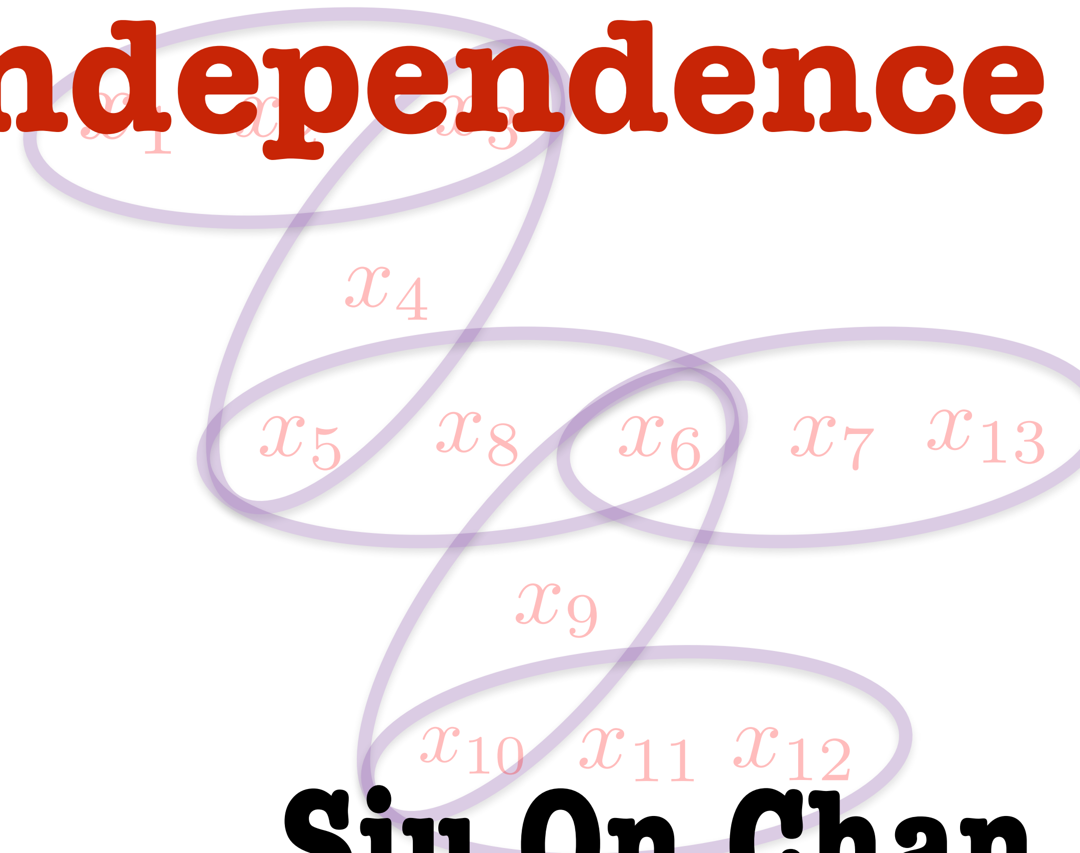
Sum of Squares Lower Bounds from Pairwise Independence

?

vs



Boaz Barak
MSR

Siu On Chan
MSR

CSP

Pravesh K. Kothari
UT Austin

Constraint Satisfaction

n Boolean **variables**.

x_1 x_2 x_3

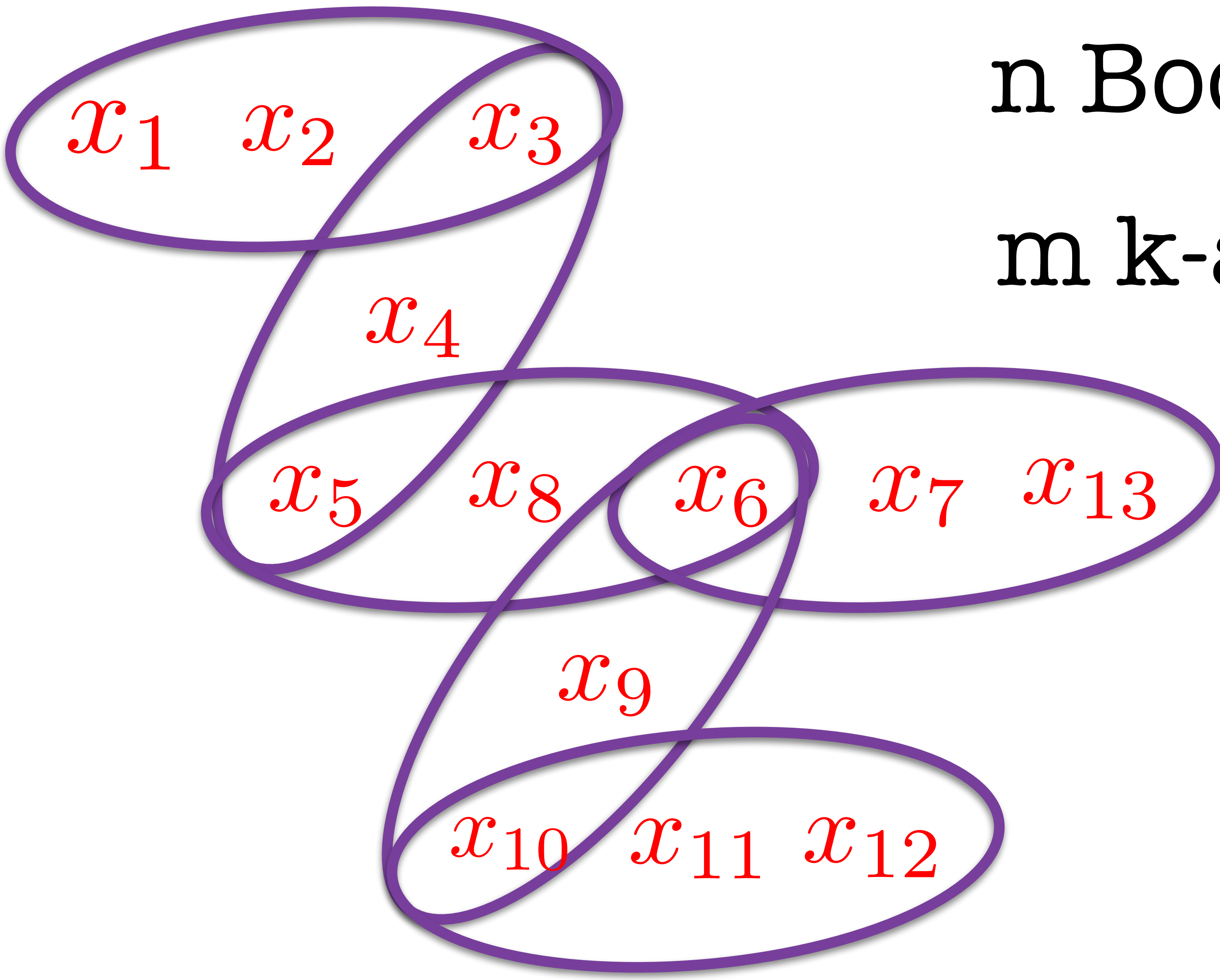
x_4

x_5 x_8 x_6 x_7 x_{13}

x_9

x_{10} x_{11} x_{12}

Constraint Satisfaction

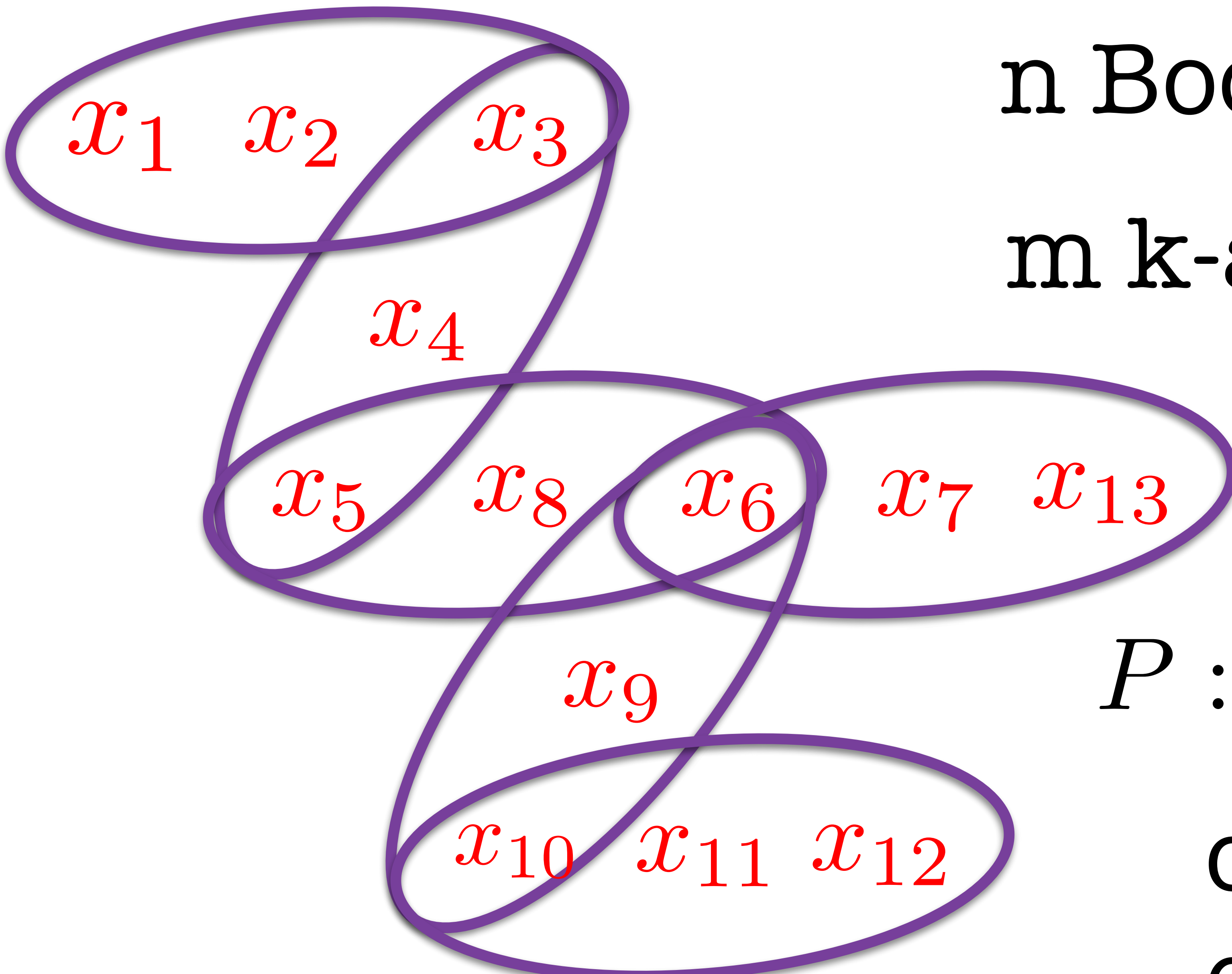


n Boolean **variables**.

m k-ary **constraints**

k = 3

Constraint Satisfaction



CSP(P)

n Boolean **variables**.

m k-ary **constraints**

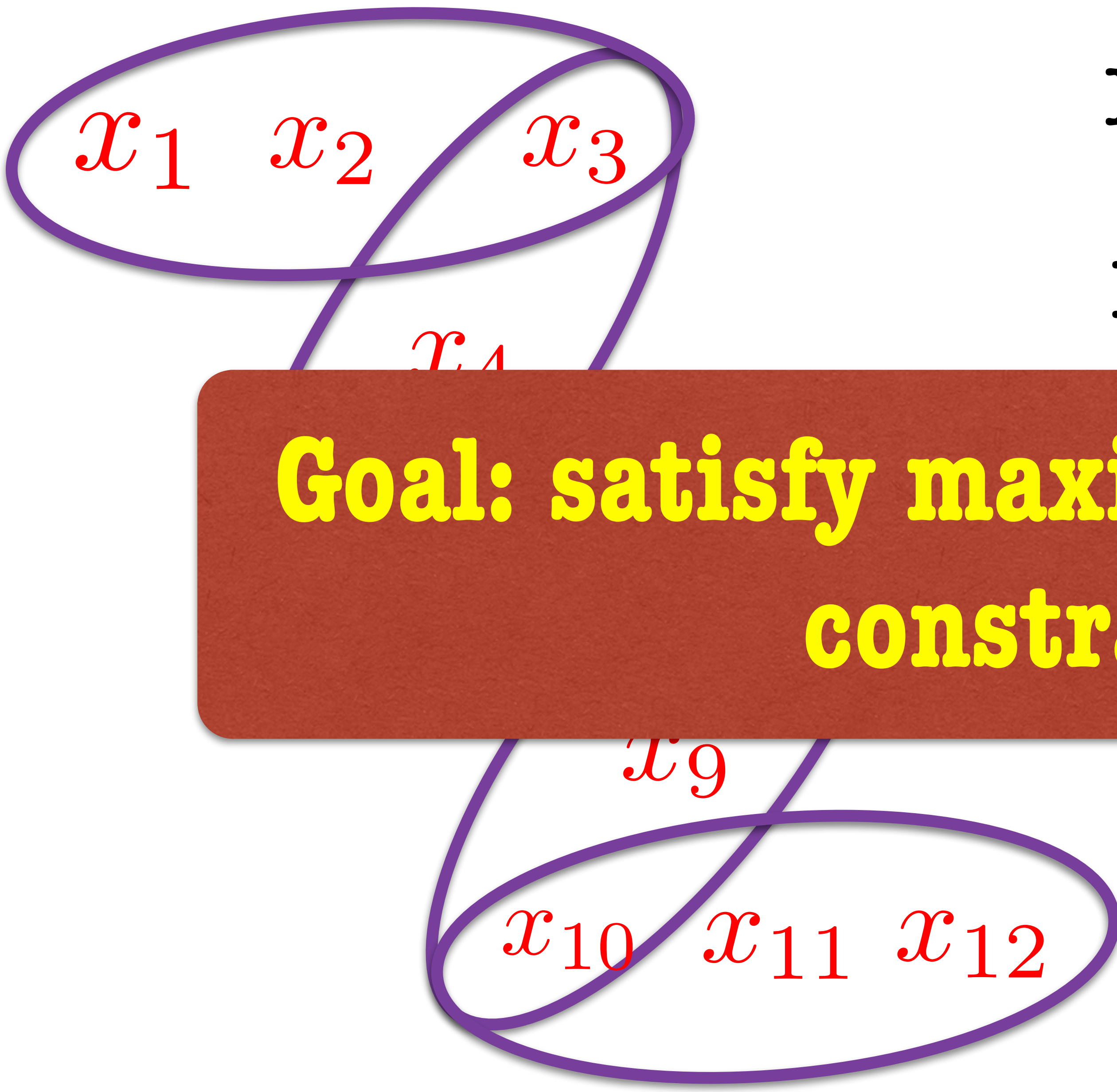
Predicate

$$P : \{0,1\}^k \to \{0,1\}$$

on **literals** on constraints.

$\ell_i = x_i$ or $\neg x_i$

Constraint Satisfaction



CSP(P)

n Boolean **variables**.

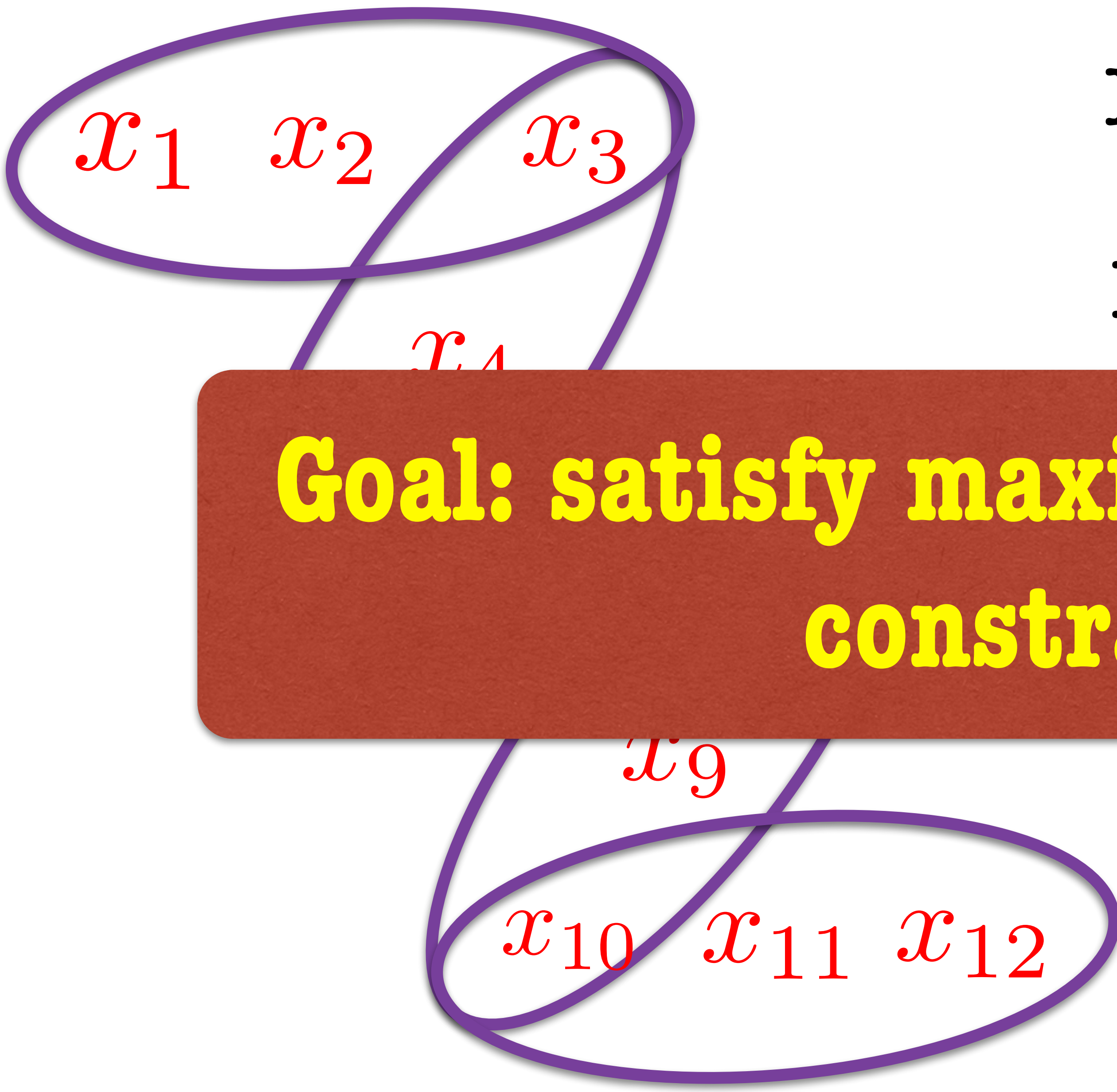
m k-ary **constraints**

$P : \{0, 1\} \to \{0, 1\}$

on **literals** on constraints.

Goal: satisfy maximum fraction of constraints.

Constraint Satisfaction



n Boolean **variables**.

m k-ary **constraints**

Goal: satisfy maximum fraction of constraints.

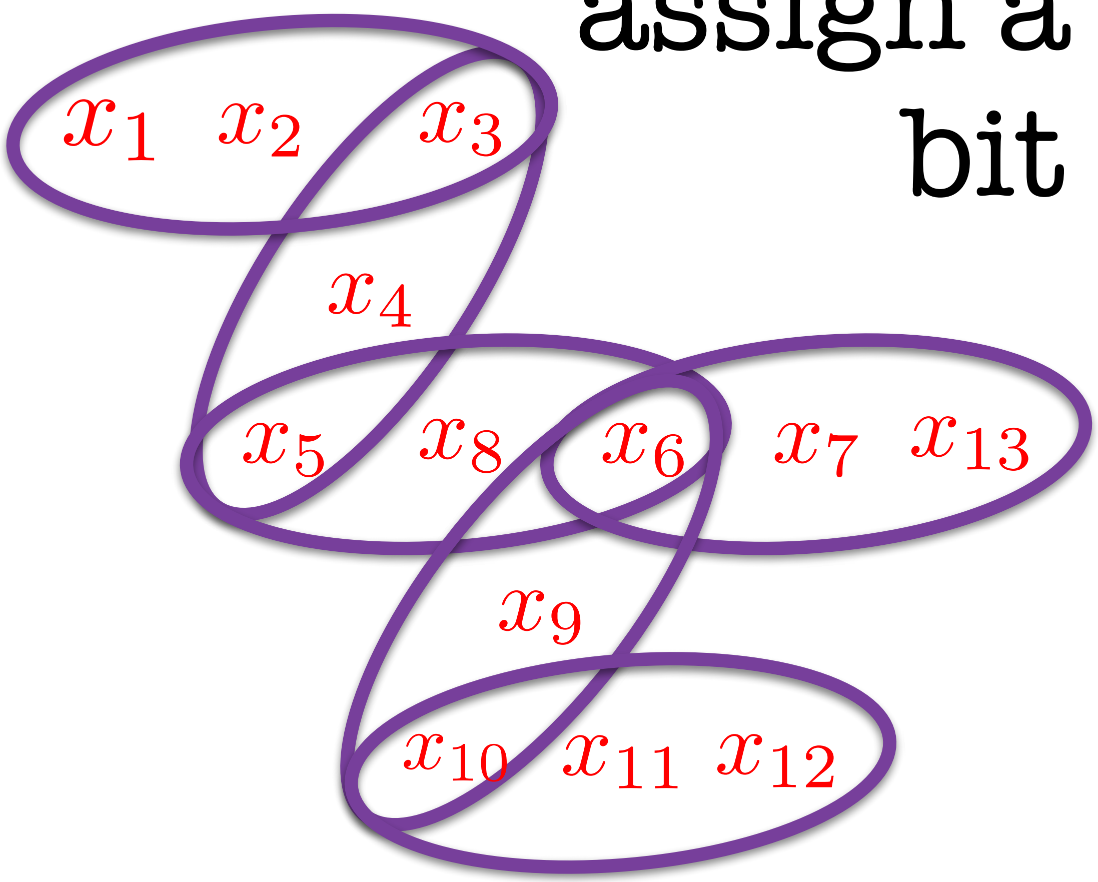
$P : \{0, 1\} \to \{0, 1\}$

on **literals** on constraints.

CSP(P)

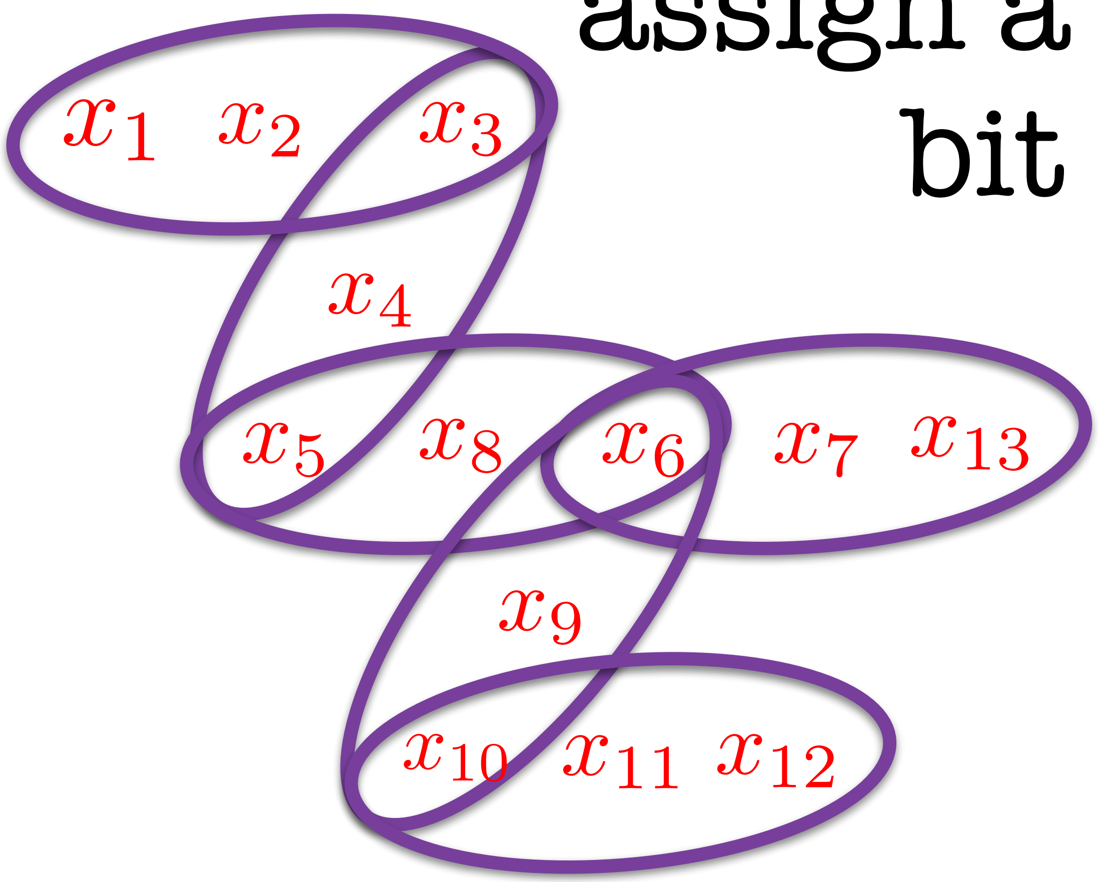
MAX-3SAT, MAX-3XOR, ...

Random Assignment



assign a uniform, independent bit to each variable.

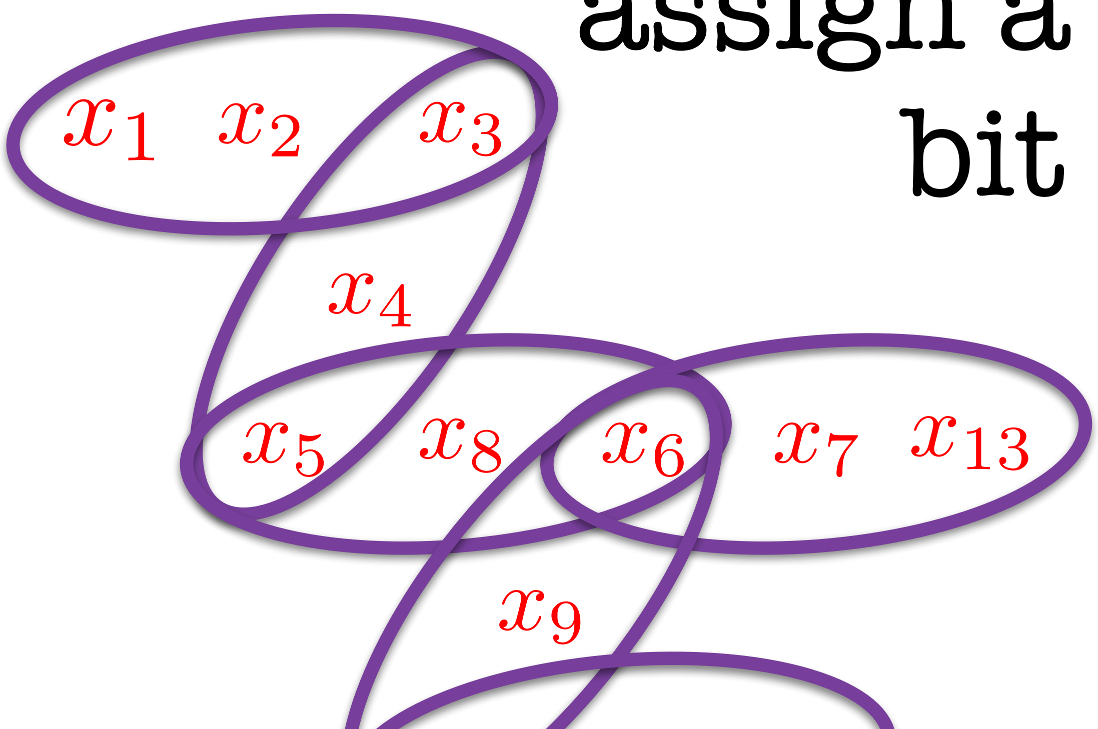
Random Assignment



assign a uniform, independent bit to each variable.

satisfies $\frac{|P^{-1}(1)|}{2^k}$ fraction of constraints.

Random Assignment

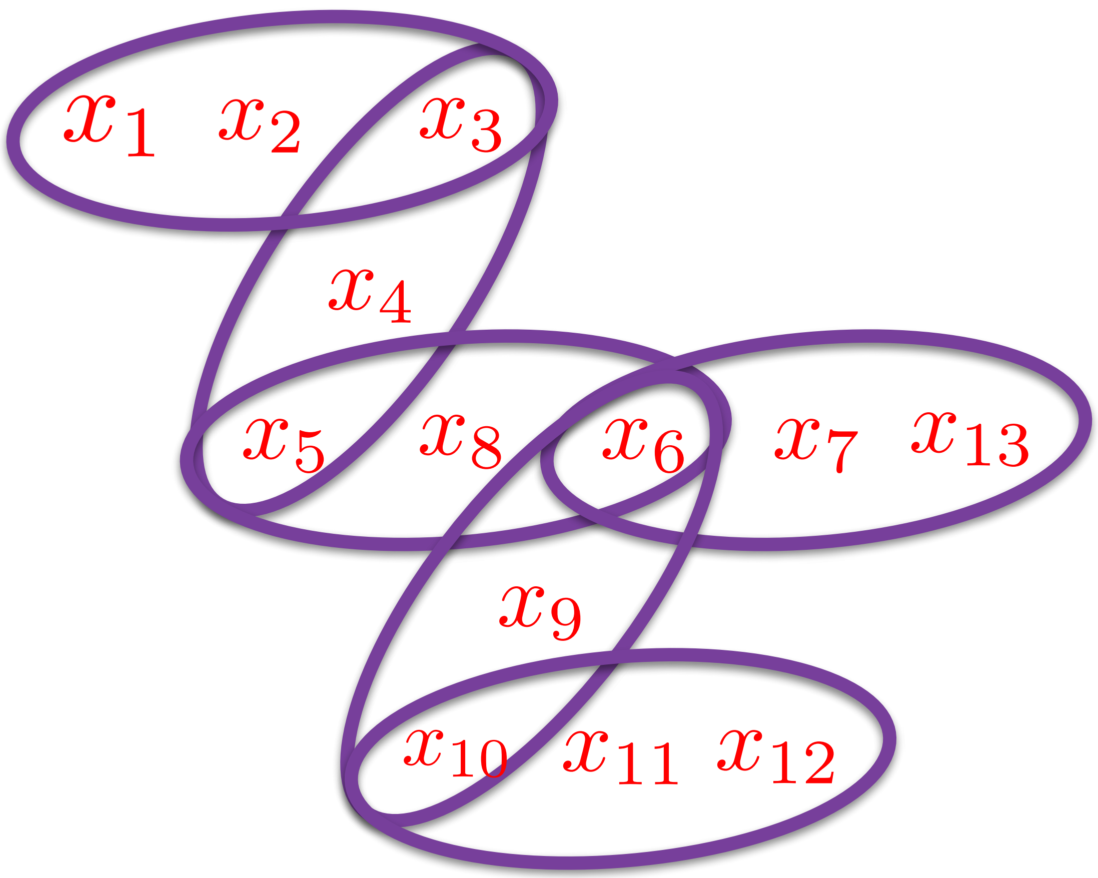


assign a uniform, independent bit to each variable.

satisfies $\frac{|P^{-1}(1)|}{2^k}$ fraction of constraints.

Can we do better?

Random Assignment

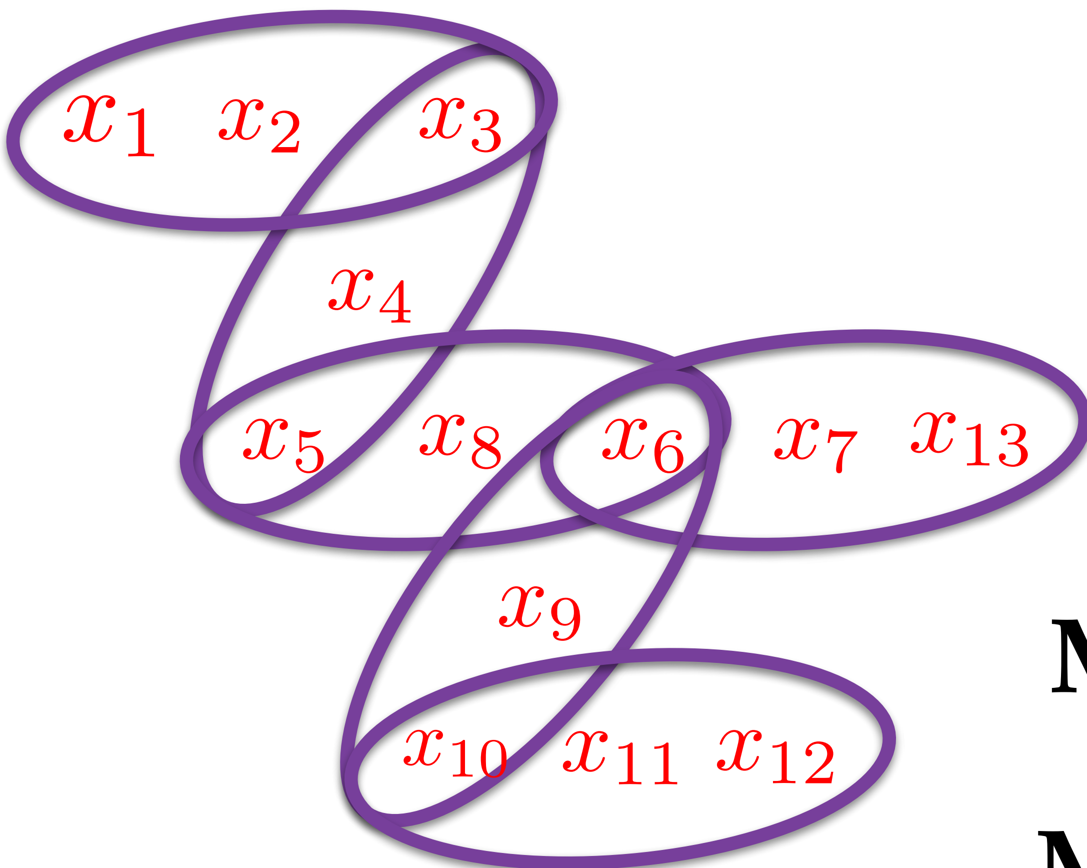


MAX CUT Yes! [GW95]

MAX 2LIN Yes! [AEH01]

MAX 3MAJ Yes! [Zwi98]

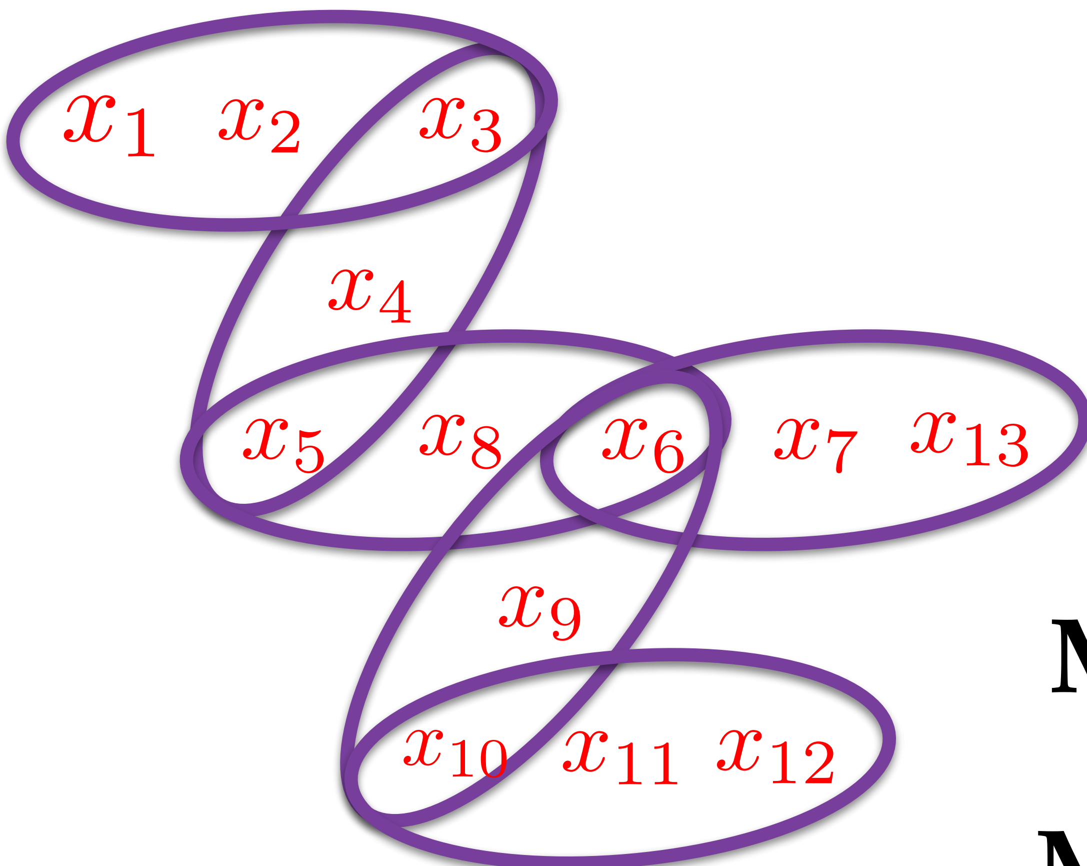
Random Assignment



MAX CUT	Yes!	[GW95]
MAX 2LIN	Yes!	[AEH01]
MAX 3MAJ	Yes!	[Zwi98]
MAX k-SAT*	No!	[Hås99]
MAX k-XOR*	No!	[Hås99]
MAX TSA	No!	[KH01]

k = 3 and above

Random Assignment

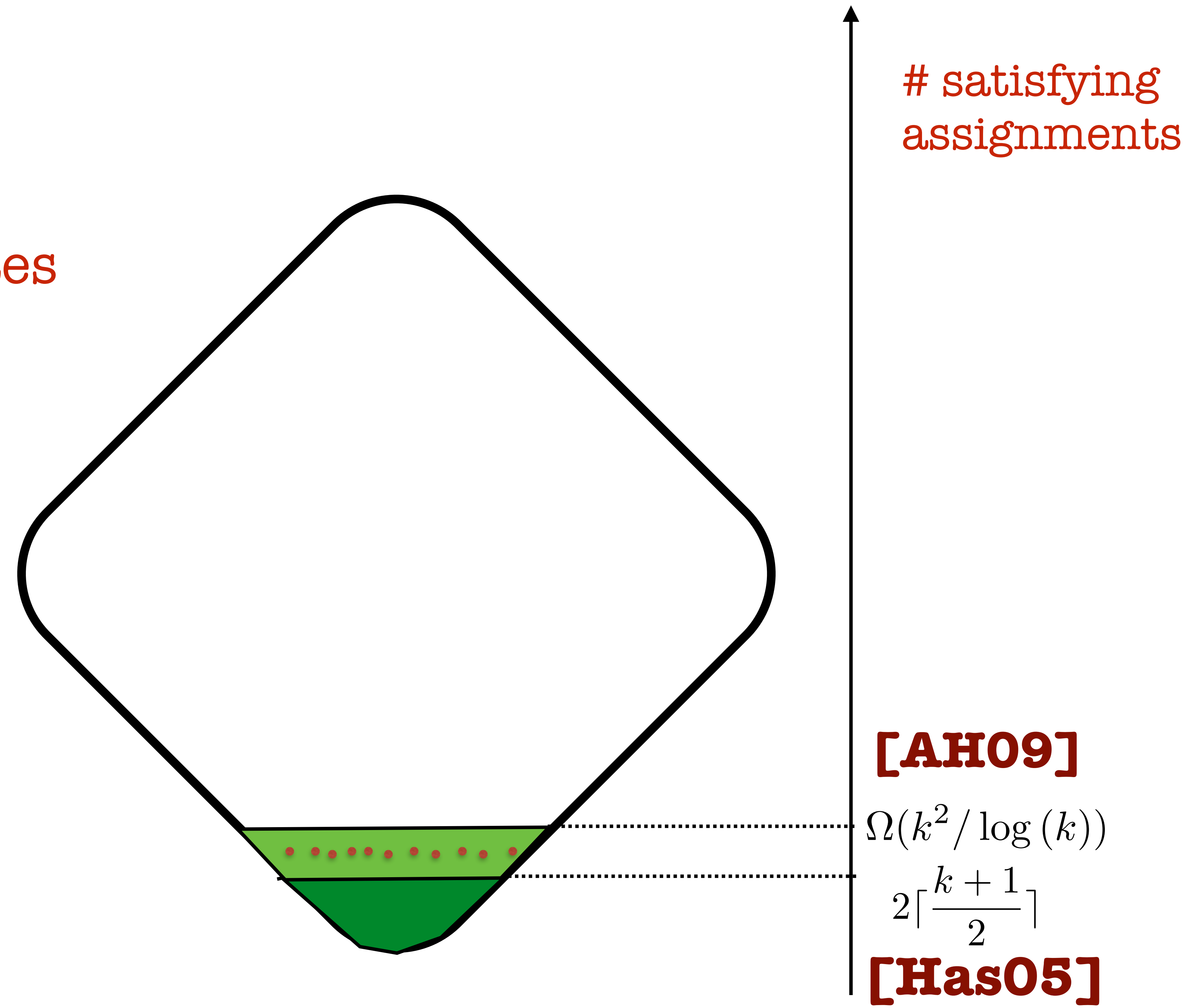


MAX CUT	Yes!	[GW95]
MAX 2LIN	Yes!	[AEH01]
MAX 3MAJ	Yes!	[Zwi98]
MAX k-SAT*	No!	[Hås99]
MAX k-XOR*	No!	[Hås99]
MAX TSA	No!	[KH01]

"approximation resistance"

Green = can efficiently beat the random assignment.
Red = approximation resistant.

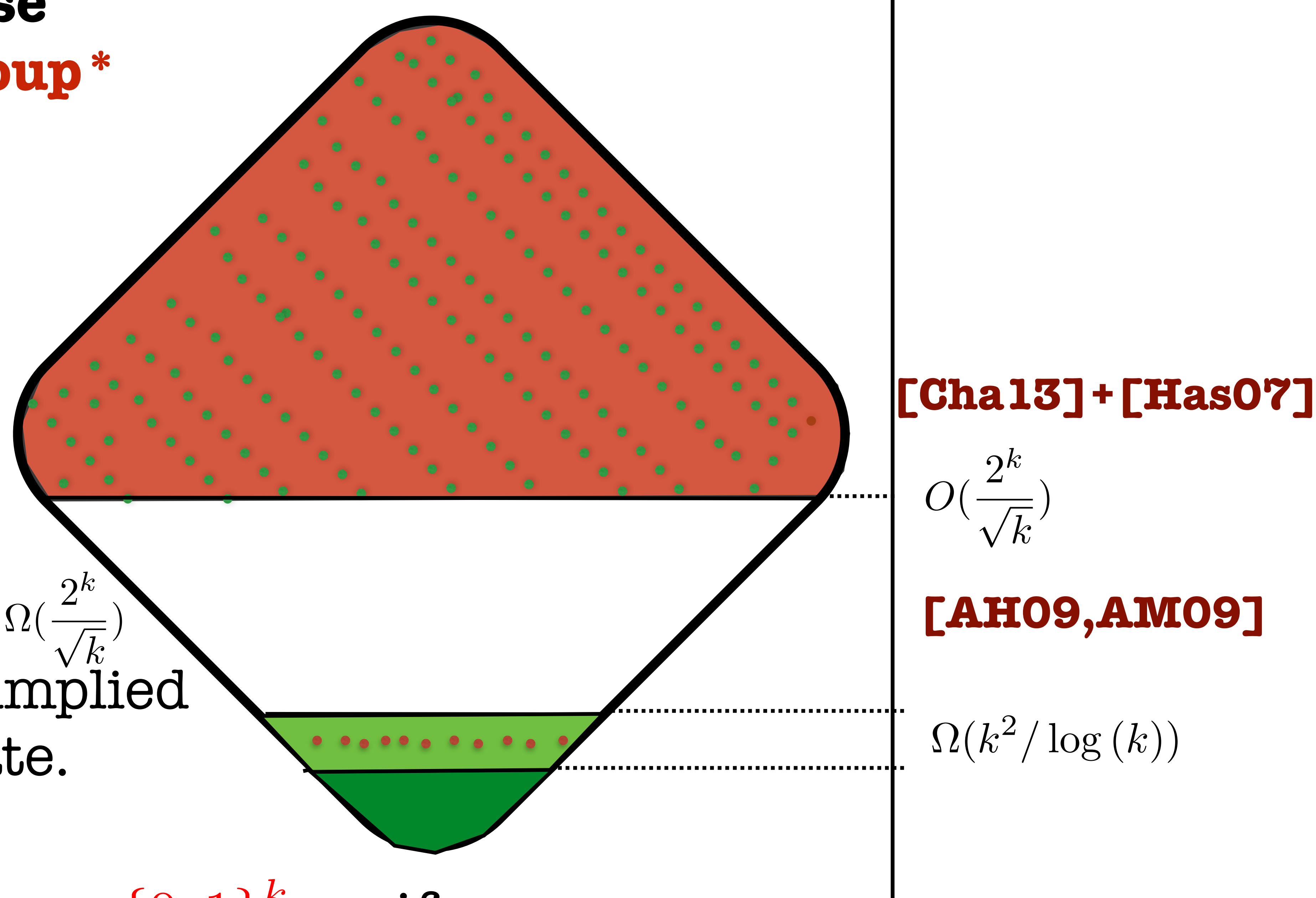
can beat random
on **most** predicates
with sparsity
$O(k^2/\log(k))$



Green = can efficiently beat the random assignment.
Red = approximation resistant.

[Cha13]
NP hard to beat random for CSP(P) whenever P supports a **pairwise independent subgroup** *

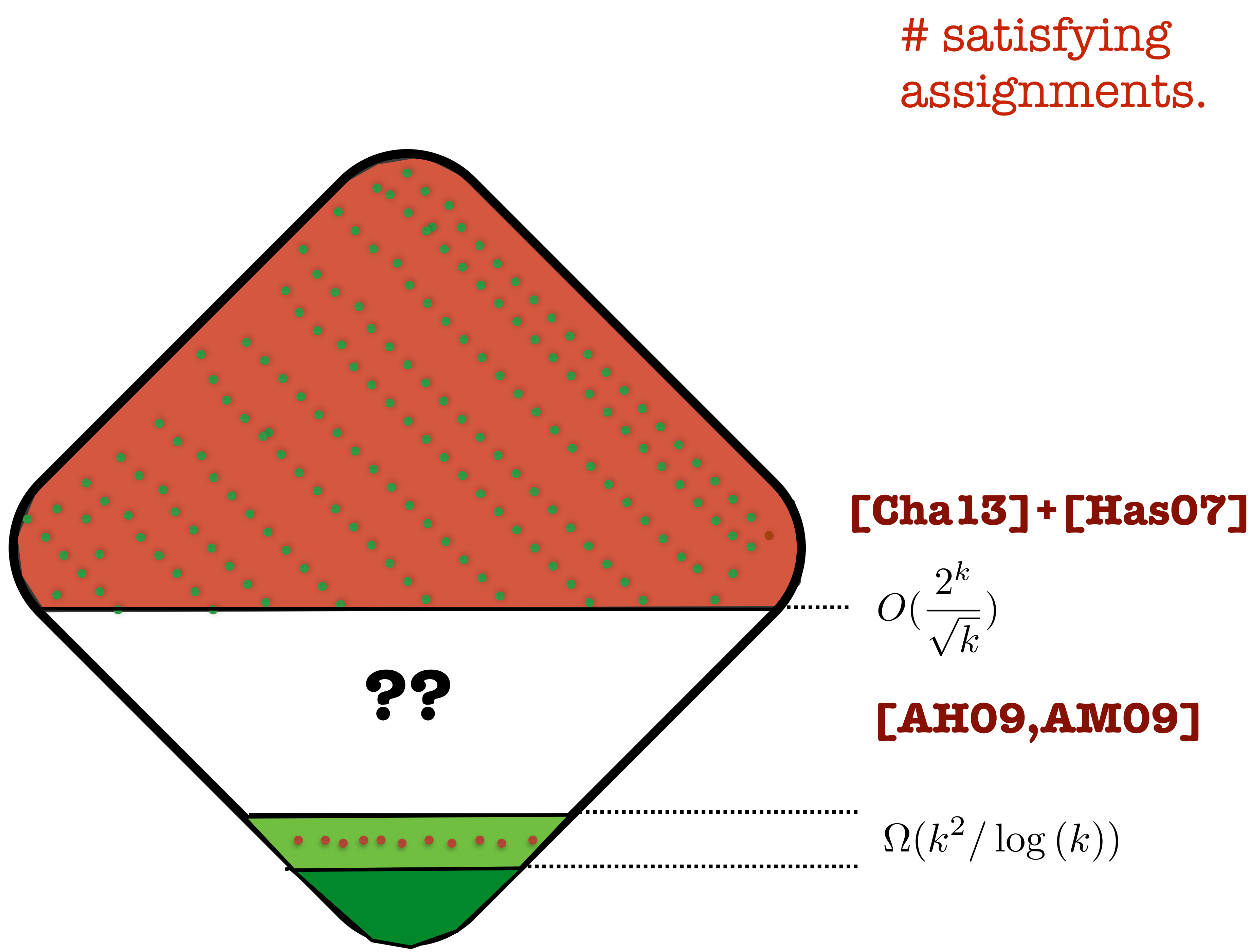
satisfying assignments.



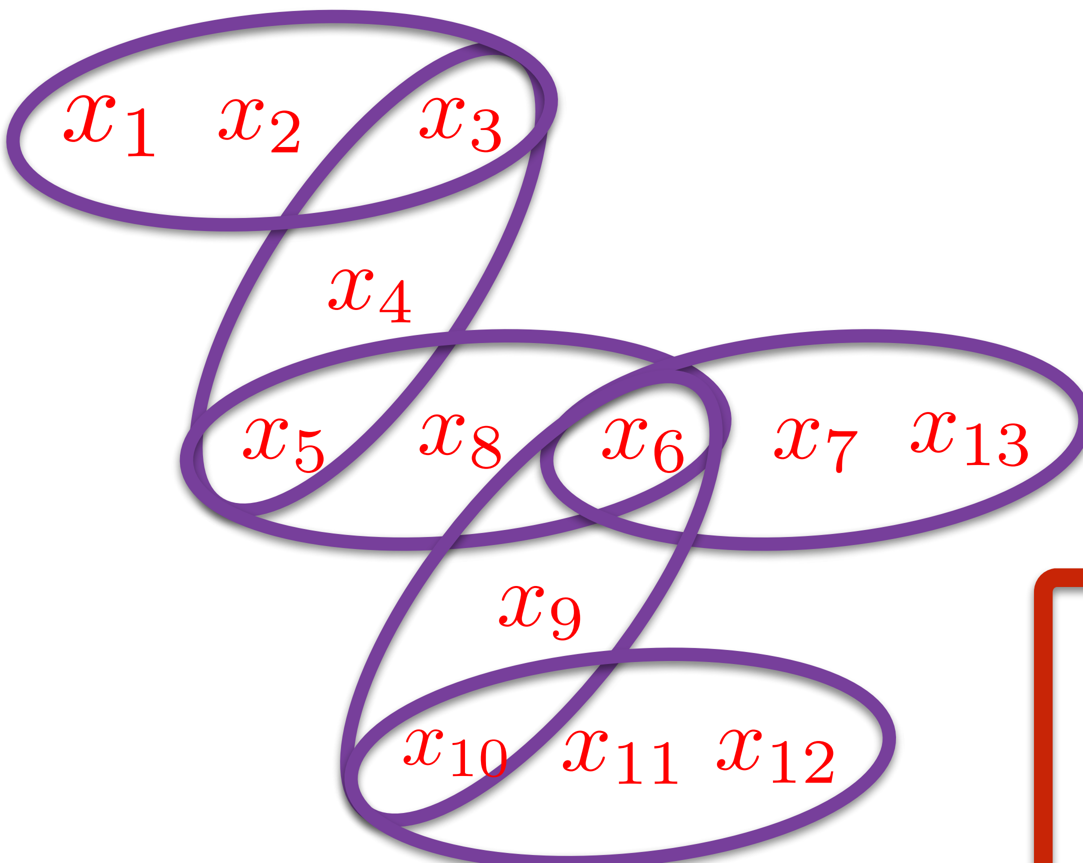
[Has07]
most predicates with $\Omega(\frac{2^k}{\sqrt{k}})$ sat. assignments are implied by Hadamard predicate.

* $\exists$ a prob. dist. μ on $\{0,1\}^k$, uniform on a subgroup inside $P^{-1}(1)$ s.t. $\Pr_{x \sim \mu}[x_i = a, x_j = b] = 1/4$

Green = can efficiently beat the random assignment.
Red = approximation resistant.



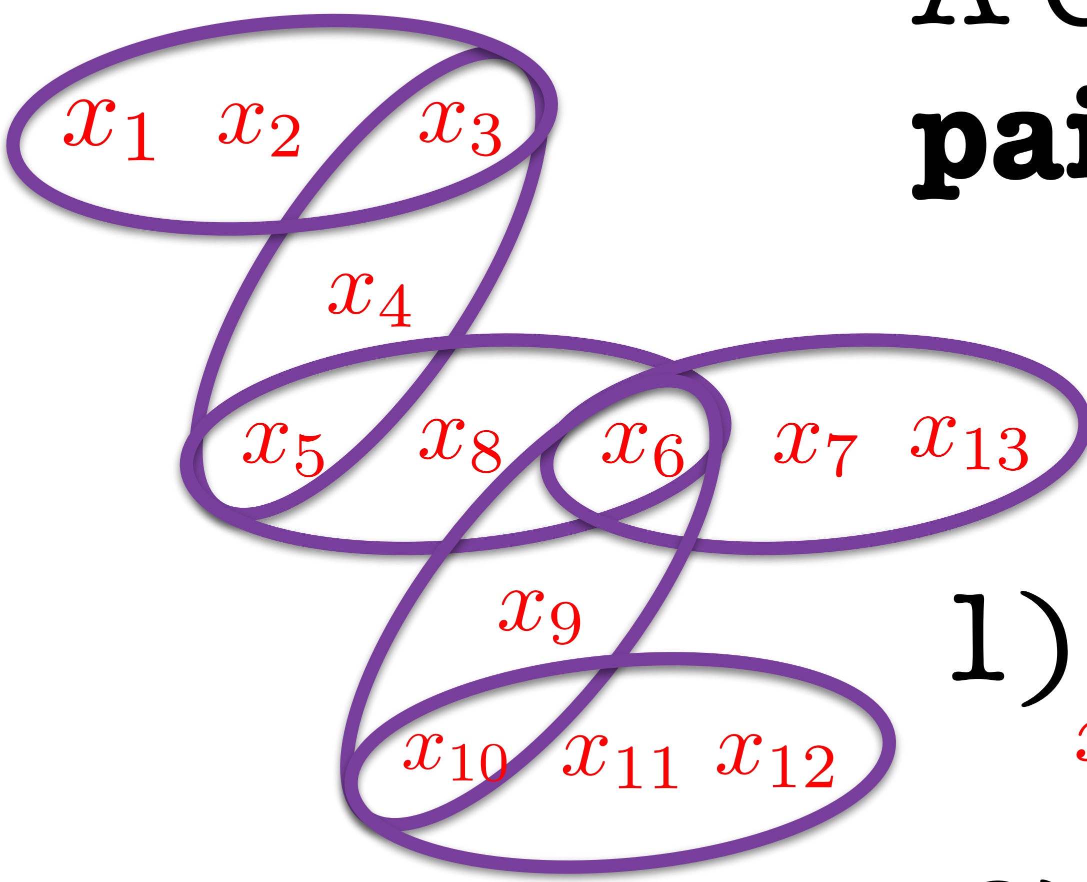
Random Assignment



MAX CUT	Yes!	[GW95]
MAX 2LIN	Yes!	[AEH01]
MAX 3MAJ	Yes!	[Zwi98]
MAX k-SAT	No!	[Hås09]
MAX k-XOR	No!	[Hås09]
MAX TSA	No!	[KH01]

“pairwise independence”

Pairwise Independence



$P : \{0,1\}^k \to \{0,1\}$

A CSP(P) is **pairwise independent** if

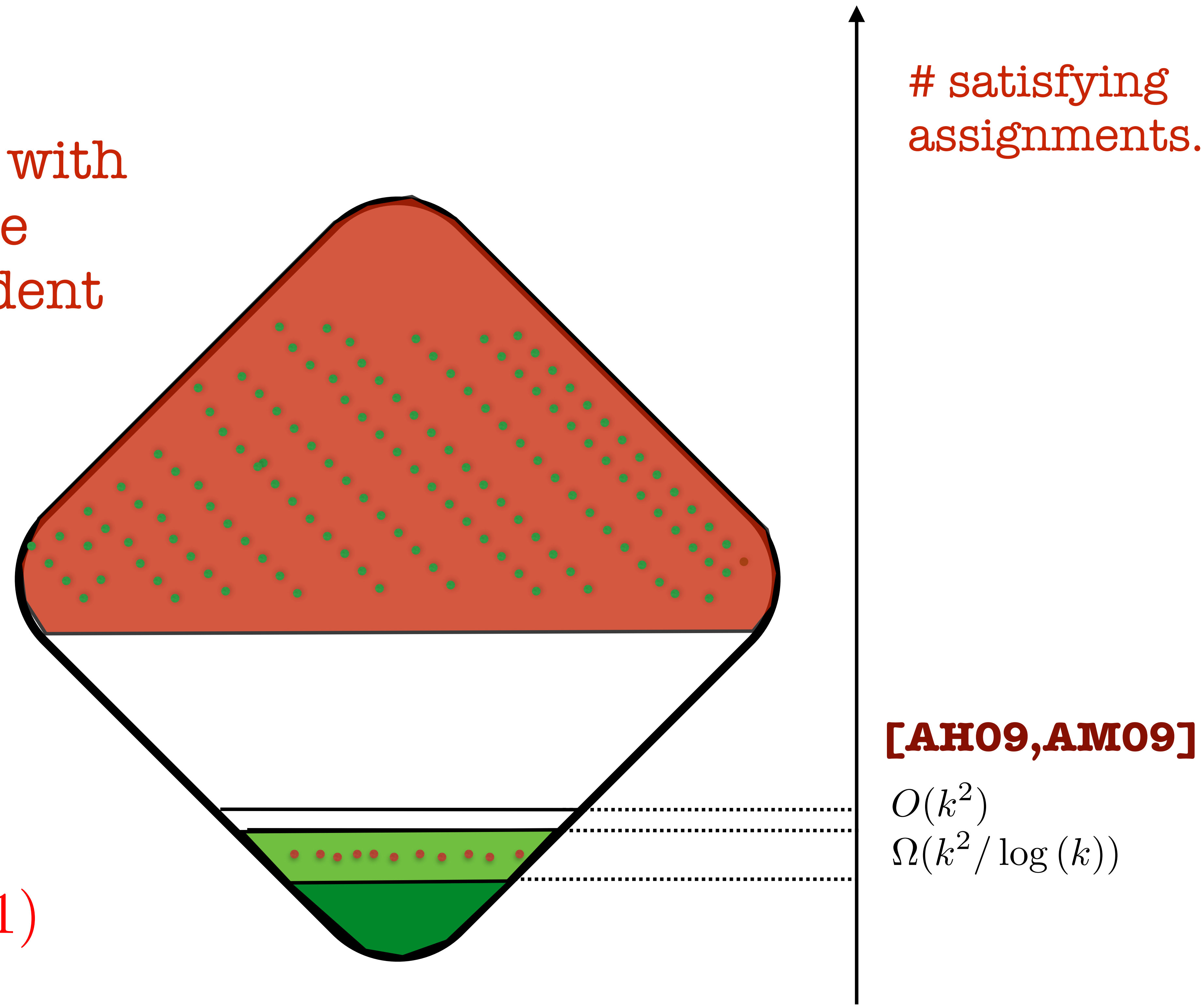
there is a prob. dist. μ on $\{0,1\}^k$

1) $\Pr_{x \sim \mu}[x_i = a, x_j = b] = 1/4$

2) supported on $P^{-1}(1)$

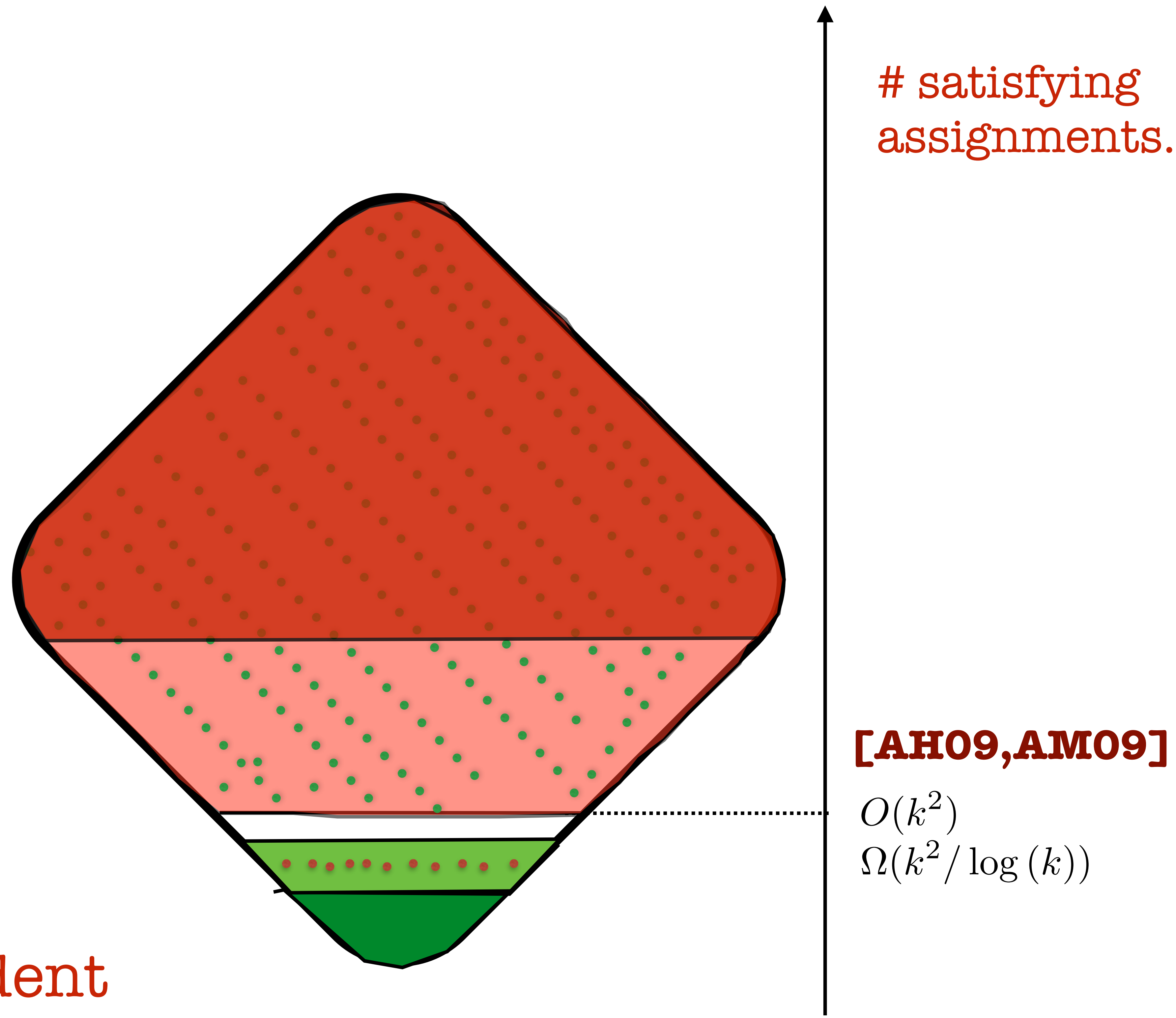
Green = can efficiently beat the random assignment.
Red = approximation resistant.

[AH09]
most* predicates with sparsity $\Omega(k^2)$ are pairwise independent



*__most__ = $1 - o_k(1)$

Green = can efficiently beat the random assignment.
Red = approximation resistant.



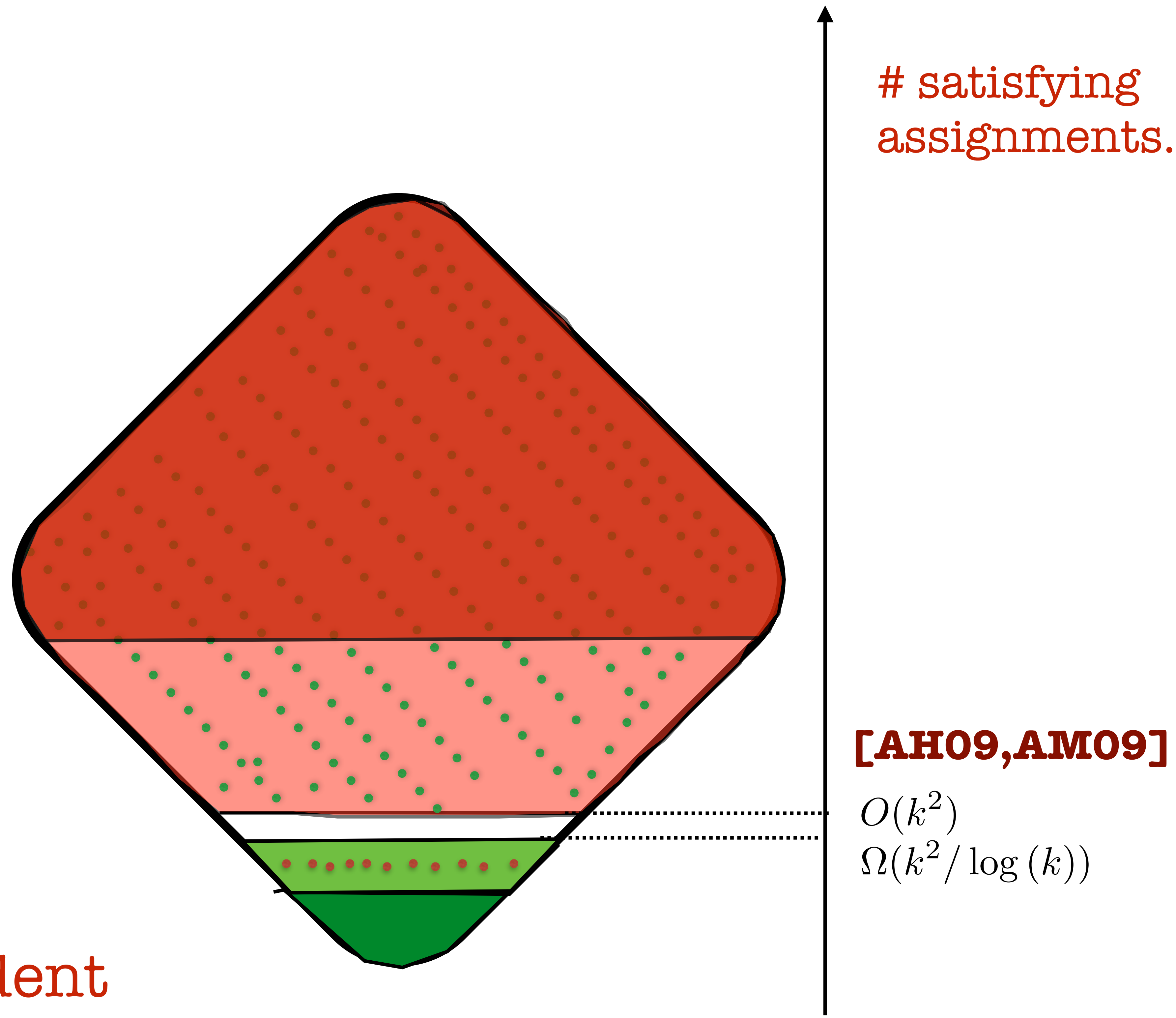
[AM09]
UGC

pairwise independent CSPs are approximation resistant.

Green = can efficiently beat the random assignment.
Red = approximation resistant.

basic sdp
cannot
beat random

[Rag08]



[AM09]
UGC

pairwise independent
CSPs are approximation resistant.

Green = can efficiently beat the random assignment.
Red = approximation resistant.

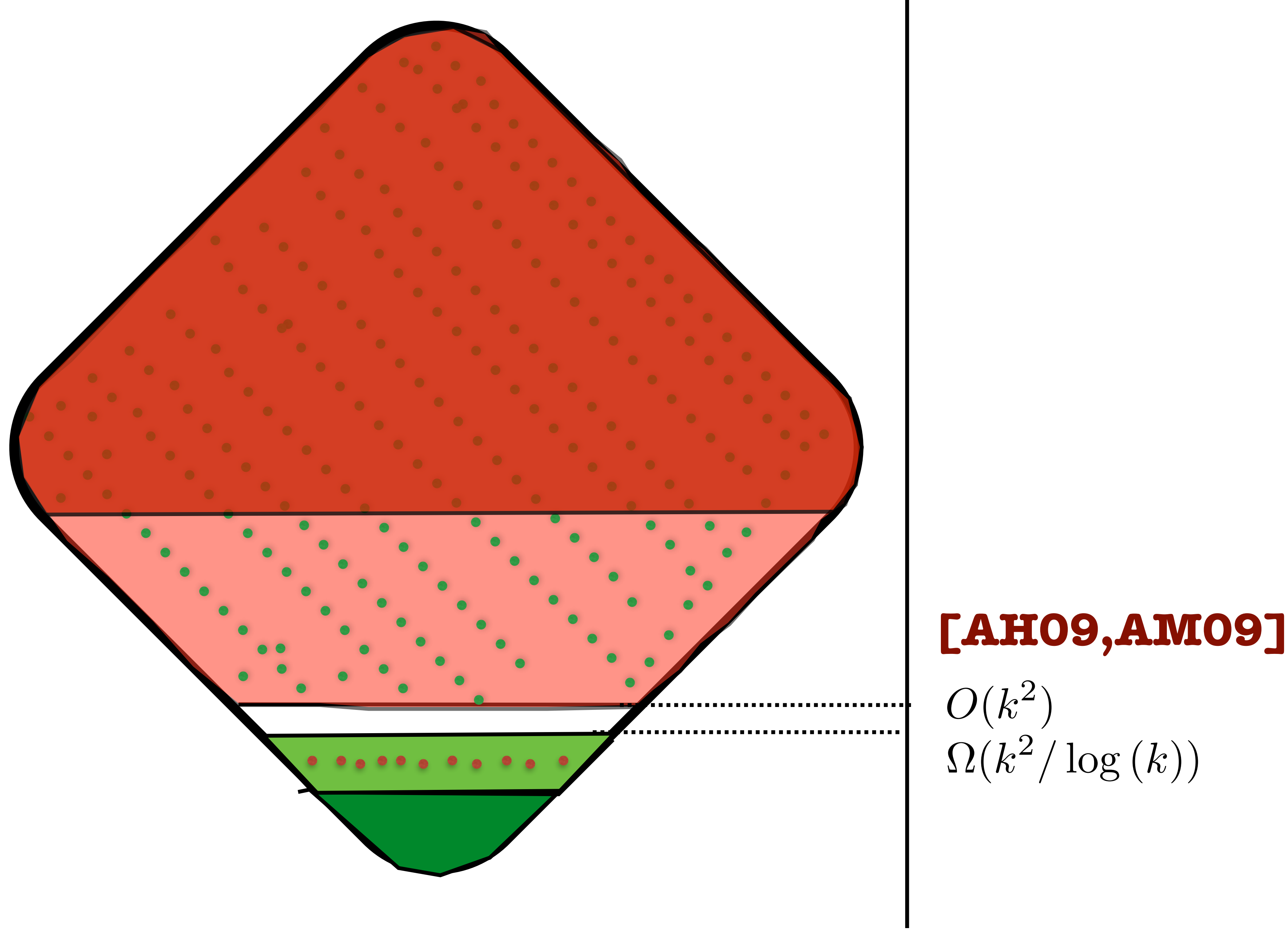
$\Omega(n)$ **round Sherali Adams/Lovasz Schrijver**

$+$

basic sdp
cannot
beat random

[BGMT09,TW13]

satisfying assignments.



Sum of Squares SDP

[Grigoriev-Vorobjov99,Shor87,Nesterov99,Parillo00,Lasserre01]

hierarchical strengthening
of the basic SDP.

Sum of Squares SDP

[Grigoriev-Vorobjov99,Shor87,Nesterov99,Parillo00,Lasserre01]

highly successful convex relaxation for comb. optimization.

SPARSEST CUT **[ARV04]**

SPARSE CODING **[BKS14]**

Sum of Squares SDP

[Grigoriev-Vorobjov99,Shor87,Nesterov99,Parillo00,Lasserre01]

highly successful convex relaxation for comb. optimization.

SPARSEST CUT [ARV04]

SPARSE CODING [BKS14]

UNIQUE GAMES [ABS10, BRS12, GS12]

Sum of Squares SDP

[Grigoriev-Vorobjov99,Shor87,Nesterov99,Parillo00,Lasserre01]

highly successful convex relaxation for comb. optimization.

SPARSEST CUT [ARV04]

SPARSE CODING [BKS14]

UNIQUE GAMES [ABS10, BRS12, GS12]

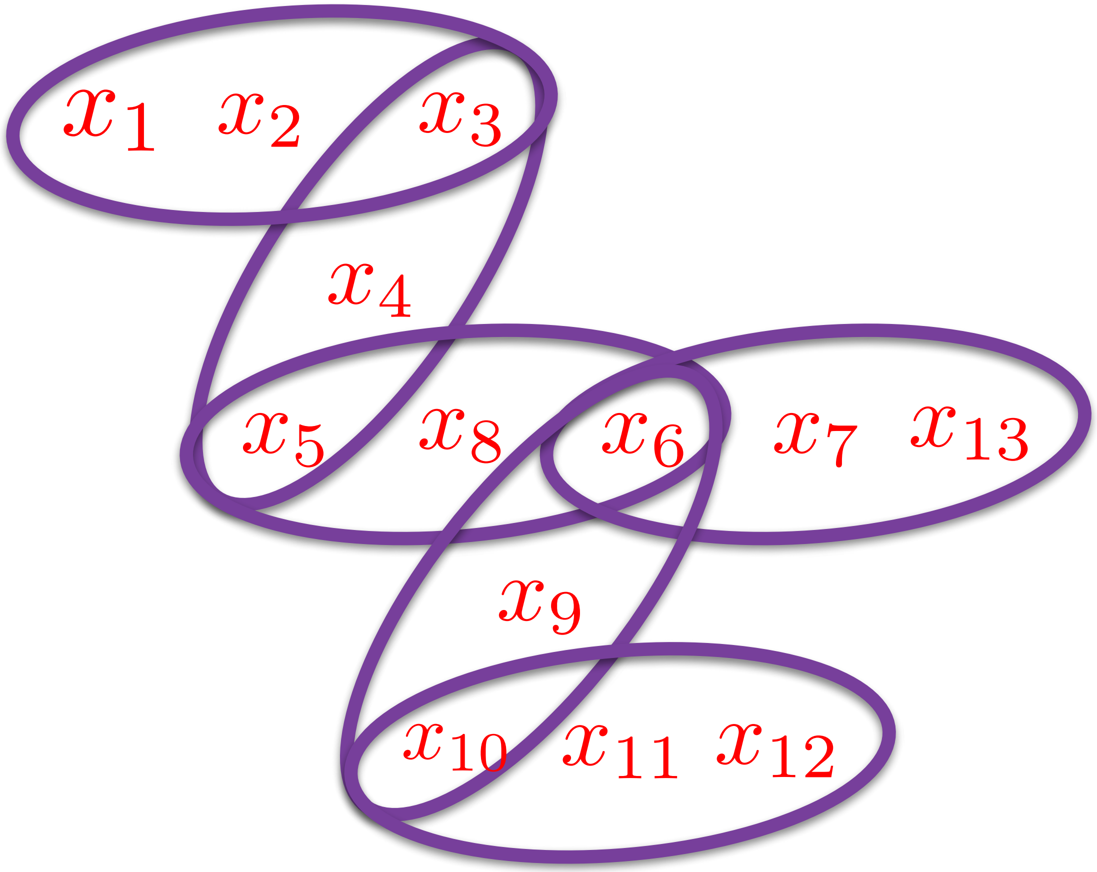
....solves all known hard instances for weaker hierarchies in **polynomial** time of

UNIQUE GAMES, BALANCED SEPARATOR, MAX CUT

[BBHKSZ12, OZ13,DMN13]

SOS vs CSP?

?
vs



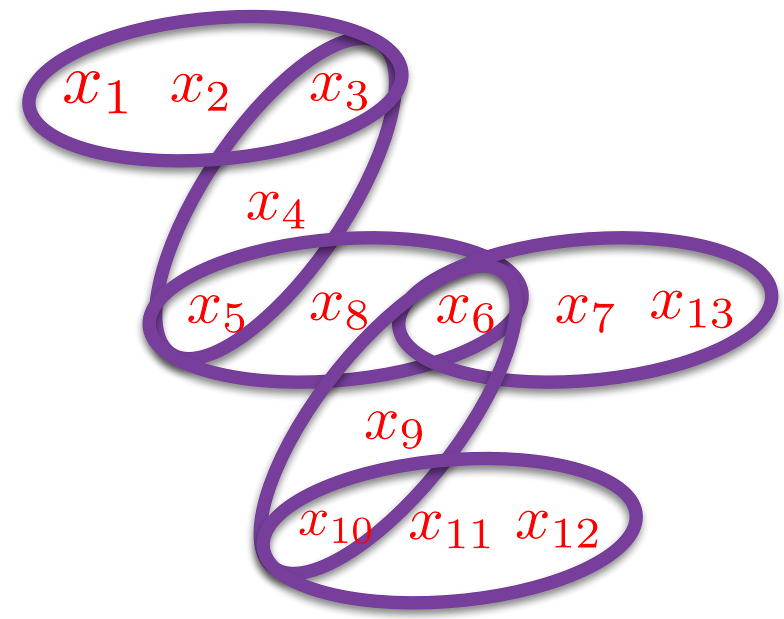
SOS

CSP

SOS vs CSP?

?

vs



[Gri01,Sch08] k-XOR requires **exponential** SOS.

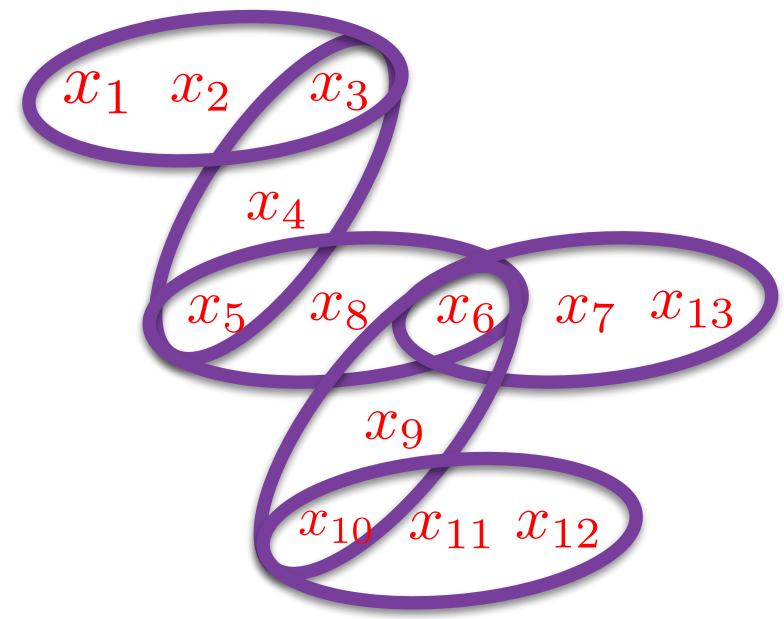
[Tul09,Cha13] Predicates supporting a **pairwise independent subgroup*** require exponential SOS.

* μ is uniform on a subgroup inside $P^{-1}(1)$

SOS vs CSP?

?

vs



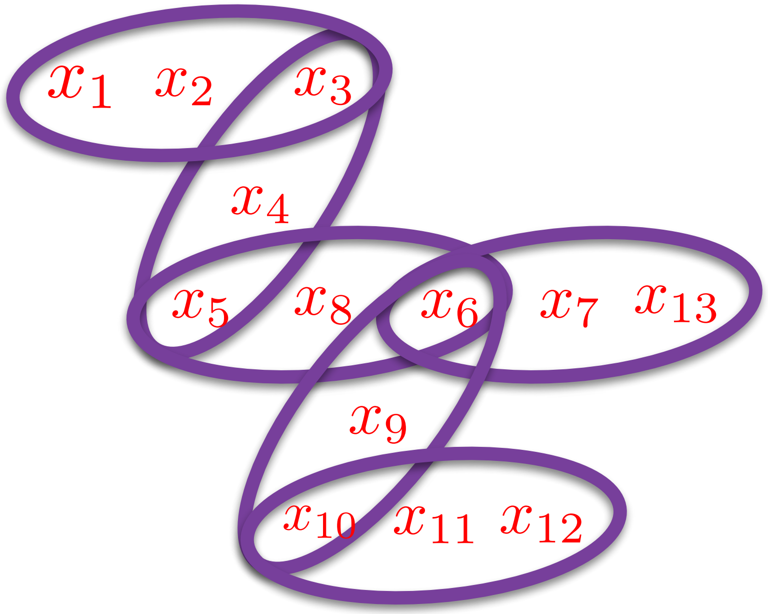
[Gri01,Sch08] k-XOR requires **exponential** SOS.

[Tul09,Cha13] Predicates supporting a **pairwise independent subgroup*** require exponential SOS.

Strong algebraic condition. Almost no sparse predicates satisfy!

Our Result

? vs

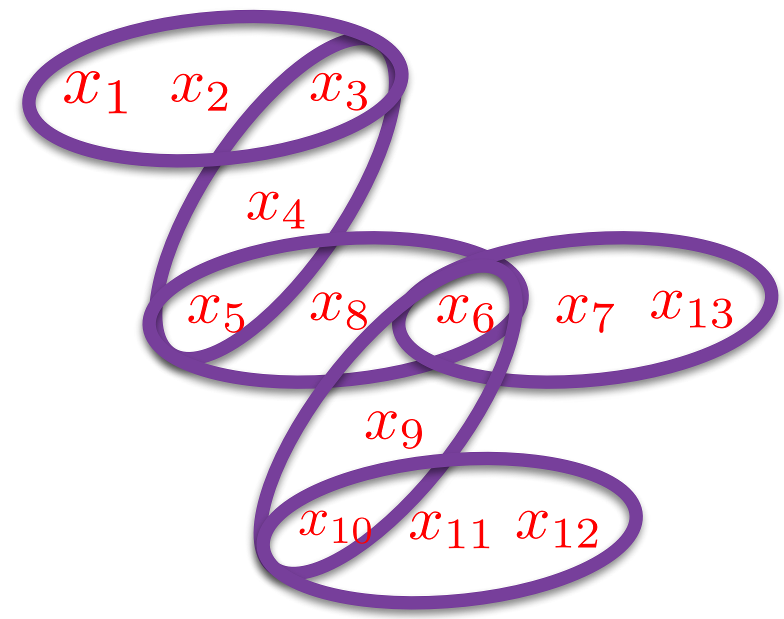


SOS needs exponential time to beat random for pairwise independent CSPs.

Our Result

?

vs



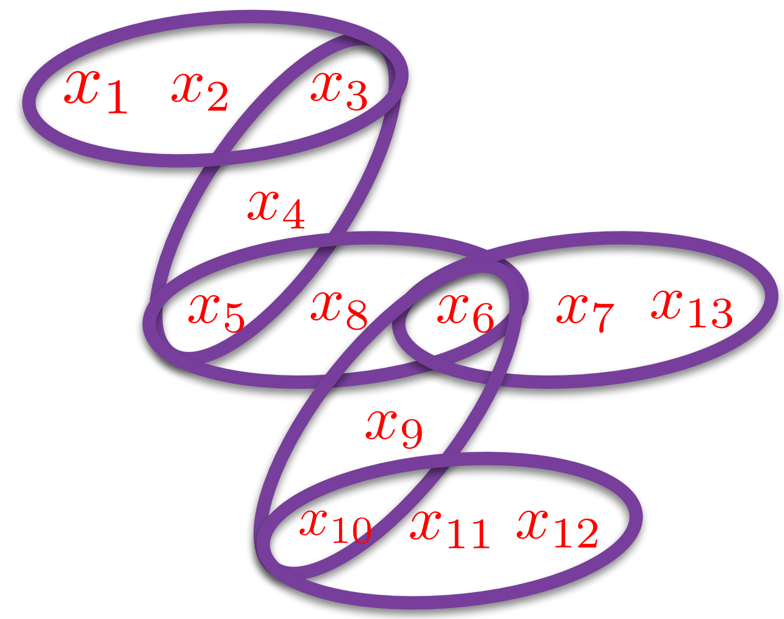
Previous Proofs

Reduction to **Resolution Width** lower bounds algebraic structure **crucial.**

Our Result

?

vs



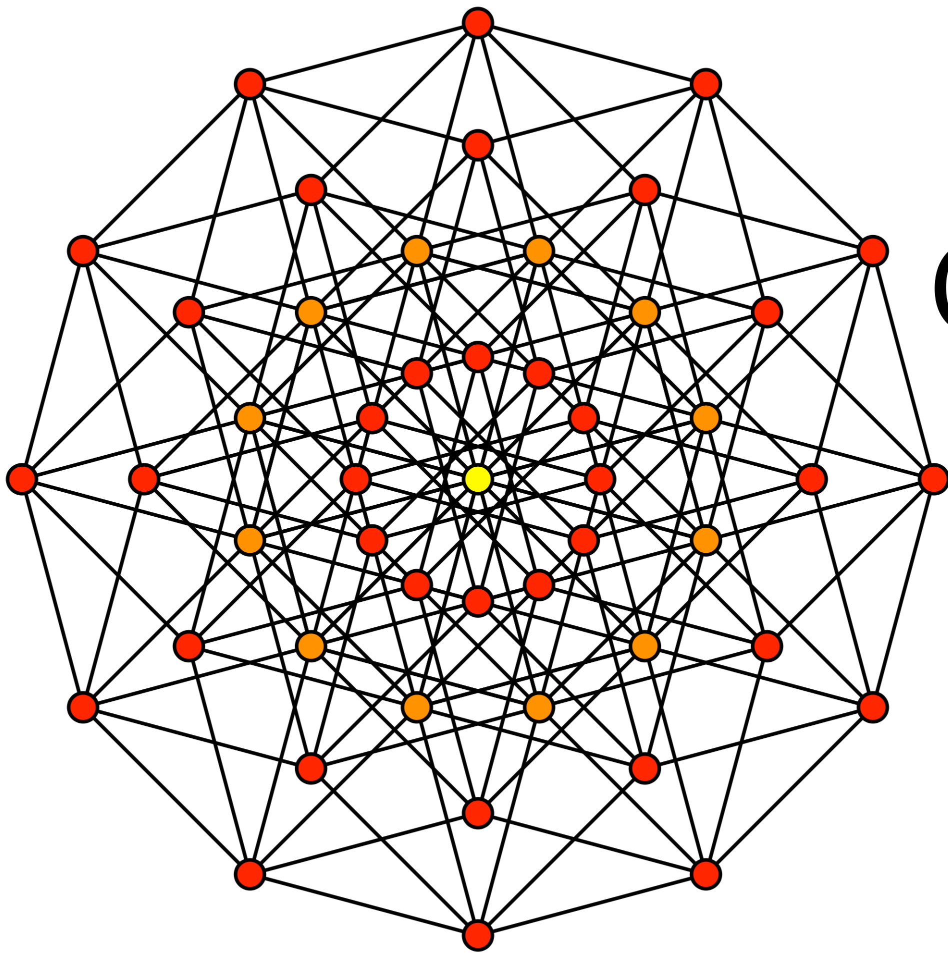
Previous Proofs — Reduction to **Resolution Width** lower bounds algebraic structure **crucial.**

Our Proof — **"Local Gram Schmidt"**

Proof Sketch

Sum of Squares SDP

.....in a slide



CSPs: optimize over $\{0,1\}^n$

solution:

a probability dist. on $\{0,1\}^n$

search over $\exp(n)$ parameters!

Sum of Squares SDP

.....in a slide

CSPs: optimize over $\{0,1\}^n$

solution:

~~a probability dist. on~~ $\cancel{\{0,1\}^n}$ ~~.~~

Idea: search over "pseudo distributions"

satisfy only a subset of constraints of being a distribution!

search over $\exp(n)$ parameters!

Sum of Squares SDP

.....in a slide

CSPs: optimize over $\{0,1\}^n$

solution:

~~a probability dist. on~~ $\cancel{\{0,1\}^n}$~~.~~

Idea: search over "pseudo distributions"

satisfy only a subset of constraints of being a distribution!

~~search over~~ $\cancel{\exp(n)}$ ~~parameters!~~

search over $n^{O(d)}$ parameters.

Sum of Squares SDP

.....in a slide

CSPs: optimize over $\{0,1\}^n$

solution:

~~a probability dist. on $\{0,1\}^n$.~~

Idea: search over “pseudo distributions”

satisfy only a subset of constraints of being a distribution!

~~search over~~

“degree/rounds”

search over $n^{O(d)}$ parameters.

Sum of Squares SDP

.....in a slide

CSPs: optimize over $\{0,1\}^n$

solution:

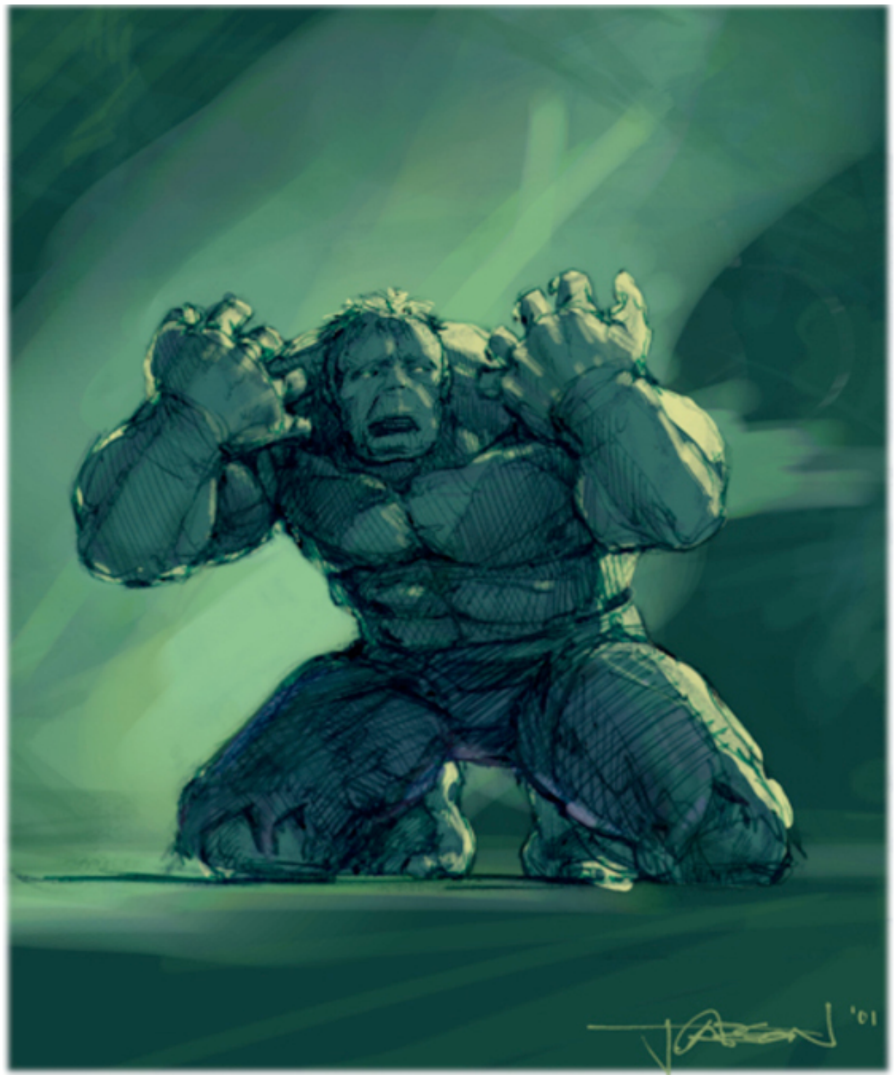
~~a probability dist. on~~ $\cancel{\{0,1\}^n}$~~.~~

Idea: search over "pseudo distributions"

satisfy only a subset of constraints of being a distribution!

= a linear operator **"pseudo-expectation"** on degree d polynomials.

SOS Lower Bound



Show that there is an instance **I** with m constraints:

1) val(**I**) $\sim m \cdot \frac{|P^{-1}(1)|}{2^k}$

2) there is pseudo expectation $\tilde{\mathbb{E}}$ of deg d s.t.

$$\tilde{\mathbb{E}}[\#\text{ I's constraints sat}] = m$$

Deg $d = \varepsilon n$ SOS "thinks" **I** is satisfiable and **I** is far from it!

Proof Outline

Instance **I** such that:

1) val(**I**) $\sim m \cdot \frac{|P^{-1}(1)|}{2^k}$

2) $\tilde{\mathbb{E}}[\#\text{ I's constraints sat}] = m$

3) $\tilde{\mathbb{E}}[p^2] \geq 0\ \forall$ p of deg $\boldsymbol{d} = \varepsilon n$

Proof Outline

Instance **I** such that:

1) val(**I**) $\sim m \cdot \frac{|P^{-1}(1)|}{2^k}$

2) $\tilde{\mathbb{E}}[\#\text{ I's constraints sat}] = m$

3) $\tilde{\mathbb{E}}[p^2] \geq 0\ \forall$ p of deg $\boldsymbol{d} = \varepsilon n$

An actual probability dist. satisfies 3) for polynomials of all degrees up to n.

Proof Outline

Instance **I** such that:

1) val(**I**) $\sim m \cdot \frac{|P^{-1}(1)|}{2^k}$

2) $\tilde{\mathbb{E}}[\#\text{ I's constraints sat}] = m$

3) $\tilde{\mathbb{E}}[p^2] \geq 0 \; \forall$ p of deg $\boldsymbol{d} = \varepsilon n$

I :random, $m = \Theta(n)$, as in **[BGMT09,TW13,...]**

$\tilde{\mathbb{E}}$:same as **[BGMT09,TW13]**

1) and 2) already known to be true!

This work: Showing (3)

Proof Outline

Instance **I** such that:

1) val(**I**) $\sim m \cdot \frac{|P^{-1}(1)|}{2^k}$

2) $\tilde{\mathbb{E}}[\#\text{ I's constraints sat}] = m$

3) $\tilde{\mathbb{E}}[p^2] \geq 0\ \forall$ p of deg $\boldsymbol{d} = \varepsilon n$

I :random, $m = \Theta(n)$, as in **[BGMT09,TW13,...]**

$\tilde{\mathbb{E}}$:same as **[BGMT09,TW13]**

Find $\chi_1, \chi_2, \ldots \chi_N$ **such that:**

1) their span is degree d polys.

2) $\tilde{\mathbb{E}}[\chi_i^2] \geq 0\ \forall i$ **3)** $\tilde{\mathbb{E}}[\chi_i \cdot \chi_j] = 0\ \forall i \neq j$

Proof Outline

Instance **I** such that:

1) val(**I**) $\sim m \cdot \frac{|P^{-1}(1)|}{2^k}$

2) $\tilde{\mathbb{E}}[\#\text{ I's constraints sat}] = m$

3) $\tilde{\mathbb{E}}[p^2] \geq 0 \ \forall$ p of deg $\boldsymbol{d} = \varepsilon n$

I :random, $m = \Theta(n)$, as in **[BGMT09,TW13,...]**

$\tilde{\mathbb{E}}$:same as **[BGMT09,TW13]**

Find $\chi_1, \chi_2, \ldots \chi_N$ **such that:**

1) their span is degree d polys.

2) $\tilde{\mathbb{E}}[\chi_i^2] \geq 0 \ \forall i$ **3)** $\tilde{\mathbb{E}}[\chi_i \cdot \chi_j] = 0 \ \forall i \neq j$

$$\tilde{\mathbb{E}}[p^2] = \tilde{\mathbb{E}}[(\sum_{i=1}^{N} p_i \chi_i)^2] = \sum_{i=1}^{n} p_i^2 \tilde{\mathbb{E}}[\chi_i^2] \geq 0$$

Proof Outline

Find $\chi_1, \chi_2, \ldots \chi_N$ **such that:**

1) their span is degree d polys.

2) $\tilde{\mathbb{E}}[\chi_i^2] \geq 0 \ \forall i$ **3)** $\tilde{\mathbb{E}}[\chi_i \cdot \chi_j] = 0 \ \forall i \neq j$

1) start from the Fourier characters in some order.

2) run Gram-Schmidt using $\tilde{\mathbb{E}}$ inner product.

Proof Outline

Find $\chi_1, \chi_2, \ldots \chi_N$ **such that:**

1) their span is degree d polys.

2) $\tilde{\mathbb{E}}[\chi_i^2] \geq 0 \ \forall i$ **3)** $\tilde{\mathbb{E}}[\chi_i \cdot \chi_j] = 0 \ \forall i \neq j$

1) start from the Fourier characters in some order.

2) run Gram-Schmidt using $\tilde{\mathbb{E}}$ inner product.

[BGMT09]: $\tilde{\mathbb{E}}$ is valid inner product restricted to any d coordinates.

"corresponds to an **actual** expectation associated with a local prob. dist."

Proof Outline

Find $\chi_1, \chi_2, \dots \chi_N$ **such that:**

1) their span is degree d polys.

2) $\tilde{\mathbb{E}}[\chi_i^2] \geq 0 \ \forall i$ **3)** $\tilde{\mathbb{E}}[\chi_i \cdot \chi_j] = 0 \ \forall i \neq j$

1) start from the Fourier characters in some order.

2

Gram Schmidt → **long range dependence!**

is valid inner product restricted to any d coordinates.

Circular argument!

Proof Outline

“expansion in the hyper graph I”

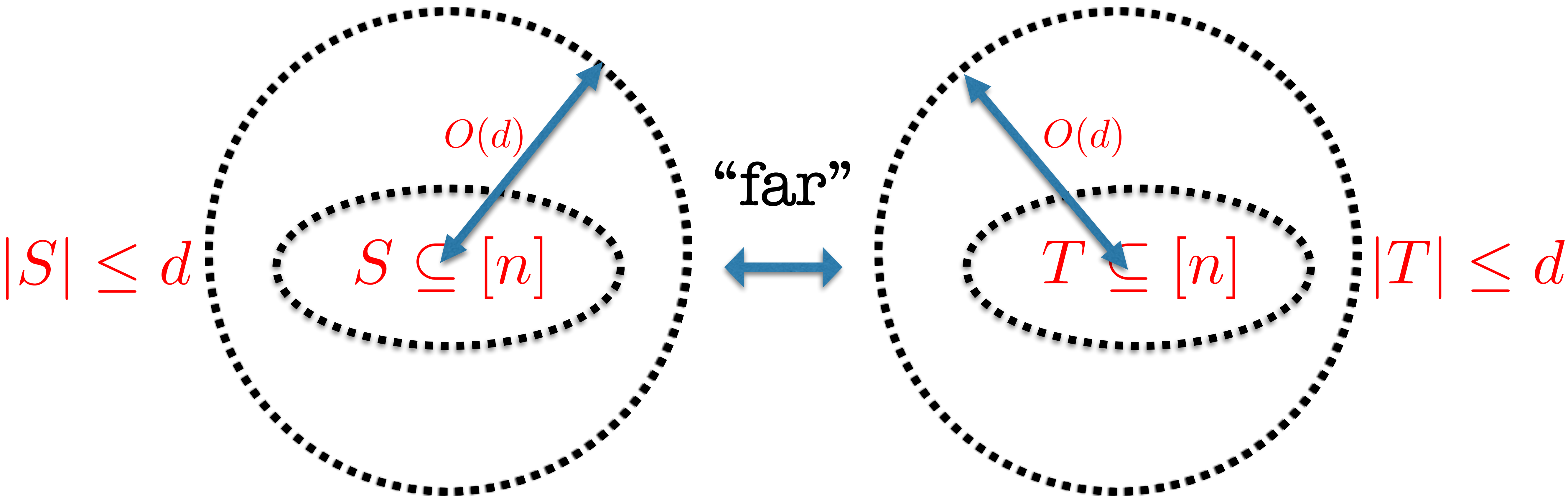
(used in all previous works)

All subsets C of $\sim \varepsilon n$ constraints cover

$(k-1-o(1))|C|$ variables.

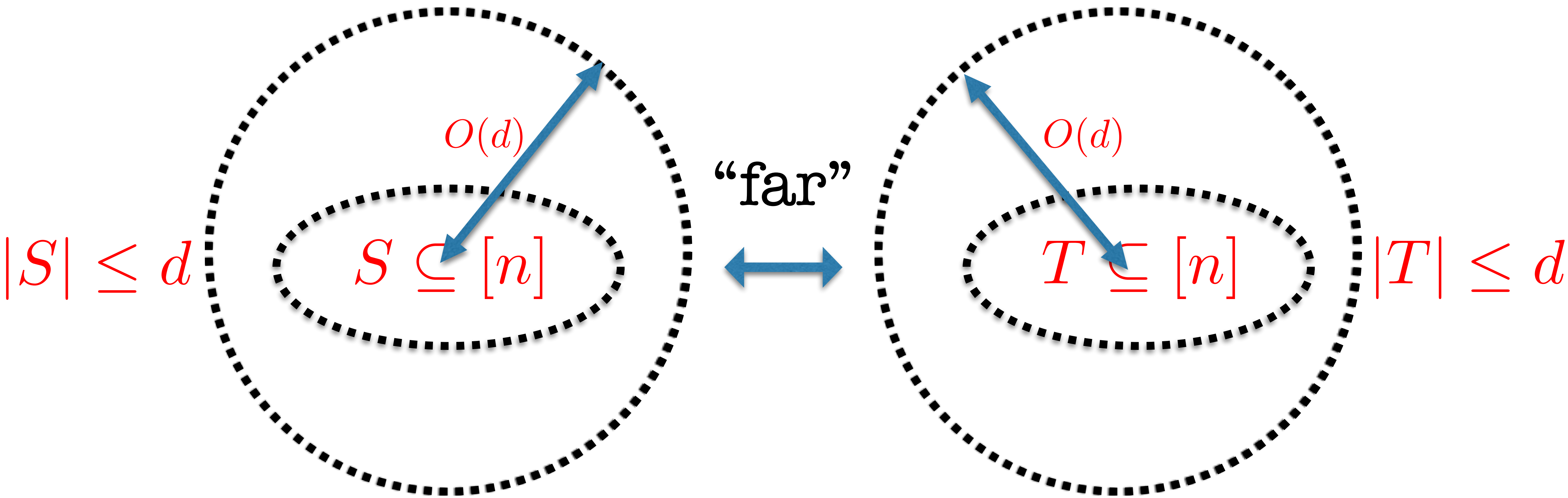
Proof Outline

dist measure based on the hypergraph defining **I**.



Proof Outline

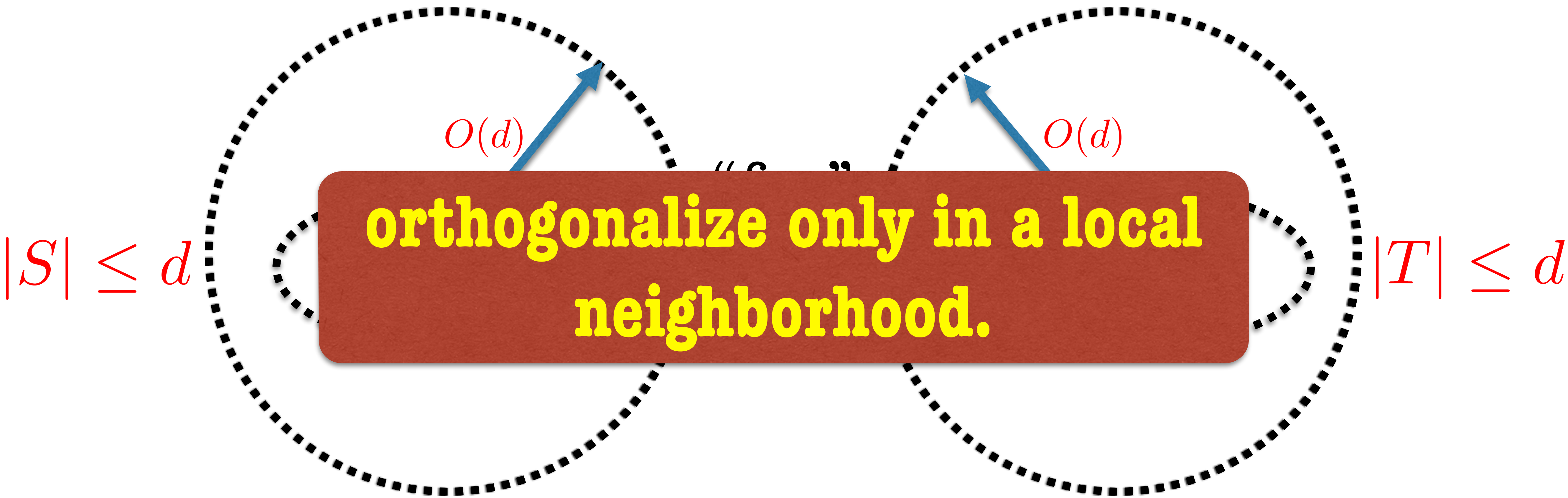
dist measure based on the hypergraph defining $\mathbf{I}$.



S, T “far enough” $\Rightarrow$ **local distributions on S and T given by $\tilde{\mathbb{E}}$ are “independent”.**

Proof Outline

dist measure based on the hypergraph defining **I**.



S, T
“far enough”

local distributions on S and T given by $\tilde{\mathbb{E}}$ **are “independent”.**

Proof Outline

Find $\chi_1, \chi_2, \dots \chi_N$ **such that:**

1) their span is degree d polys.

2) $\tilde{\mathbb{E}}[\chi_i^2] \geq 0 \ \forall i$ **3)** $\tilde{\mathbb{E}}[\chi_i \cdot \chi_j] = 0 \ \forall i \neq j$

1) start :Fourier characters in a **carefully chosen** order.

Proof Outline

Find $\chi_1, \chi_2, \ldots \chi_N$ **such that:**

1) their span is degree d polys.

2) $\tilde{\mathbb{E}}[\chi_i^2] \geq 0 \ \forall i$ **3)** $\tilde{\mathbb{E}}[\chi_i \cdot \chi_j] = 0 \ \forall i \neq j$

1) start :Fourier characters in a **carefully chosen** order.

2) run a step of Gram-Schmidt using $\tilde{\mathbb{E}}$ inner product

....but only subtract components on characters that belong the local neighborhood.

Proof Outline

Find $\chi_1, \chi_2, \dots \chi_N$ **such that:**

1) their span is degree d polys.

2) $\tilde{\mathbb{E}}[\chi_i^2] \geq 0 \; \forall i$ **3)** $\tilde{\mathbb{E}}[\chi_i \cdot \chi_j] = 0 \; \forall i \neq j$

1) start :Fourier characters in a **carefully chosen** order.

2) run a step of Gram-Schmidt using $\tilde{\mathbb{E}}$ inner product

....but only subtract components on characters that belong the local neighborhood.

3) Hope for the best.

Proof Outline

Find $\chi_1, \chi_2, \dots \chi_N$ **such that:**

1) their span is degree d polys.

2) $\tilde{\mathbb{E}}[\chi_i^2] \geq 0 \ \forall i$ **3)** $\tilde{\mathbb{E}}[\chi_i \cdot \chi_j] = 0 \ \forall i \neq j$

1) start :Fourier characters in a **carefully chosen** order.

2) run a step of Gram-Schmidt using $\tilde{\mathbb{E}}$ inner product

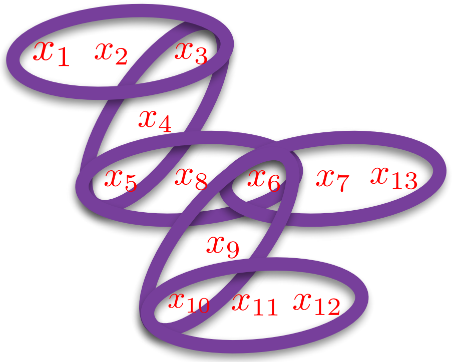
....but only subtract components on characters that belong the local neighborhood.

3) every little thing is gonna be alright!

Summary

?

vs



SOS needs exponential time to beat random for pairwise independent CSPs.

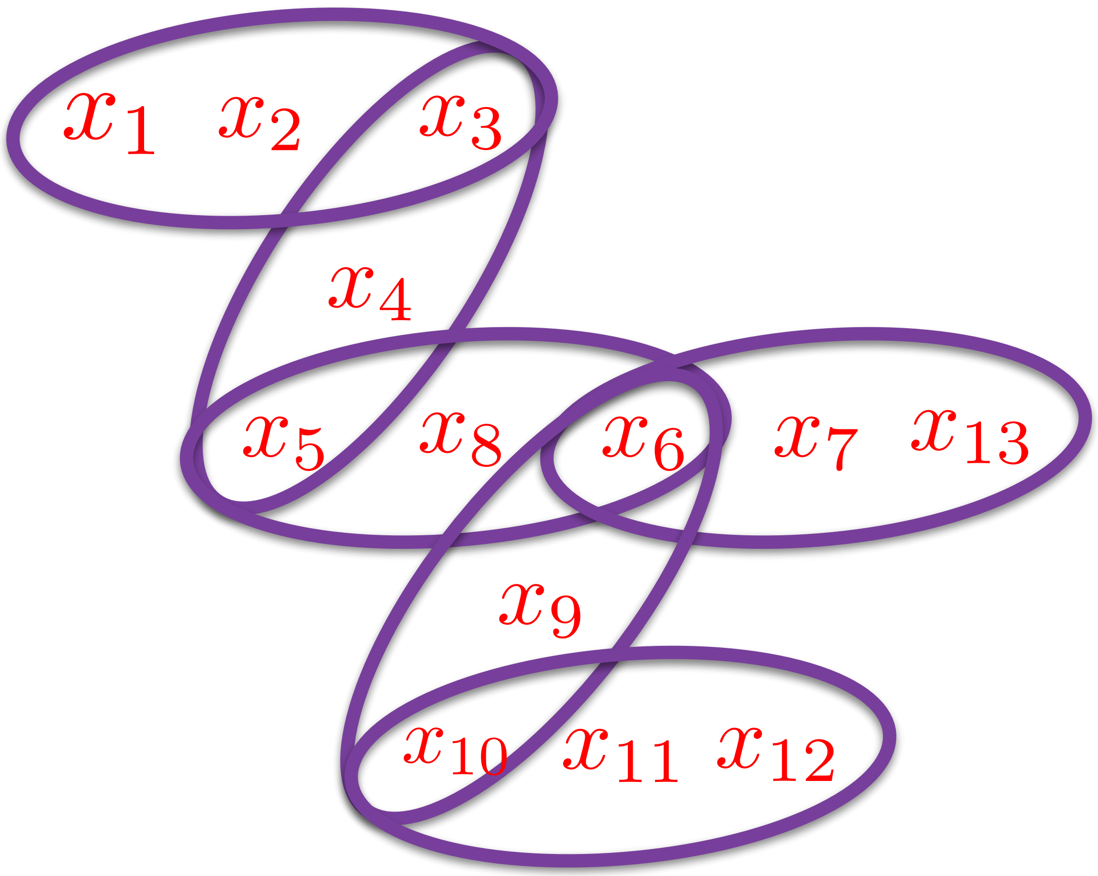
Summary

SOS needs exponential time to beat random for pairwise independent CSPs.

Open: NP hardness?

Questions?

Pairwise Independence



A CSP(P) is **pairwise independent** if

there is a prob. dist. μ

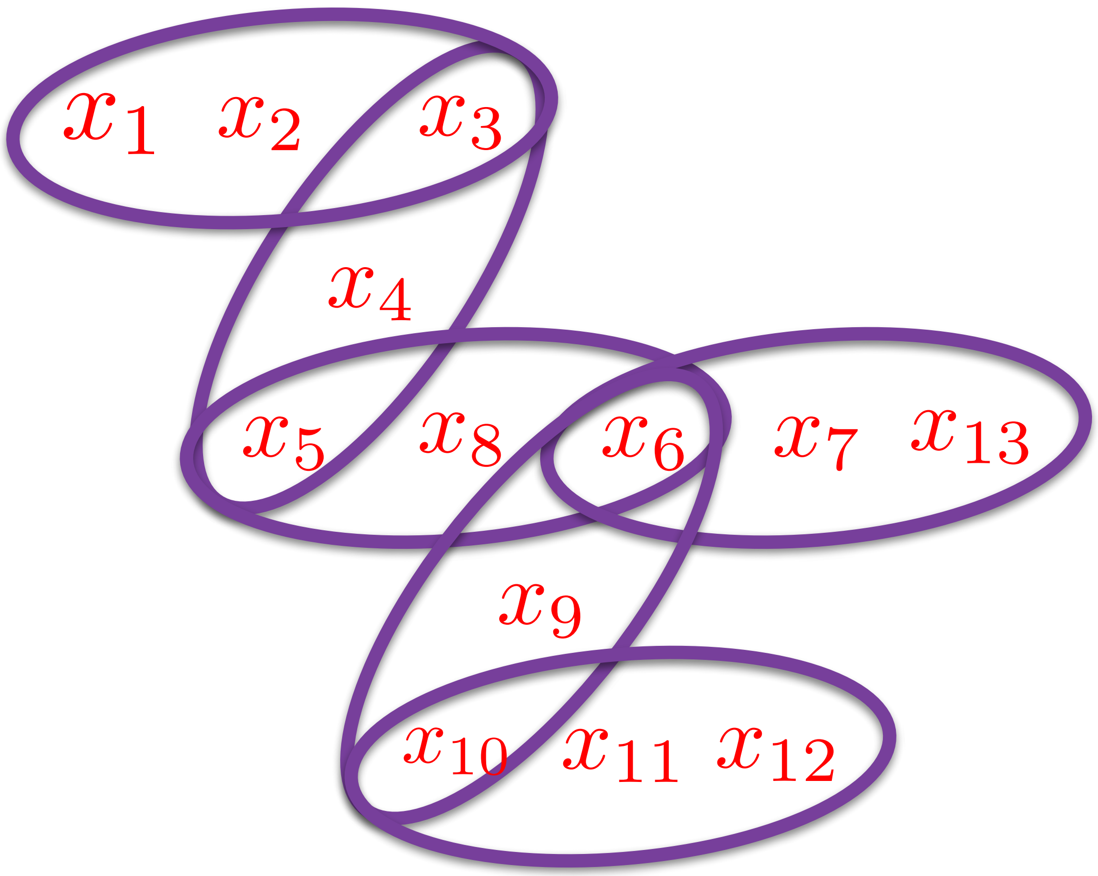
1) pairwise independent.

2) supported on $P^{-1}(1)$

[AH09]

random predicates with $\Omega(k^2)$ satisfying assignments are pairwise independent.

Pairwise Independence



A CSP(P) is **pairwise independent** if

there is a prob. dist. μ

1) pairwise independent.

2) supported on $P^{-1}(1)$

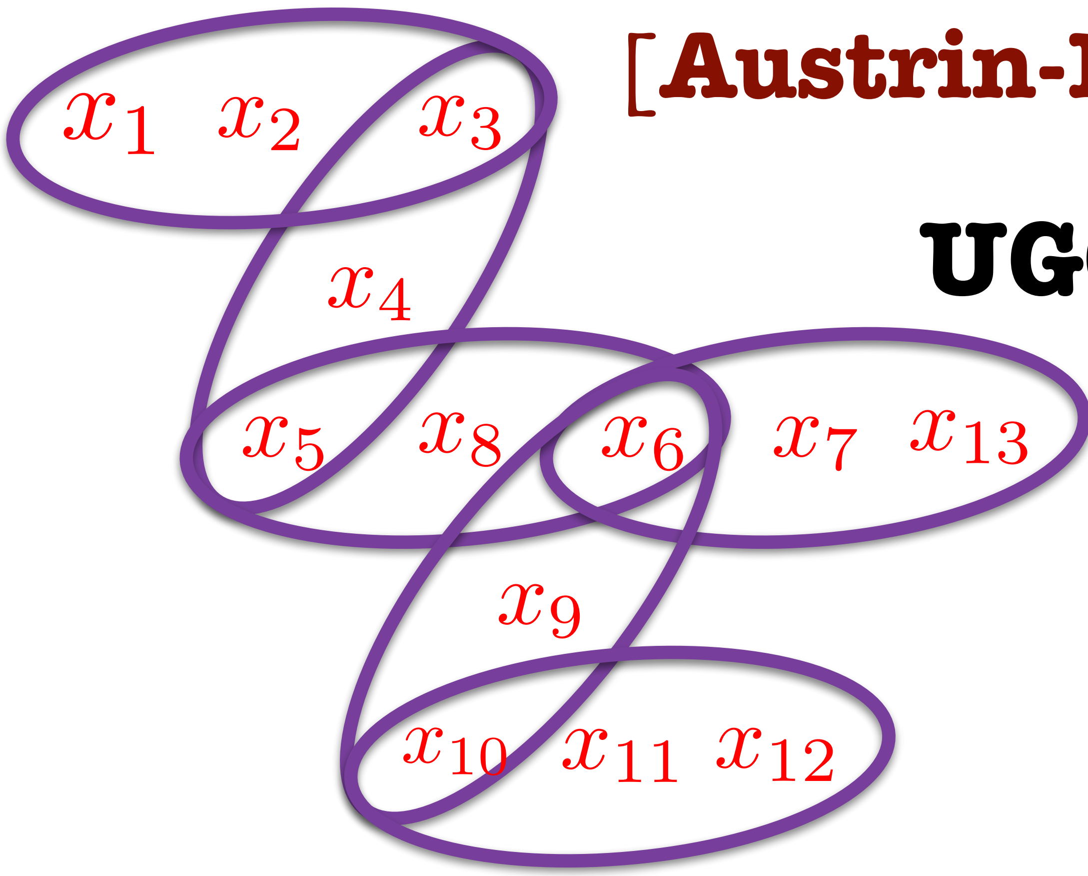
[AH09] sparse p.i. preds are abundant.

[AM09]

UGC → pairwise independent CSPs are approximation resistant*.

*can't beat random assignment.

Pairwise Independence



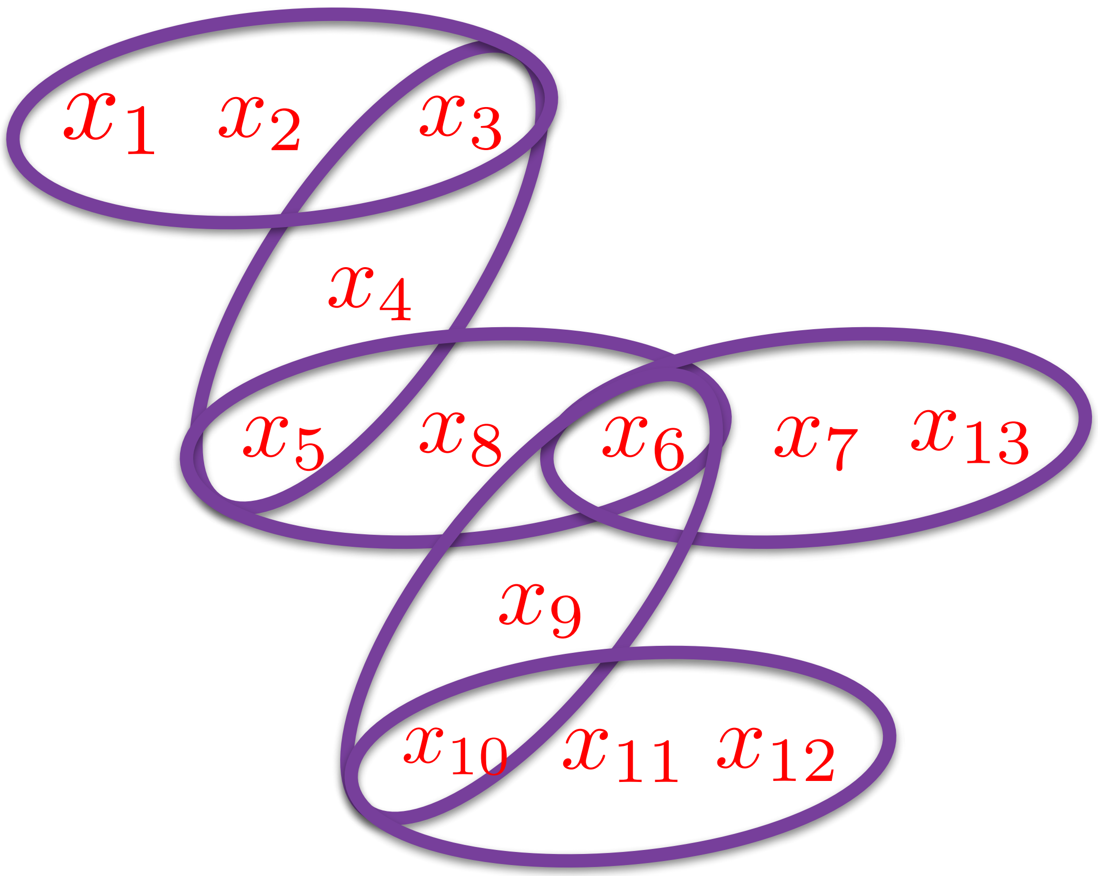
[**Austrin-Mossel09**]

UGC → p.i. CSPs are hard.

[Raghavendra08] For CSPs

UGC Hardness ↔ integrality gaps for “basic” SDP

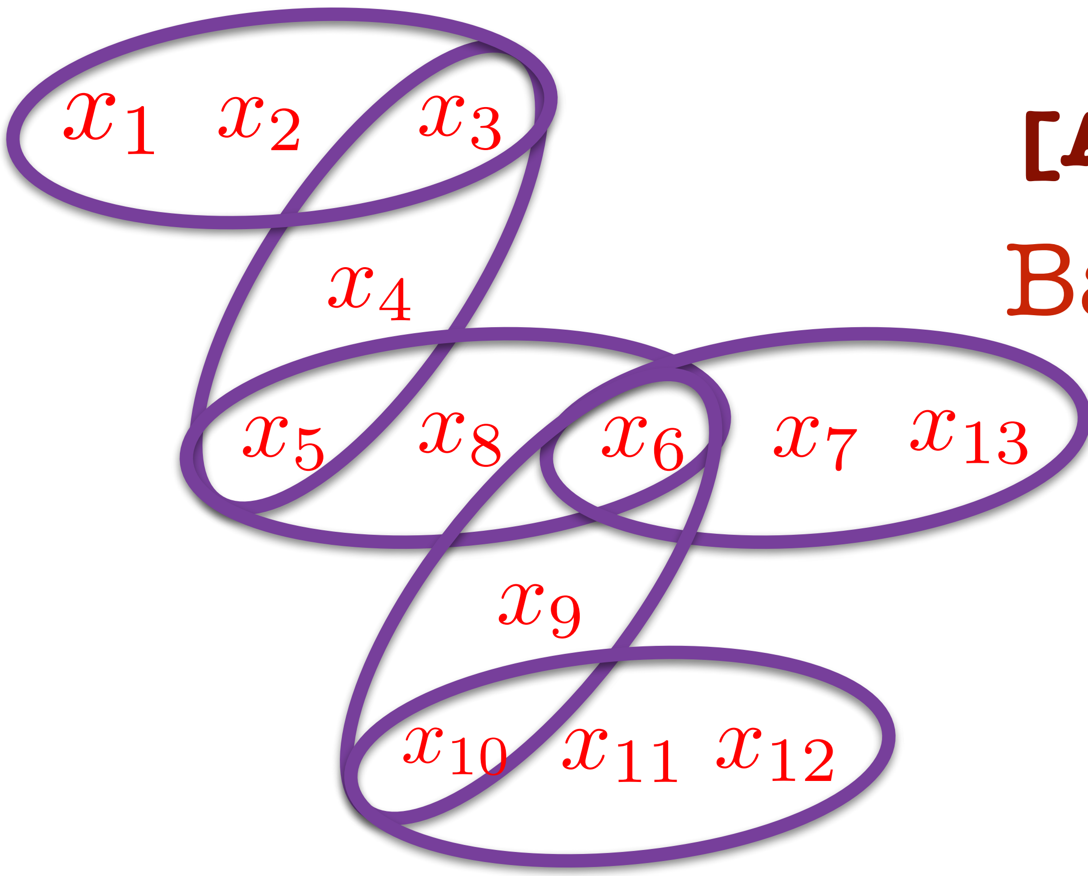
Pairwise Independence



[AM09]

Basic SDP can't beat random assignment for p.i. CSPs.

Pairwise Independence



[AM09]

Basic SDP doesn't beat random assignment for p.i. CSPs.

[BGMT09] **[TW13]**

Sherali Adams/Lovasz-Schrjiver SDPs of **exponential** size do not beat random assignment for p.i. CSPs.

Sum of Squares SDP

.....in a slide

Pseudoexpectations

A linear operator $\tilde{\mathbb{E}}$ on degree d polynomials on $\{0,1\}^n$ satisfying:

$$\tilde{\mathbb{E}}[p^2] \geq 0$$

for every degree d/2 polynomial p.

Sum of Squares SDP

.....in a slide

Pseudo expectations

A linear operator $\tilde{\mathbb{E}}$ on degree d polynomials on $\{0, 1\}^n$ satisfying:

$\tilde{\mathbb{E}}[p^2$

for every degree d/2 p

degree k instance/objective polynomial

Value of CSP(P) at an assignment x $\leftrightarrow$ **I(x)**