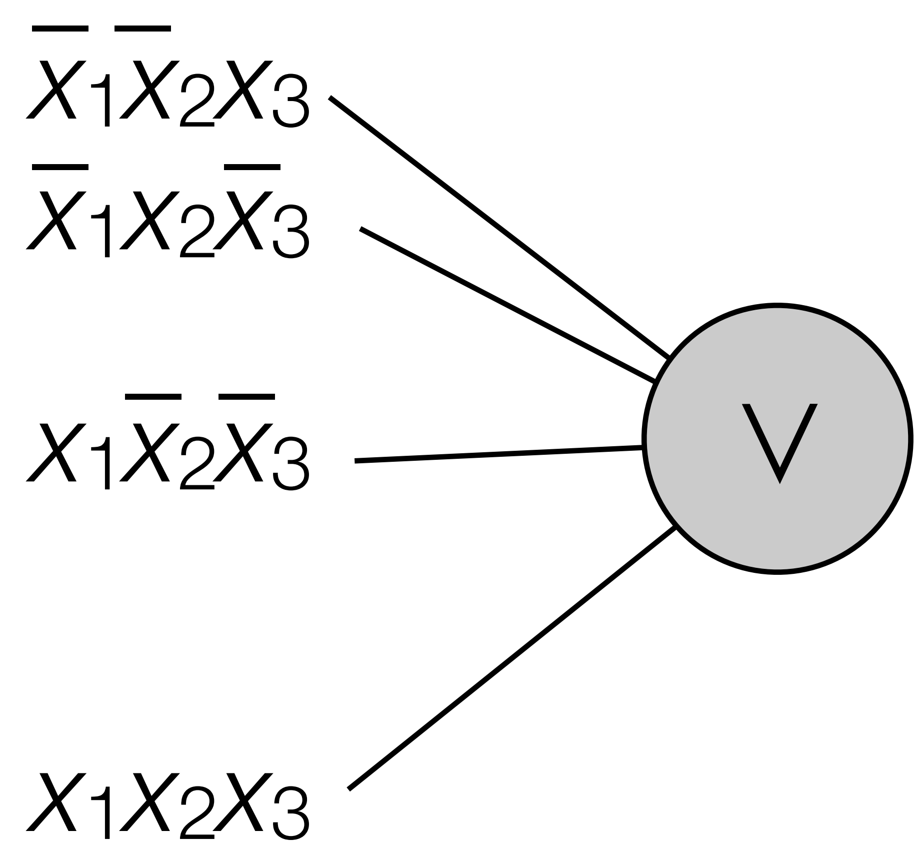


DNFs are very well understood

Theorem (Lupanov '61). Every function f:$\{0,1\}^n \rightarrow \{0,1\}$ can be computed by a DNF of size $2^{n\text{-}1}$ and width n.

Proof.

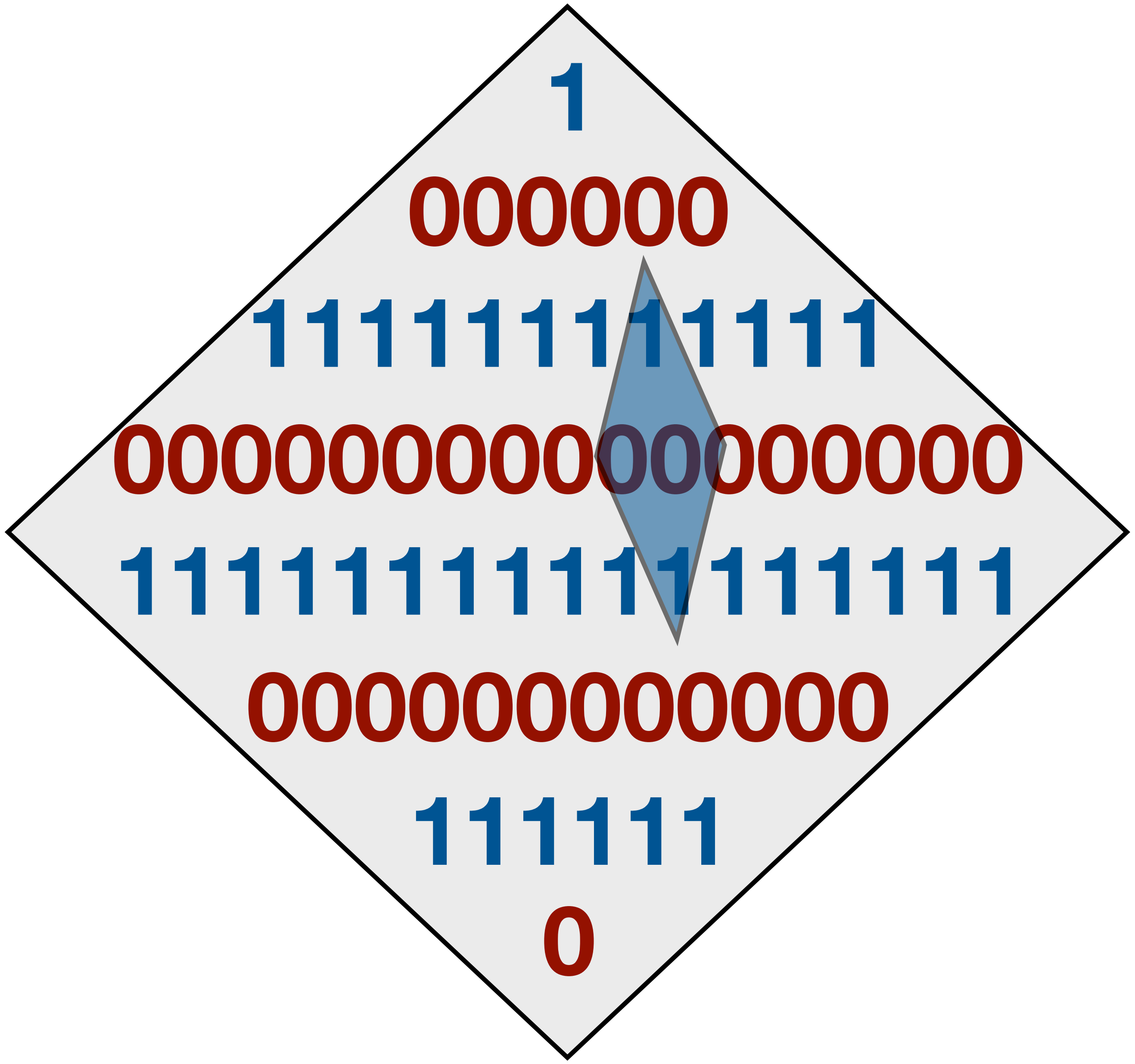
x_1	x_2	x_3	$f(x)$
0	0	0	**0**
0	0	1	**1**
0	1	0	**1**
0	1	1	**0**
1	0	0	**1**
1	0	1	**0**
1	1	0	**0**
1	1	1	**1**



Parity function

Theorem (Lupanov '61). Every DNF computing the function $f = x_1 \oplus x_2 \oplus \ldots \oplus x_n$ has size 2^{n-1} and width n.

Proof.



Random function

Theorem (Korshunov-Kuznetsov ’83). For almost every function f:$\{0,1\}^n \rightarrow \{0,1\}$, every DNF computing f has size $\Theta(2^n/\log n \log \log n)$ and width n - $\log(3n)$.

Approximating functions

Main Question. What if we only require our DNF to compute f correctly on *most* inputs?

Def'n. The functions $f,g : \{0,1\}^n \to \{0,1\}$ are ε-close if

$$|\{x \in \{0,1\}^n : f(x) \neq g(x)\}| \leq \varepsilon 2^n.$$

Def'n. A DNF ε-approximates $f : \{0,1\}^n \to \{0,1\}$ if the function it computes is ε-close to f.

Approximating Parity

Fact. Every function can be .01-approximated by a DNF of size $0.99*2^{n-1}$ and width n.

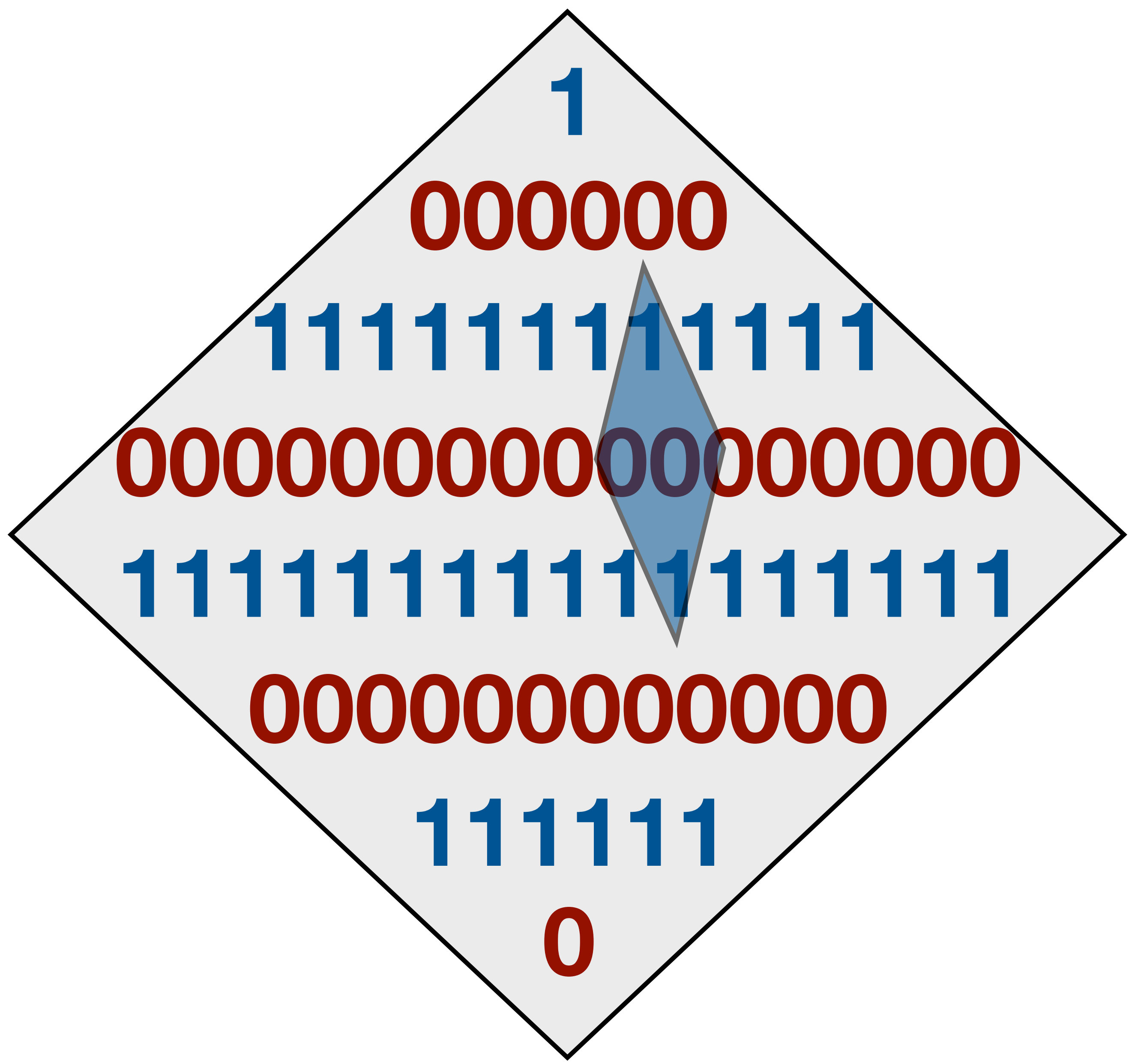
Theorem (Boppana-Hastad '97). Every DNF that .01-approximates Parity has size at least $2^{n/16}$ and width at least $n/16$.

How tight are these bounds?

Approximating parity

(Bold) Conjecture. Every DNF that .01-approximates Parity has size at least $\Omega(2^n)$ and width at least n - O(1).

Intuition.



1. Upper bound on DNF size

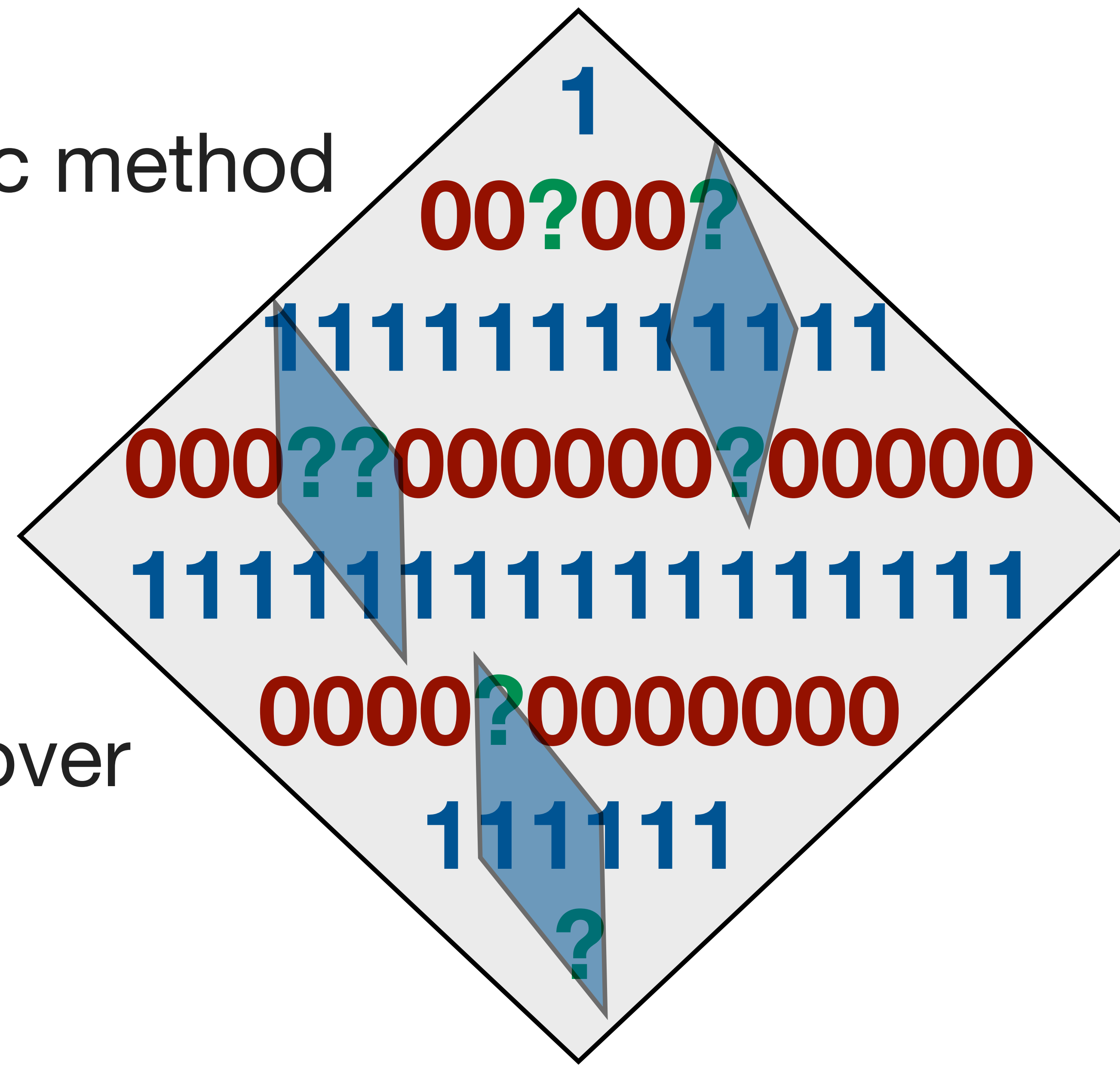
Approximating parity

Theorem. There is a DNF of size $O(2^n/\log n)$ that .01-approximates the parity function.

Proof strategy: Probabilistic method

1. Flip each **0** to **?** with probability .01.

2. Add all the subcubes of dimension log log *n* that cover only **1** and **?**.



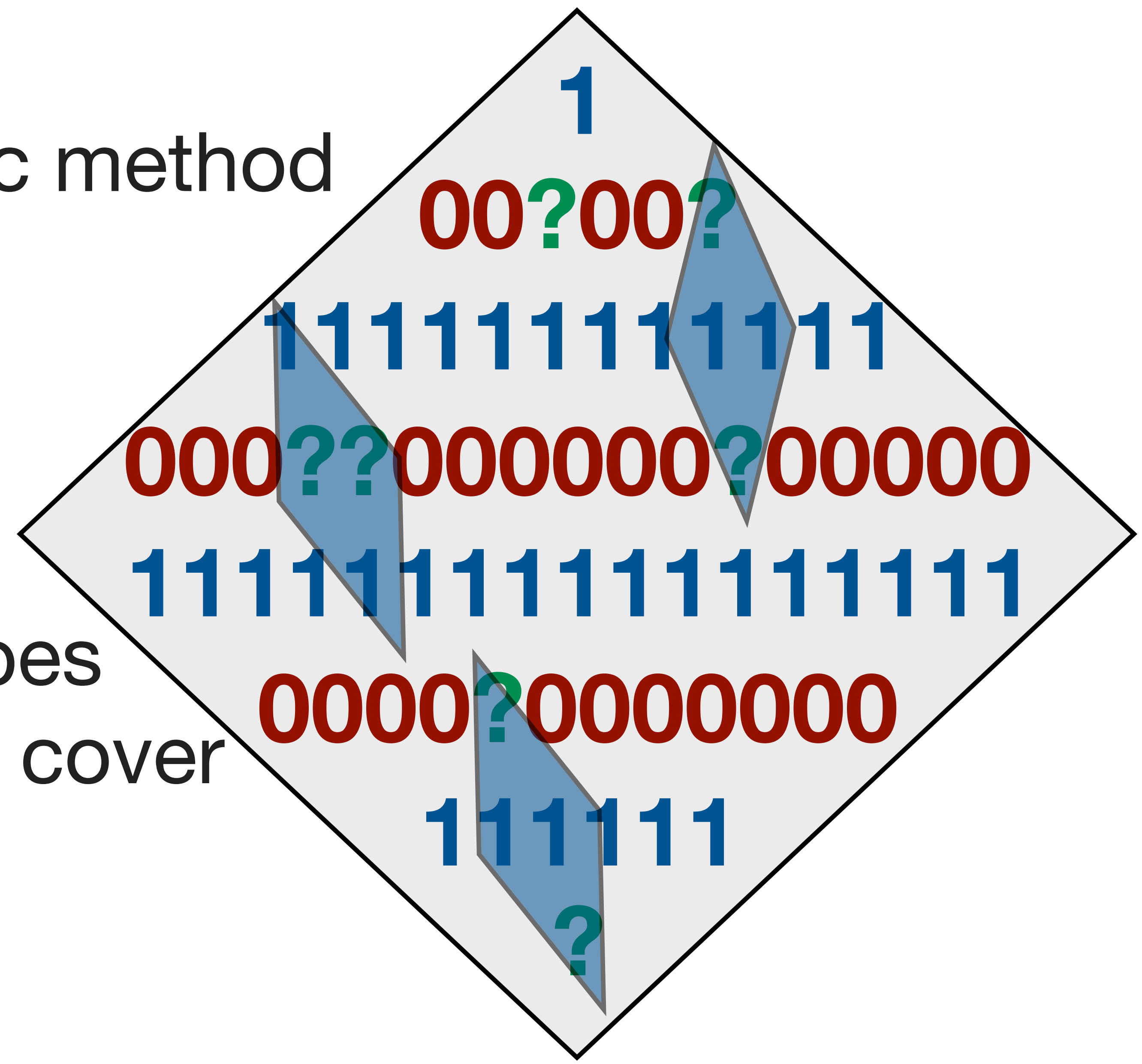
Approximating parity

Theorem. There is a DNF of size O(2^n/log n) that .01-approximates the parity function.

Proof strategy: Probabilistic method

1. Flip each **0** to **?** with probability .01.

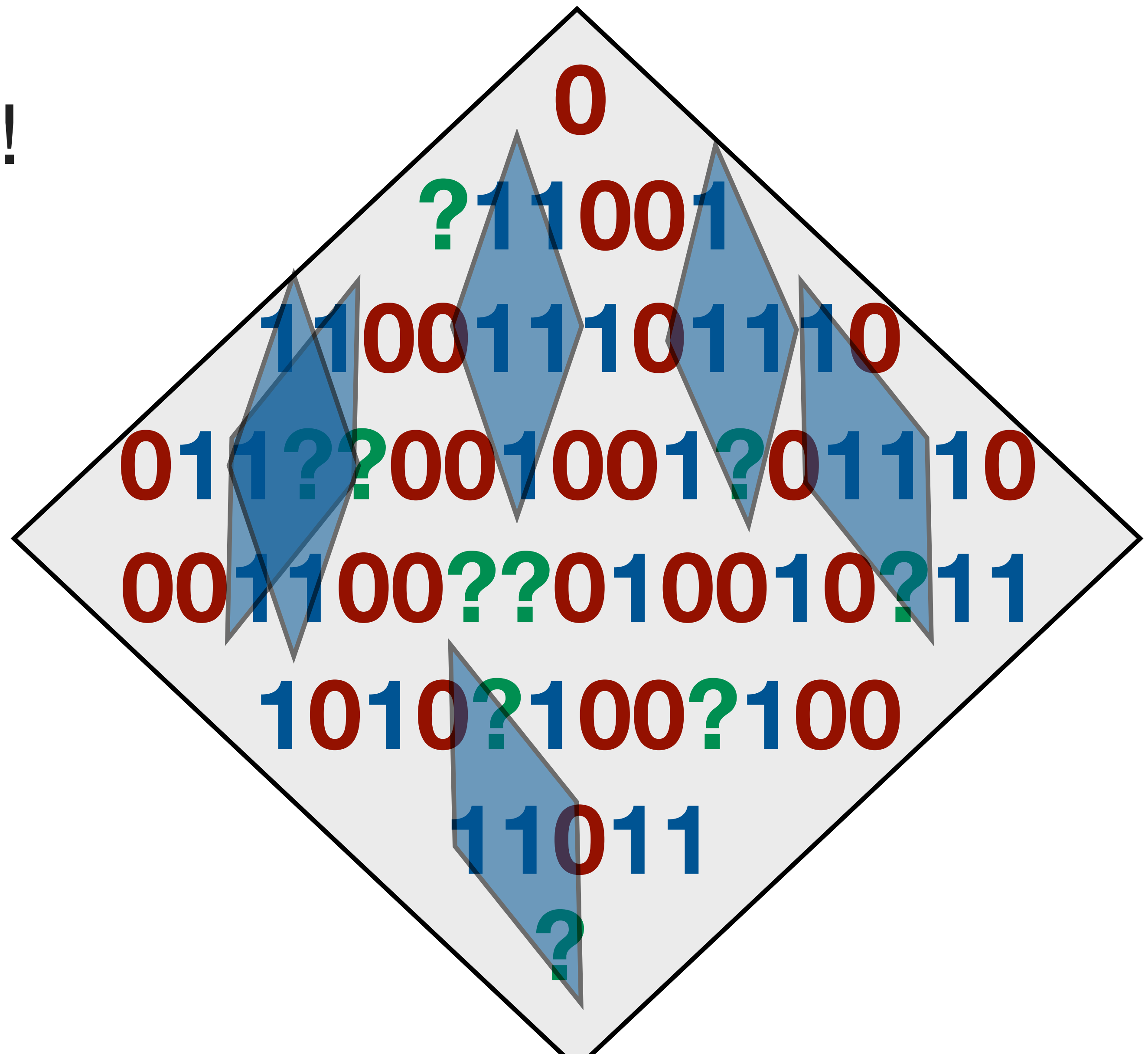
2. Add some of the subcubes of dimension log log n that cover only **1** and **?**.



Approximating any function

Theorem. For every function f:$\{0,1\}^n \to \{0,1\}$, there is a DNF of size $O(2^n/\log n)$ that .01-approximates f.

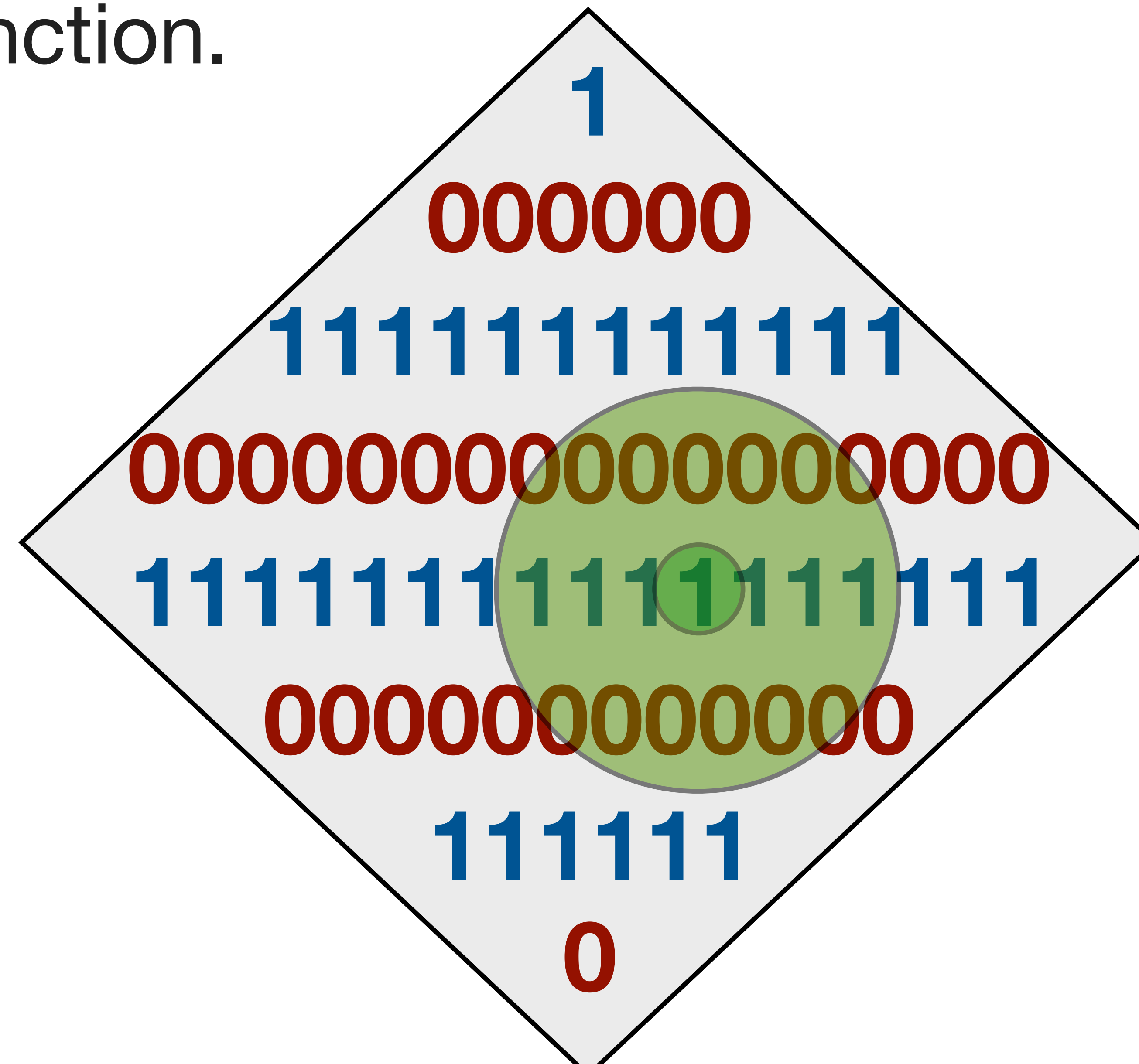
Proof. Same!



2. Upper bound on DNF width

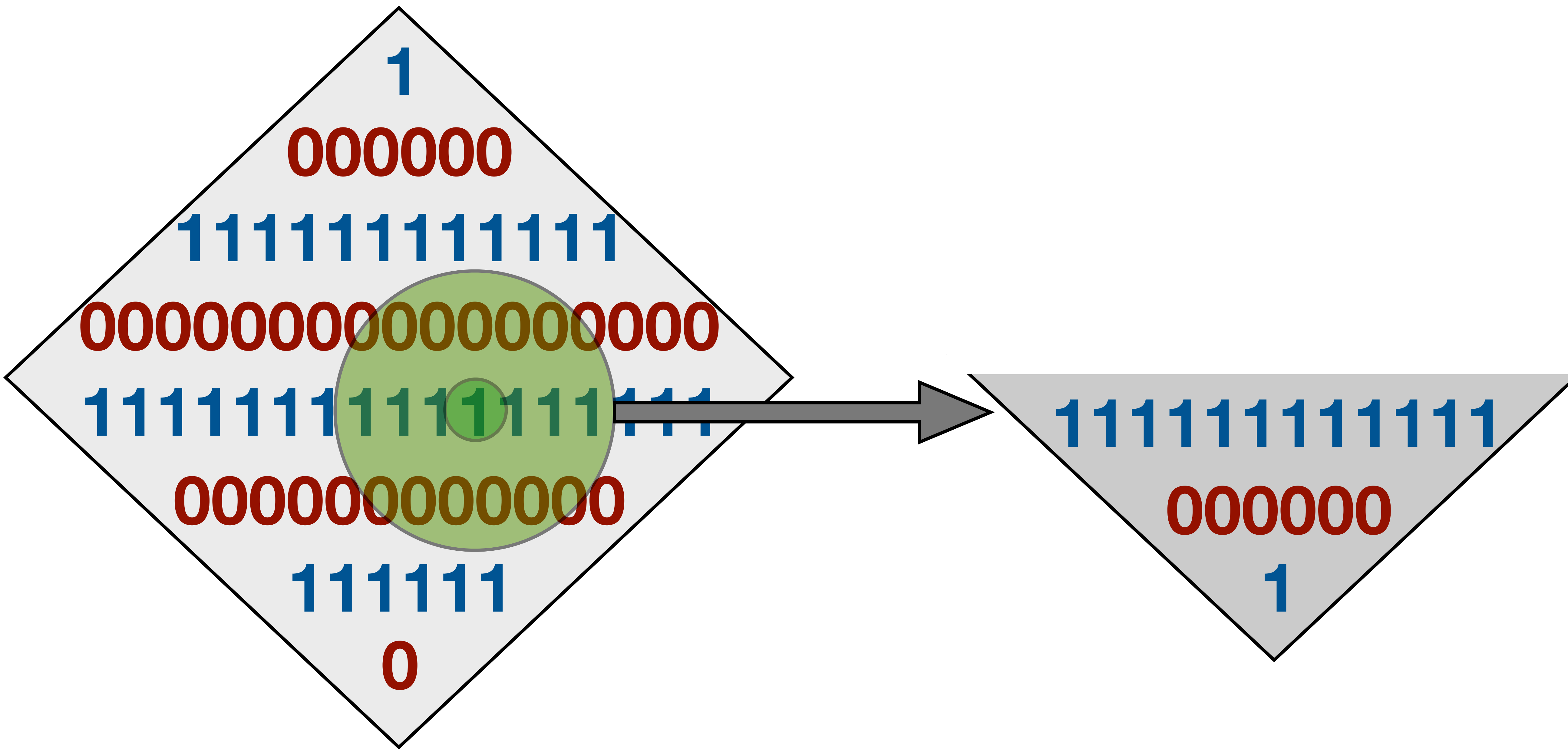
Approximating parity with small width

Theorem. There is a DNF of width n-Ω(n) that .01-approximates the parity function.

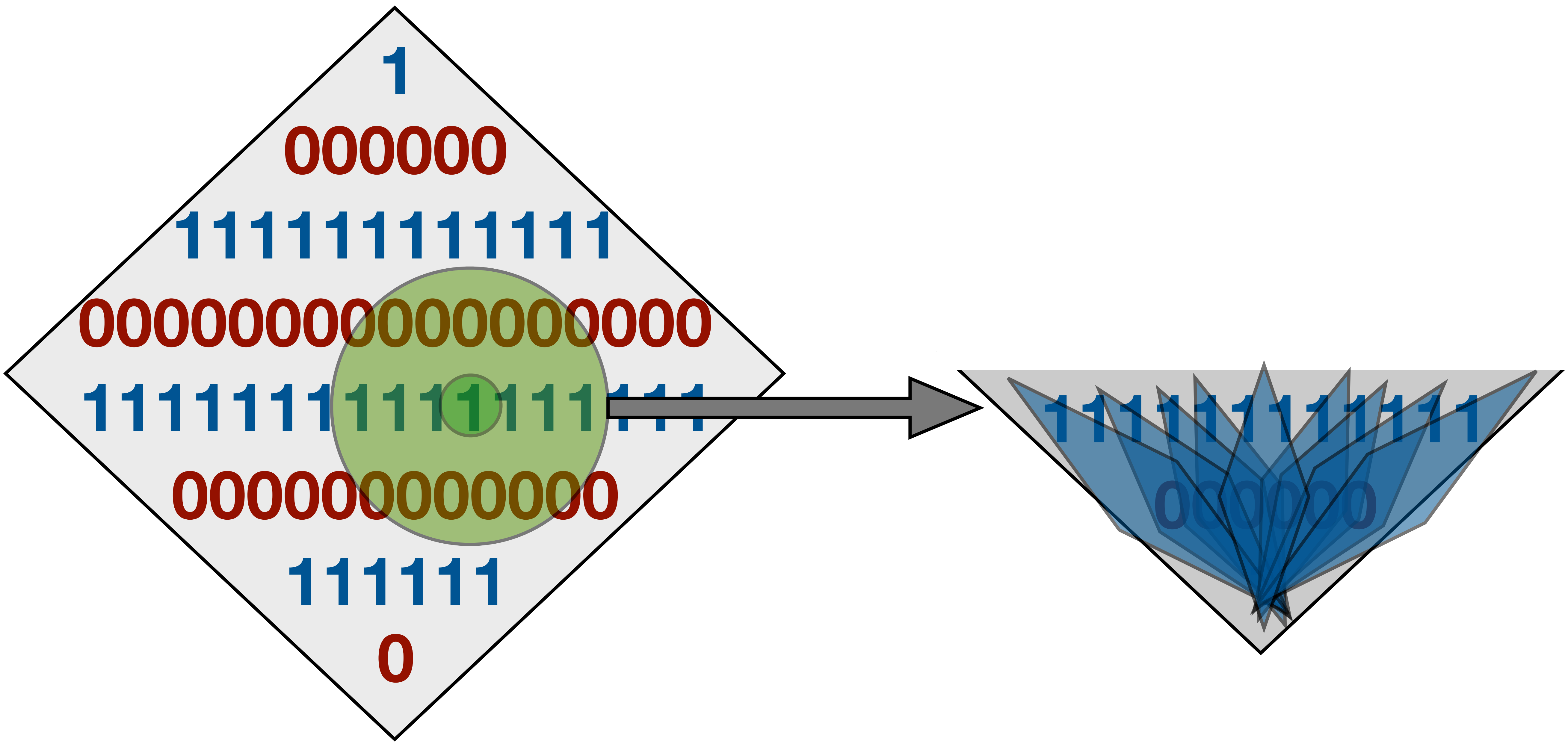


Proof strategy.
Design DNFs that
1. Approximate parity well on a fixed Hamming ball.
2. Evaluate to 0 elsewhere.

Approximating parity with small width



Approximating parity with small width



Approximating parity with small width

Theorem. There is a DNF of width n-Ω(n) that .01-approximates the parity function.

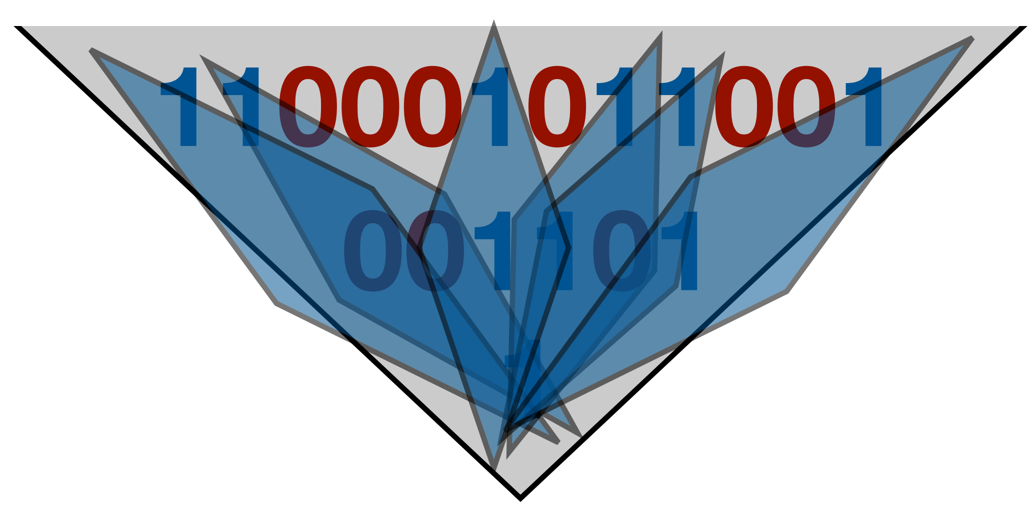
To complete the proof: We want to cover 99.9% of the hypercube with Hamming balls that overlap very little.

Lemma. There is a collection of O(2^n/Vol(d)) Hamming balls of radius d that cover 99.9% of the hypercube.

Approximating any function with small width

Theorem. For every function f:$\{0,1\}^n \rightarrow \{0,1\}$, there is a DNF of width n-Ω(n) that .01-approximates f.

Proof. Same!



3. Improved upper bounds for Parity

Approximating parity even better

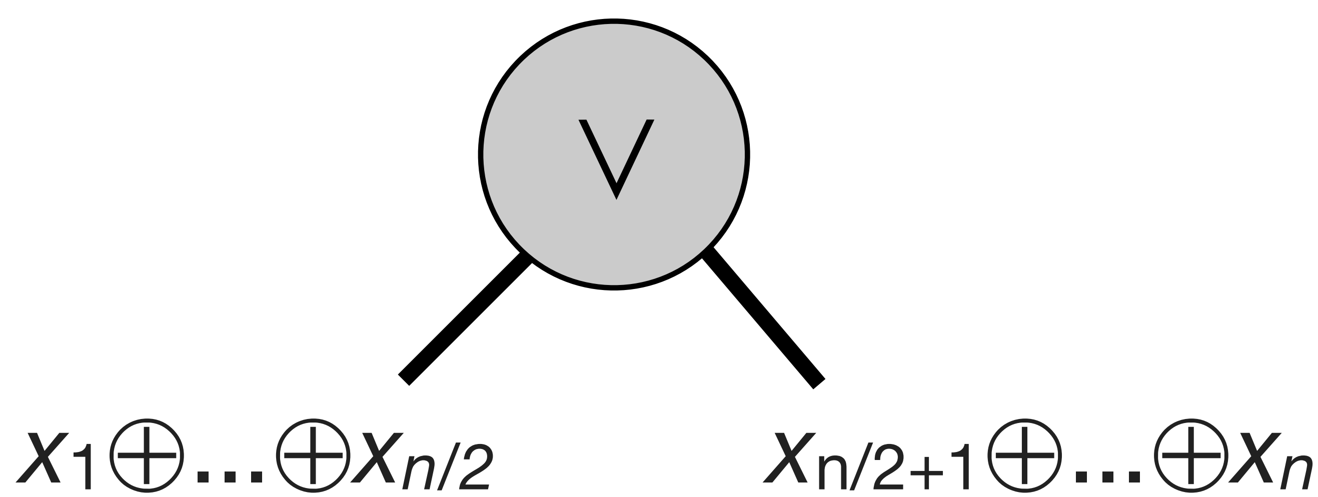
Theorem. There is a DNF of size $2^{.98n}$ and width $.98n$ that .01-approximates the parity function.

Proof strategy. Divide and conquer!

Approximating parity even better

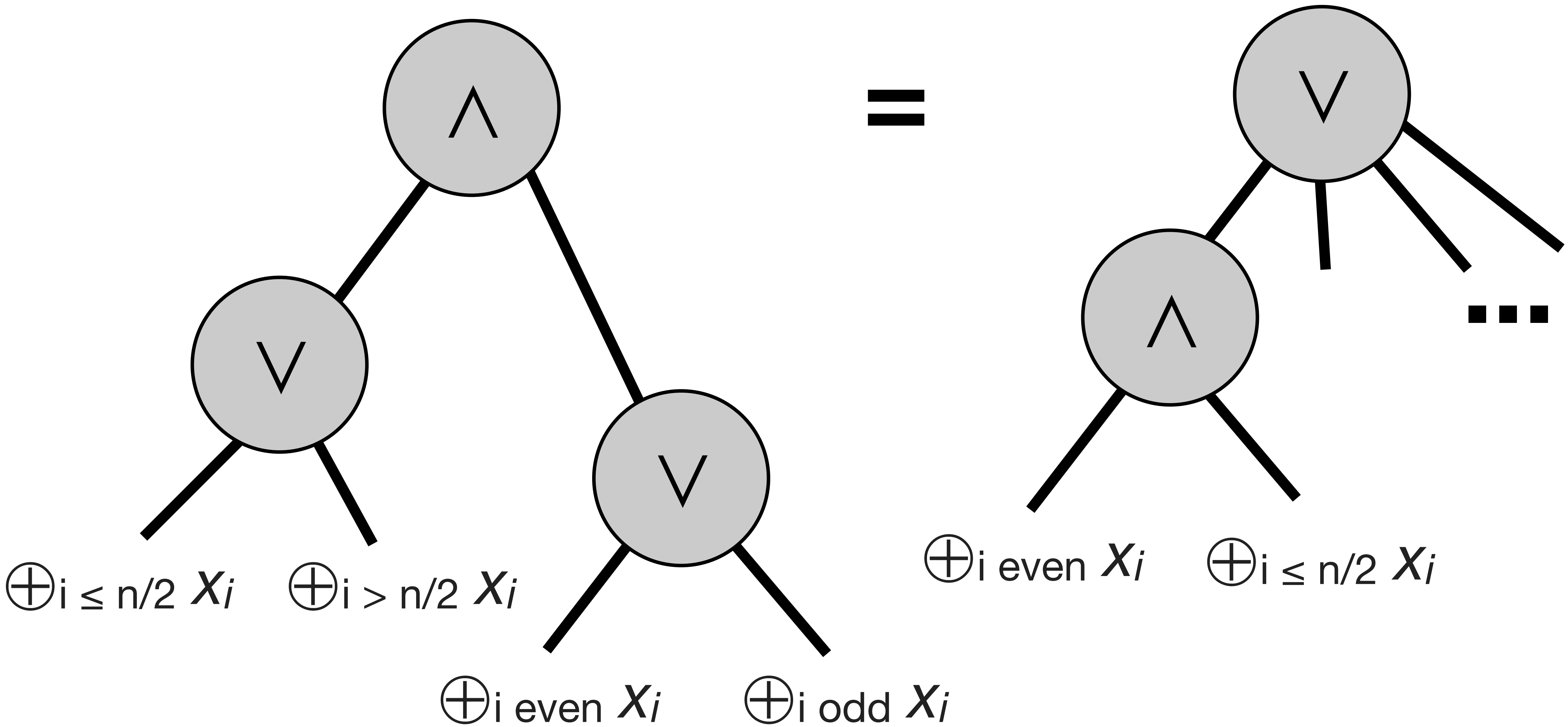
Theorem. There is a DNF of size $2^{.98n}$ and width $.98n$ that .01-approximates the parity function.

Proof strategy. Divide and conquer!



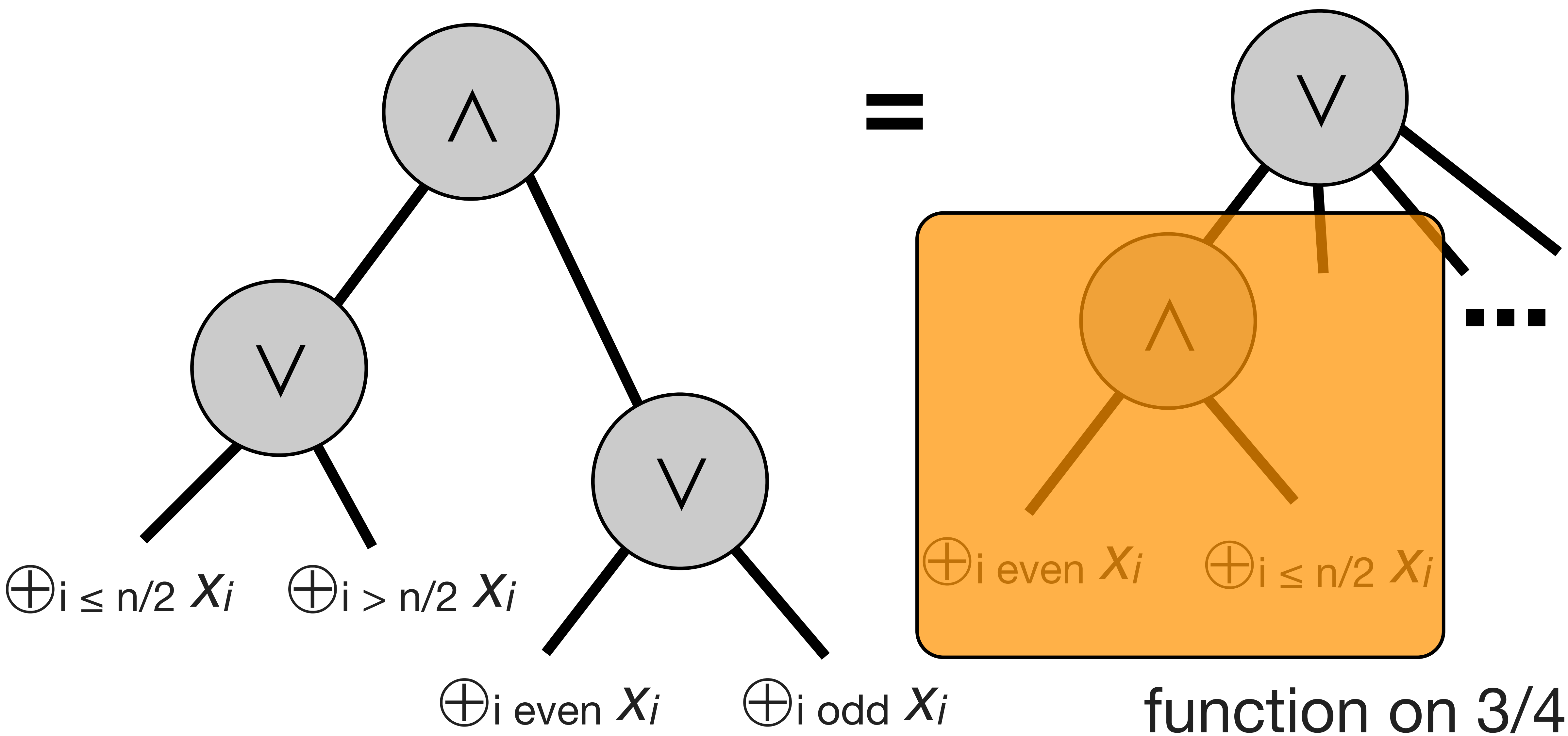
Approximating parity even better

Lowering the error probability...



Approximating parity even better

Lowering the error probability...



function on 3/4*n variables!

4. Lower bound on DNF size

Lower bounds

Theorem. For almost every function $f : \{0,1\}^n \to \{0,1\}$, every DNF computing f has size $\Omega(2^n/n)$.

Proof strategy. Entropy method.

Fact 1. If X is a random variable that can take m possible values, $H(X) \leq \log m$. Equality holds iff X is uniformly distributed among the m values.

Fact 2. $H(X,Y) = H(X) + H(Y \mid X)$.

Fact 3. $H(X \mid Y) \leq H(X)$.

Lower bounds

Let f be a random function.
Let $T = (T_1,\ldots,T_{3n}) \in \{0,1\}^{3n}$ denote which terms are in the smallest ε-approximating DNF for f.

Fact. $H(f,T) = H(f) = 2^n$.
Fact. $H(f \mid T) \leq H(\varepsilon) * 2^n$.
Corollary. $H(T) = H(f,T) - H(f \mid T) \geq (1-H(\varepsilon))\ 2^n$.

Finally:
$H(T) \leq \sum H(T_i) \leq 3^n\ H(E[\sum T_i] / 3^n) \leq E[\sum T_i] \log(3^n)$.

Lots more to explore!

Conjecture. Every function can be .01-approximated by a DNF of size $O(2^n/\log n \log \log n)$ and this bound is tight for almost every function.

Open Question. Find an *explicit* function f for which every DNF that .01-approximates f is larger than the DNF that approximates the parity function.

Lots more to explore!

Theorem (Quine ’54). Every monotone function f can be computed by a DNF of size $O(2^n/\sqrt{n})$.

- Maximum attained by Majority.
- Negations do not help.

Theorem (O’Donnell-Wimmer ’07). The majority function can be .01-approximated by a DNF of size $O(2^{\sqrt{n}})$.

What about universal bounds for approximating any monotone function? And do negations help?

Thank you!