Refuting CSPs: Even Covers to Eigenvalues to SDPs and Back

Pravesh Kothari
CMU

Jackson Abascal
CMU

Venkat Guruswami
CMU/Berkeley

Peter Manohar
CMU

Refuting CSPs

Refutation Algorithm:

Input: An instance $\boldsymbol{\phi}$ of k-SAT with **m** clauses on **n** variables.

Output: A value $\boldsymbol{v} \in [\mathbf{0}, \mathbf{1}]$.

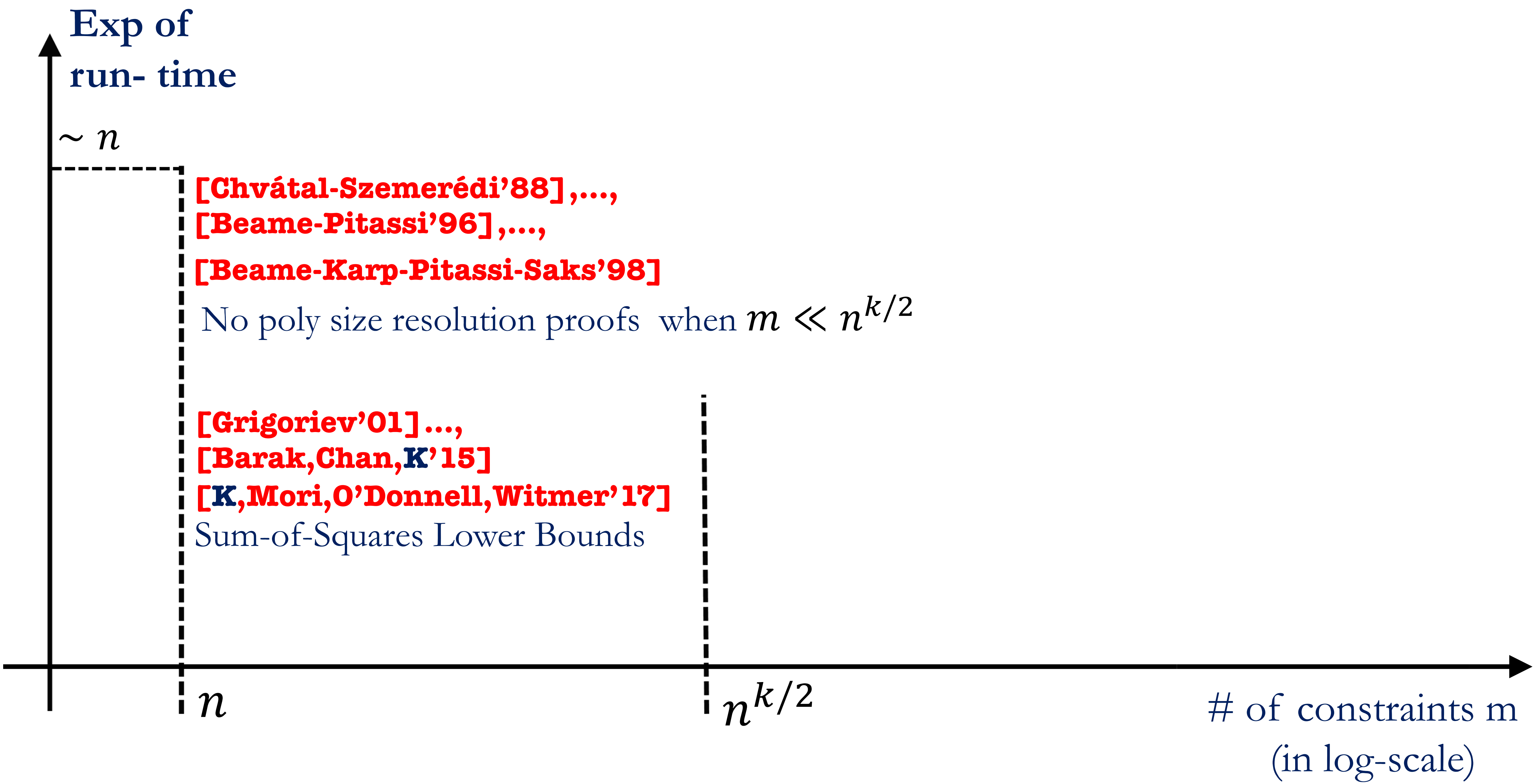
Correctness: $\boldsymbol{val(\phi) \leq v}$. "$val(\phi)$ = max frac of constraints satisfiable"

The algorithm *weakly refutes* a formula ϕ if $v < 1$.
strongly refutes …. if $v < 1 - \delta$ $\delta > 0$, abs. const.

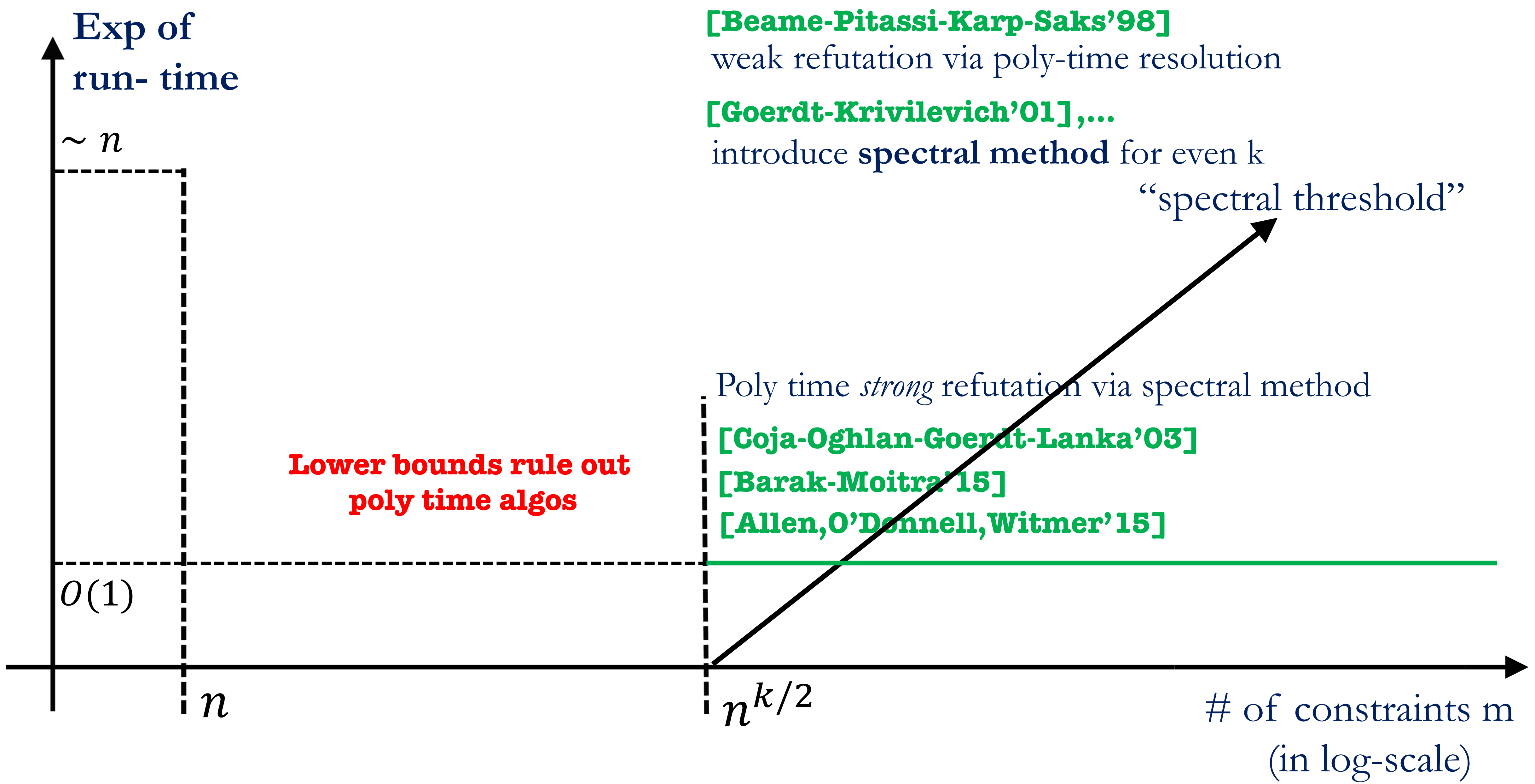
Goal: refute largest possible family of instances ϕ: $val(\phi) < 0.99$.

Random: clause structure (instance hypergraph), literal patterns uniform random
Semi-random: clause structure worst-case, literal patterns uniform random
Smoothed: worst-case instance → perturb each literal w.p. 0.01

The story of random k-SAT



The story of random k-SAT



Strongly refuting *random* 4-XOR

Over $x \in \{\pm 1\}^n$, **4**-XOR constraints are of the form: $\{\, x_1 x_2 x_3 x_4 = \pm 1, \ldots \}$

Instance: A 4-uniform hypergraph $\mathcal{H}$ and a set of "RHS" b_C for each $C \in \mathcal{H}$.

$$\phi(x) = \frac{1}{m} \sum_{C \in \mathcal{H}} b_C x_{C_1} x_{C_2} x_{C_3} x_{C_4} = \frac{1}{m} \sum_{C \in \mathcal{H}} b_C x_C$$

…is a deg 4 polynomial that computes "advantage over ½" of assignment x.

Goal: Certify that $\phi(x) \leq \epsilon$ for all $x \in \{\pm 1\}^n$

Strongly refuting *random* 4-XOR

Goal: Certify that $\phi(x) = \frac{1}{m}\sum_{C\in\mathcal{H}} b_C x_C \leq \epsilon$ for all $x \in \{\pm 1\}^n$

Idea: write $\phi(x)$ as the quadratic form of some matrix! [Goerdt,Krivilevich'01…]

$$A = \begin{array}{cc} & \{k,\ell\} \\ \{i,j\} & \begin{bmatrix} & & \\ & b_{\{i,j,k,\ell\}} & \\ & & \end{bmatrix} \end{array}$$

Then, $\phi(x) = \frac{1}{6}\left(x^{\odot 2}\right)^{\top} \mathrm{A}\left(x^{\odot 2}\right)$.

$\leq \frac{1}{6}\left\|\left(x^{\odot 2}\right)\right\|_2^2 \left\|\mathrm{A}\right\|_2$.

Analysis: Succeeds in refuting if $m \gtrsim n^2$.

Matrix Chernoff, trace method,…all work easily to bound $\|A\|_2$

Strongly refuting *random* 3-XOR

Goal: Certify that $\phi(x) = \frac{1}{m} \sum_{C \in \mathcal{H}} b_C x_C \leq \epsilon$ for all $x \in \{\pm 1\}^n$

Which matrix?

"rectangulum non quadratum"

-Marcus Aurelius, 150 AD **maybe**

Strongly refuting *random* 3-XOR

Goal: Certify that $\phi(x) = \frac{1}{m}\sum_{C\in\mathcal{H}} b_C x_C \le \epsilon$ for all $x \in \{\pm 1\}^n$

Which matrix? the n^2 by n^2 matrix of the "Cauchy-Schwarzed" instance…

Entries no longer independent ($\sim n^2$ non-zeros, $\sim n^{1.5}$ indep bits…)

Analyze spectral norm using the trace moment method,
Needs randomness in both the clauses and the RHS

the *trace moment* method

Strongly refuting *semirandom* k-XOR

[Fei'07] *Weak* ref for *smoothed* 3-SAT with $\tilde{O}(n^{1.5})$ clauses.

- Relies on spectral methods and even covers
- Extends to 3-CSPs but not to strong ref or >3-CSPs.

[Abascal-Guruswami-K'21]

Poly time strong refutation of semirandom k-XOR with $m \sim n^{\frac{k}{2}}$ constraints.

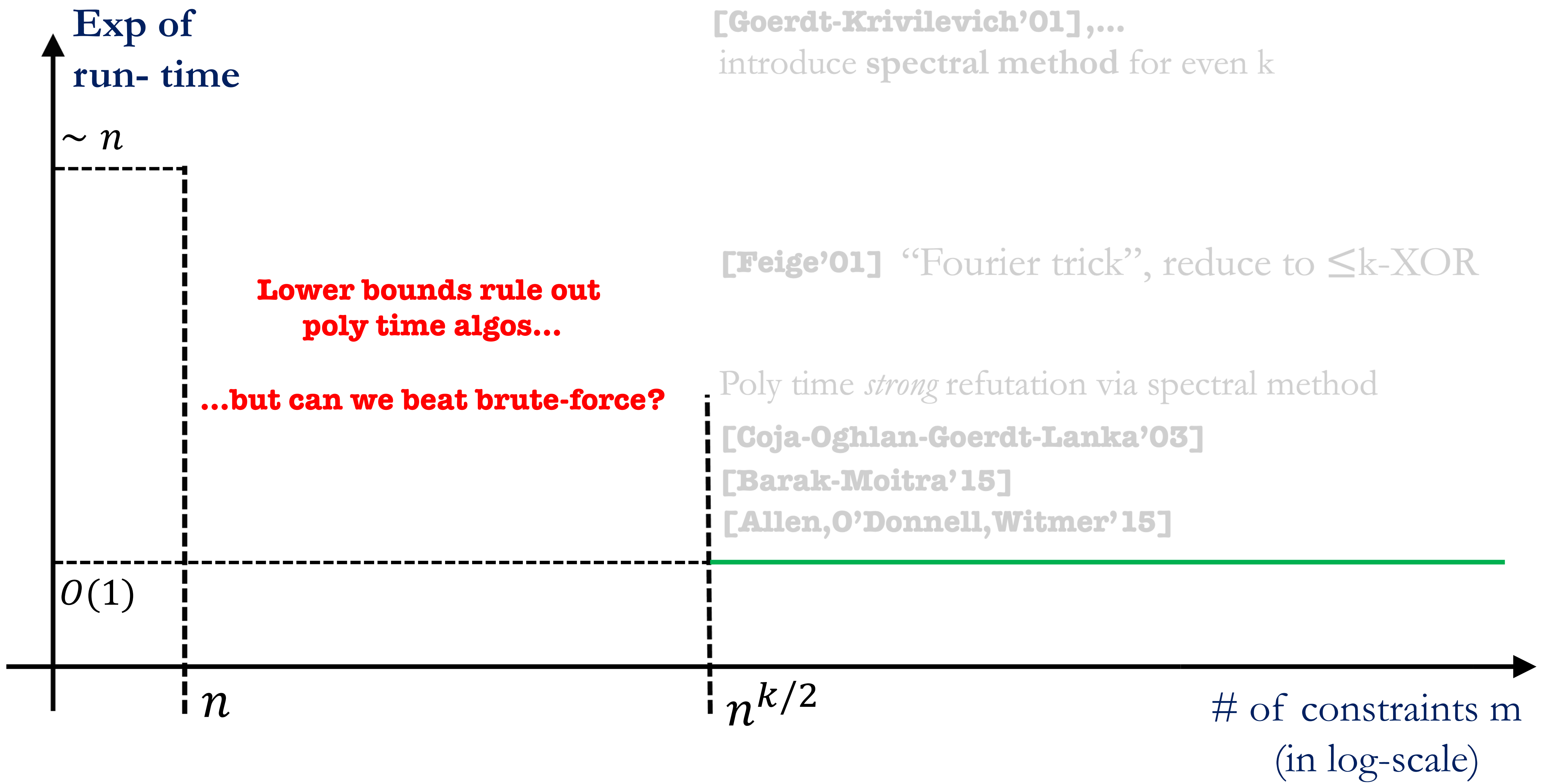
A new analysis for the *random* case that only needs standard matrix conc.

Extend via SDP/SoS Proofs to semirandom setting with:

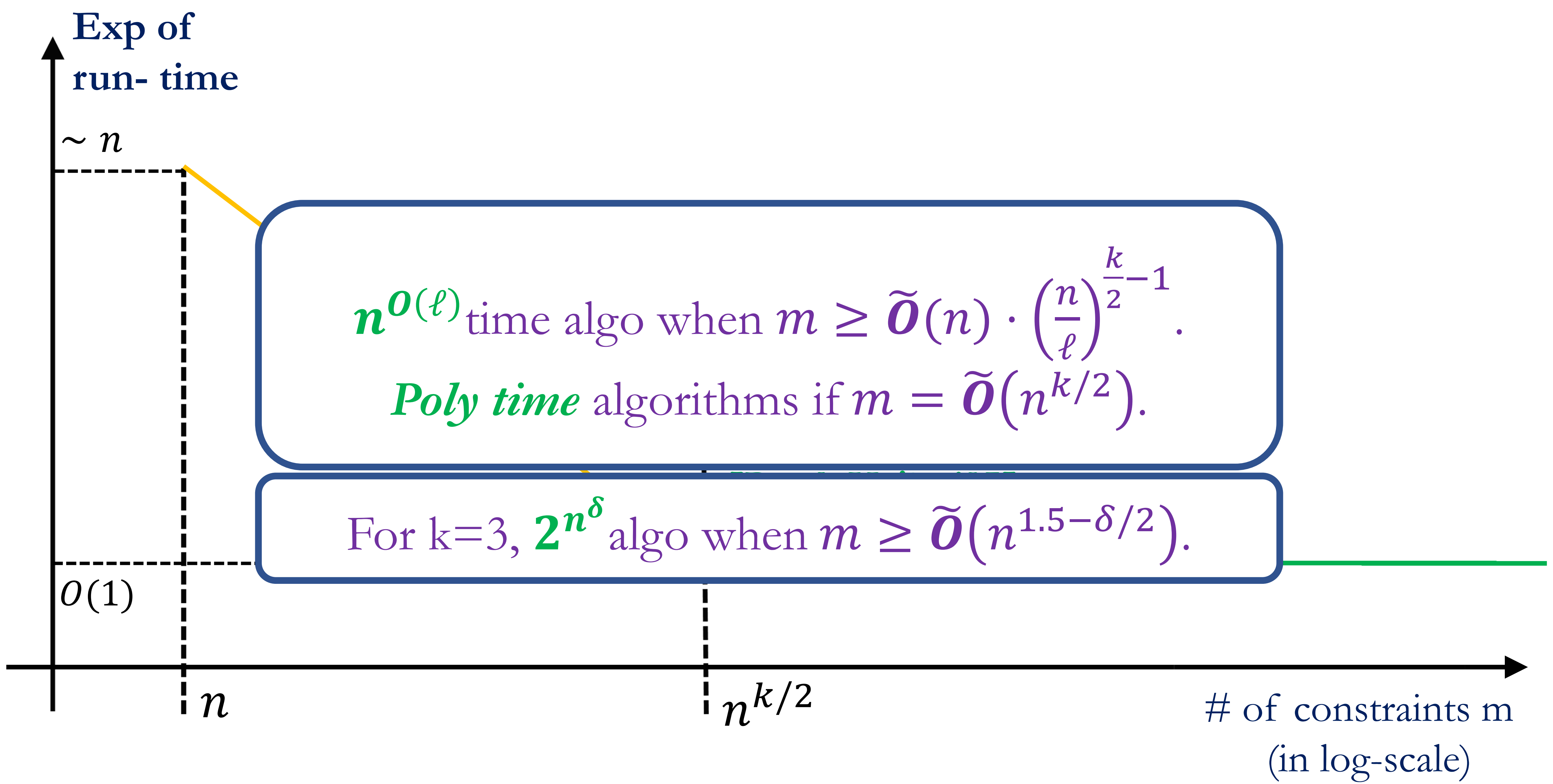
instance decomposition

+ *row bucketing*

The story of random k-SAT



The story of random k-SAT



$\boldsymbol{n^{O(\ell)}}$ time algo when $m \geq \tilde{\boldsymbol{O}}(n) \cdot \left(\frac{n}{\ell}\right)^{\frac{k}{2}-1}$.

Poly time algorithms if $m = \tilde{\boldsymbol{O}}\left(n^{k/2}\right)$.

For k=3, $\boldsymbol{2^{n^{\delta}}}$ algo when $m \geq \tilde{\boldsymbol{O}}\left(n^{1.5-\delta/2}\right)$.

Strongly refuting *random* 4-XOR

Goal: Certify that $\phi(x) = \frac{1}{m}\sum_{C\in\mathcal{H}} b_C x_C \le \epsilon$ for all $x \in \{\pm 1\}^n$

Full trade-off for 4-XOR? $n^{O(\ell)}$ time vs $m \sim \frac{n^2}{\ell}$ constraints.

[RRS'16] use a "symmetrized tensor power matrix" who quad. form is $\phi(x)^{2\ell}$

Issue: Fairly technical application of the trace method
Crucially uses randomness of $\mathcal{H}$.

Two recent papers [Ahn'19,Wein-Alaoui-Moore'19] succeed in simplifying for *even k*.

[Wein-Alaoui-Moore'19] Introduce ***Kikuchi* matrix** and significantly simplify **even-arity random** k-XOR refutation.

Doesn't work for odd random k-XOR

Let's see their neat idea!

Strongly refuting *random* 4-XOR

Goal: Certify that $\phi(x) = \frac{1}{m}\sum_{C\in\mathcal{H}} b_C x_C \le \epsilon$ for all $x \in \{\pm 1\}^n$

Idea: write $\phi(x)$ as the quadratic form of a $\binom{n}{\ell}\times\binom{n}{\ell}$ matrix.

$$\binom{[n]}{\ell} \ni S, T$$

$$A_C = \begin{array}{cc} & T \\ S & \begin{pmatrix} & & \\ & \begin{cases} b_C & \text{if } S\Delta T = C \\ 0 & \text{otherwise} \end{cases} & \\ & & \end{pmatrix} \end{array} \qquad A = \sum_{C\in\mathcal{H}} A_C$$

Then, $\phi(x) = \frac{1}{D_\ell}\left(x^{\odot \ell}\right)^{\top} \mathrm{A}\left(x^{\odot \ell}\right) = \frac{1}{D_\ell}\sum_{S,T} A(S,T)x_S x_T$

$$= \frac{1}{D_\ell}\sum_{S,T} A(S,T)x_{S\Delta T} \le \frac{1}{D_\ell}\binom{n}{\ell}\|\mathrm{A}\|_2$$

Strongly refuting *random* 4-XOR

Goal: Certify that $\phi(x) = \frac{1}{m}\sum_{C\in\mathcal{H}} b_C x_C \le \epsilon$ for all $x \in \{\pm 1\}^n$

Idea: write $\phi(x)$ as the quadratic form of a $\binom{n}{\ell}\times\binom{n}{\ell}$ matrix.

$$\binom{[n]}{\ell} \ni S,\ T$$

$$A_C = \begin{array}{cc} & T \\ S & \left(\begin{cases} b_C & \text{if } S\Delta T = C \\ 0 & \text{otherwise} \end{cases}\right) \end{array}$$

$$A = \sum_{C\in\mathcal{H}} A_C$$

Then, $\phi(x) = \frac{1}{D_\ell}\left(x^{\odot \ell}\right)^{\top} \mathrm{A} \left(x^{\odot \ell}\right)$
$\le \frac{1}{D_\ell}\binom{n}{\ell} \|\mathrm{A}\|_2.$

Analysis: How can we bound $\|A\|_2$?

Strongly refuting *random* 4-XOR

How can we bound $||A||_2$?

$$A = \sum_{C \in \mathcal{H}} A_C$$

$$A = \begin{array}{cc} & T \\ S & \begin{pmatrix} & & \\ & & \\ & & \end{pmatrix} \end{array}$$

independent, random matrices.

Analysis: Apply matrix Chernoff inequality.

Succeeds in refuting if $m \gtrsim \frac{n^2}{\ell}$.

Strongly refuting *semirandom* k-XOR?

[Guruswami-K-Manohar'21]

$n^{O(\ell)}$ time strong ref. of smoothed k-XOR with $m \sim n \left(\frac{n}{\ell}\right)^{\frac{k}{2}-1}$ constraints.

A new analysis via *pruned-spectral-norm* of a Kikuchi-like matrix for random…

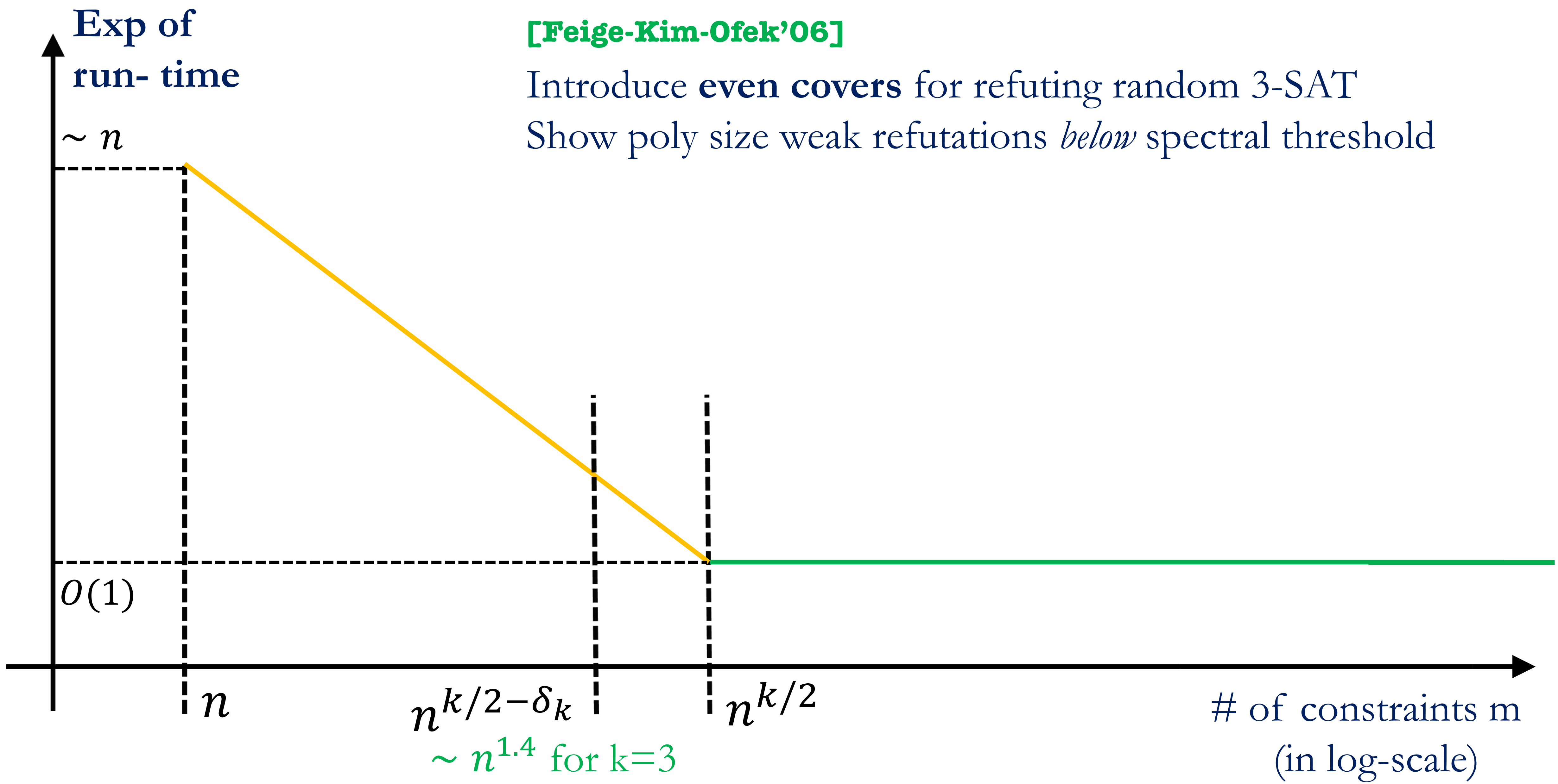
Extend via SDP/SoS Proofs to semirandom and smoothed via:
A new *hypergraph regularity* decomposition lemma
+ *pruned-spectral-norm* (via conc of polynomials with combinatorial structure)
+ *row bucketing* + *matrix conc.*

The story of random k-SAT

[Feige-Kim-Ofek'06]

Introduce **even covers** for refuting random 3-SAT

Show poly size weak refutations *below* spectral threshold

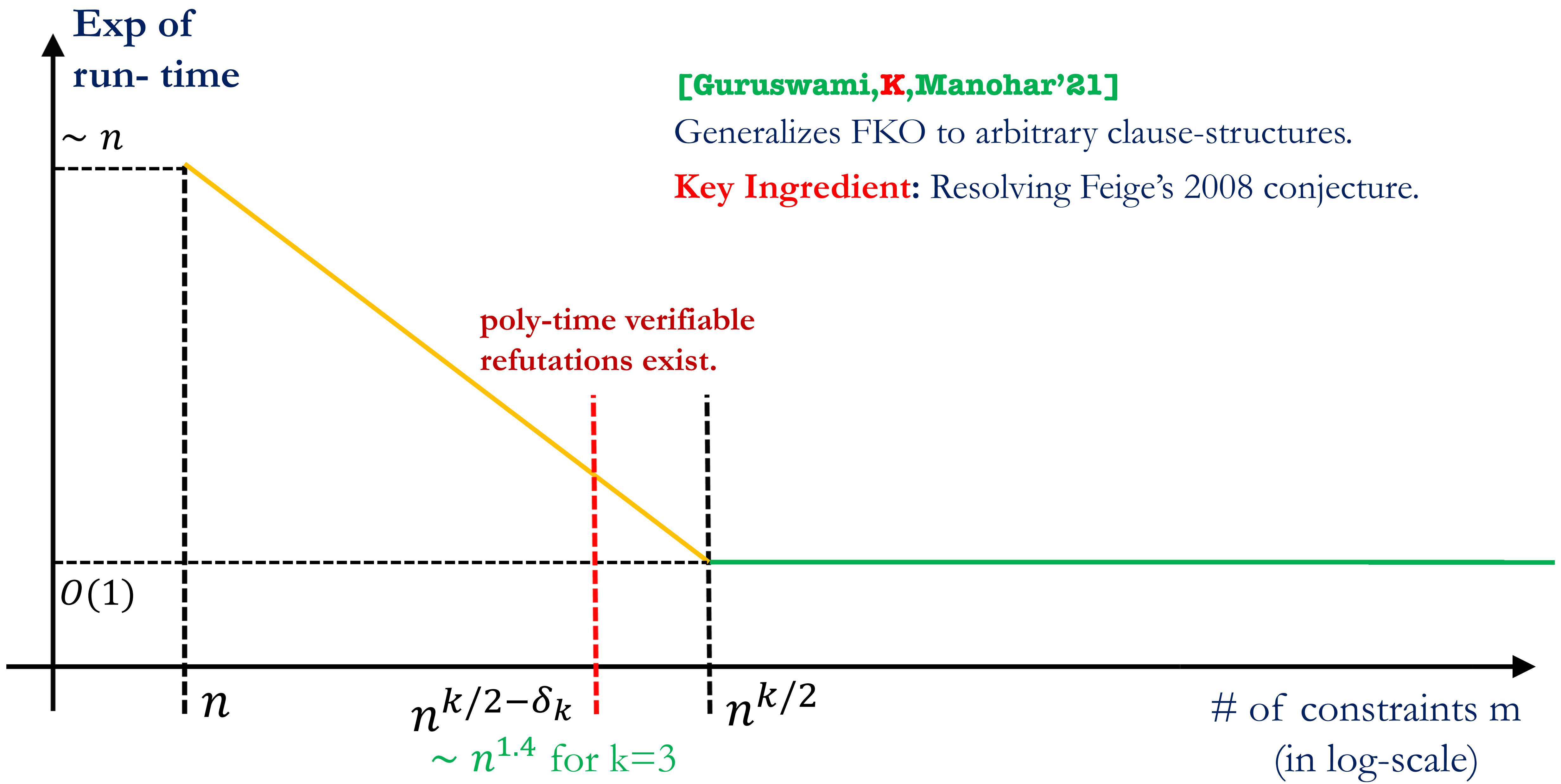


The story of smoothed k-SAT

[Guruswami, K, Manohar'21]

Generalizes FKO to arbitrary clause-structures.

Key Ingredient: Resolving Feige's 2008 conjecture.



Feige's Conjecture

An extremal conjecture about girth of hypergraphs.

Question: What's the maximum girth of a graph on n vertices and $\frac{nd}{2}$ edges?

for d=2: clearly, n (e.g., n-cycle).

for d>2: $\leq 2\log_{d-1} n+2$ **[Alon,Hoory,Linial'02]** **"Moore Bound"**

sharp up to the factor 2 (e.g., some Ramanujan graphs)

Feige's Conjecture

An extremal conjecture about girth of hypergraphs.

Moore bound: max girth of a graph on n vertices and $\frac{nd}{2}$ edges is $\sim 2\log_{d-1} n$

What about 3 (and more generally, k)-uniform hypergraphs?

A cycle is a subgraph that touches every vertex an even # of times.

Hypergraph Cycles (Even Covers)

A **hypergraph cycle** = set of hyperedges touching each vertex an. even # of times.

= size of a smallest *linearly-dependent subset* of *k-sparse* linear equations *mod 2*.

Feige's Conjecture

An extremal conjecture about girth of hypergraphs.

Moore bound: max girth of a graph on n vertices and $\frac{nd}{2}$ edges is $\sim 2\log_{d-1} n$

Hypergraph Cycles (a.k.a. even covers)

A **hypergraph cycle** = set of hyperedges touching each vertex an. even # of times.

Feige's Conjecture (2008):

Every hypergraph with $m \sim n \cdot \left(\frac{n}{\ell}\right)^{\left(\frac{k}{2}-1\right)}$ hyperedges has a cycle of length $\leq \ell \log_2 n$.

= rate-distance trade-off for linear codes with column-sparse parity-check matrices.

for $k = 3$, every hypergraph with $m \sim n\frac{\sqrt{n}}{\sqrt{\ell}}$ has a cycle of length $\leq \ell \log_2 n$.

Random hypergraphs known to achieve it (up to log factor slack in m).

Feige's Conjecture: A brief history

An extremal conjecture about girth of hypergraphs.

Feige's Conjecture (2008):

Every hypergraph with $m \geq n \cdot \left(\frac{n}{\ell}\right)^{(\frac{k}{2}-1)}$ hyperedges has a cycle of length $\leq \ell \log_2 n$.

→ there are $O(\frac{m}{\ell \log_2 n})$ hyperedge-disjoint cycles of length $\leq \ell \log_2 n$.

[Feige,Kim,Ofek'06]:

True for ***random*** k-uniform hypergraphs via a "2^{nd} moment method" argument.

→ Non-trivial weak refutation for random k-XOR.

"non-trivial weak refutation of k-XOR" → weak refutation of k-SAT.

Feige's Conjecture: A brief history

An extremal conjecture about girth of hypergraphs.

Feige's Conjecture (2008):

Every hypergraph with $m \geq n \cdot \left(\frac{n}{\ell}\right)^{\left(\frac{k}{2}-1\right)}$ hyperedges has a cycle of length $\leq \ell \log_2 n$.

[Feige,Kim,Ofek'06]:

True for *random* k-uniform hypergraphs via a "2nd moment method" argument.

[Naor-Verstraete'08],[Feige'08]:

True for all hypergraphs for $\ell = O(1)$ up to a $\log \log n$ factor slack in m.

[Alon,Feige'09]: A suboptimal trade-off for k=3: $m \sim \frac{n^2}{\ell}$ for $\ell \log_2 n$ length cycles.

[Feige,Wagner'16]: A combinatorial approach via sub-hypergraphs of bounded min-degree.

Feige's Conjecture: now a theorem!

An extremal conjecture about girth of hypergraphs.

Feige's Conjecture (2008):

Every hypergraph with $m \geq n \cdot \left(\frac{n}{\ell}\right)^{\left(\frac{k}{2}-1\right)}$ hyperedges has a cycle of length $\leq \ell \log_2 n$.

Theorem [Guruswami, K, Manohar'21]

Feige's conjecture is true **for all k and ℓ** up to a $\log^{2k} n$ factor slack in m

"Spectral double counting"

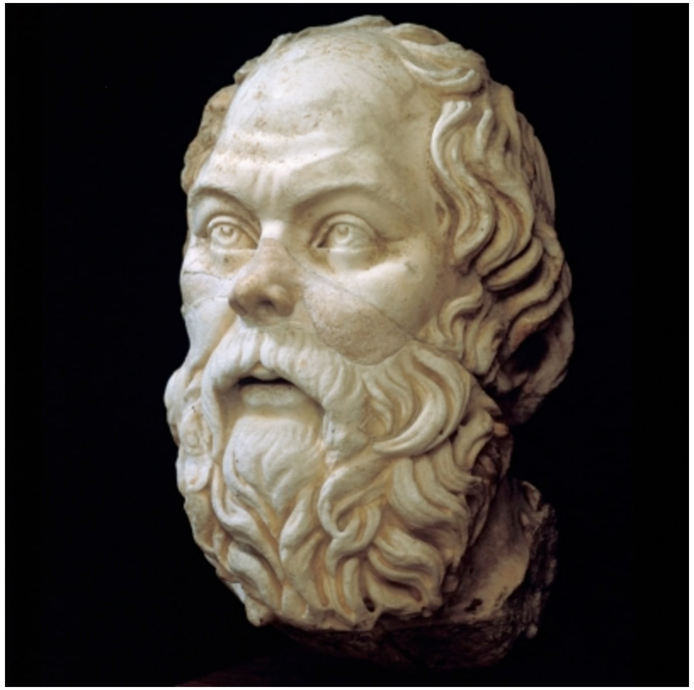
"No small *hypergraph cycle* → no *sub-exp size spectral refutations for* ***semirandom*** *k-XOR."*

Gist: 1. If you randomly perturb each literal independently with small prob, the k-SAT instance becomes **as easy as random** with same # of constraints. For both algorithms, and FKO style certificates.

2. Spectral : Random :: SDP : Semirandom/Smoothed

3. Spectral succeeds only when there are even covers…

4.



"The only true wisdom is knowing that *Kikuchi* matrices are the right object for proving stuff about CSPs…or hypergraphs…or tensors…"

-Socrates, 350 BC **maybe**

Thank you.

Small Cycles via *Spectral Double Counting*

Prop: Whp, random 4-uniform $\mathcal{H}$ with $\sim \frac{n^2}{\ell}$ hyperedges has a $\sim \ell \log_2 n$ length cycle.

Proof Idea:

If not, our refutation algo (with same ℓ) from previous slide works for ***arbitrary*** **RHS $\boldsymbol{b_C}$s**. Since there are satisfiable k-XOR instances $(b_C = 1\ \forall C)$, contradiction.

Key Step:

If there are no cycles of length $\sim \ell \log_2 n$, then regardless of $b_C s$, can prove an **upper bound on $||\boldsymbol{A}||_2$** that matches the one when b_Cs are indep. random.

fixed, deterministic matrix.

Small Cycles via *Spectral Double Counting*

Prop: Whp, random 4-uniform $\mathcal{H}$ with $\sim \frac{n^2}{\ell}$ hyperedges has a $\sim \ell \log_2 n$ length cycle.

Key Step:

If there are no cycles of length $\sim \ell \log_2 n$, then regardless of $b_C s$, can prove an **upper bound on** $||A||_2$ that matches the one when b_Cs are indep. random.

Trace Method: $||A||_2 \sim Tr\left(A^{2r}\right)^{\frac{1}{2r}}$ for $r \sim \log\binom{n}{\ell} \sim \ell \log_2 n$.

$$Tr\left(A^{2r}\right) = \sum_{(S_1, S_2, \ldots, S_{2r})} A(S_1, S_2) A(S_2, S_3) \cdots A(S_{2r}, S_1)$$

"2r-length walk" on "vertices" of the "Kikuchi Graph"

Small Cycles via *Spectral Double Counting*

Prop: Whp, random 4-uniform $\mathcal{H}$ with $\sim \frac{n^2}{\ell}$ hyperedges has a $\sim \ell \log_2 n$ length cycle.

Trace Method: $||A||_2 \sim Tr(A^{2r})^{\frac{1}{2r}}$ for $r \sim \log\binom{n}{\ell} \sim \ell \log_2 n$.

$$Tr(A^{2r}) = \sum_{(S_1,S_2,\ldots,S_{2r})} A(S_1,S_2)A(S_2,S_3)\cdots A(S_{2r},S_1)$$

Recall: $A(S_1,S_2) = b_C$ if $S_1 \Delta S_2 = C \Leftrightarrow S_1 \oplus S_2 = C$ for some $C \in \mathcal{H}$.

Each term contributes a $+1$ or 0. So RHS is the number of contributing walks.

When $b_C s$ are independent ± 1, only "even returning walks" contribute.

Returning Walk: walk that uses the same "edge" (i.e., (T, U)) an even # of times.

Observation: If $\mathcal{H}$ has no cycle of length $\sim \log\binom{n}{\ell}$, exact same set of walks contribute regardless of $b_C s$.

Small Cycles via *Spectral Double Counting*

Prop: Whp, random 4-uniform $\mathcal{H}$ with $\sim \frac{n^2}{\ell}$ hyperedges has a $\sim \ell \log_2 n$ length cycle.

$$Tr(A^{2r}) = \sum_{(S_1, S_2, \ldots, S_{2r})} A(S_1, S_2) A(S_2, S_3) \cdots A(S_{2r}, S_1)$$

Recall: $A(S_1, S_2) = b_C$ if $S_1 \Delta S_2 = C \Leftrightarrow S_1 \oplus S_2 = C$ for some $C \in \mathcal{H}$.

Observation: If $\mathcal{H}$ has no cycle of length $\sim \log\binom{n}{\ell}$, only *even returning walks* contribute.

Proof: Any contributing term $(S_1, S_2, \ldots, S_{2r})$ corresponds to $S_1, C_1, C_2, \ldots, C_{2r}$.

$$\left.\begin{aligned} S_1 \oplus S_2 &= C_1 \\ S_2 \oplus S_3 &= C_2 \\ &\cdots \\ S_{2r} \oplus S_1 &= C_{2r} \end{aligned}\right\}$$

Add both sides modulo 2,

$$C_1 \oplus C_2 \cdots \oplus C_{2r} = 0$$

Small Cycles via *Spectral Double Counting*

Prop: Whp, random 4-uniform $\mathcal{H}$ with $\sim \frac{n^2}{\ell}$ hyperedges has a $\sim \ell \log_2 n$ length cycle.

$$\boldsymbol{Tr(A^{2r}) = \sum_{(S_1,S_2,\ldots,S_{2r})} A(S_1,S_2)A(S_2,S_3)\cdots A(S_{2r},S_1)}$$

Recall: $A(S_1,S_2) = b_C$ if $S_1 \Delta S_2 = C \Leftrightarrow S_1 \oplus S_2 = C$ for some $C \in \mathcal{H}$.

Observation: If $\mathcal{H}$ has no cycle of length $\sim \log\binom{n}{\ell}$, only *even returning walks* contribute.

Proof: Any contributing term $(S_1, S_2, \ldots, S_{2r})$ corresponds to $S_1, C_1, C_2, \ldots, C_{2r}$.

$$C_1 \oplus C_2 \cdots \oplus C_{2r} = 0$$

If all C_is are distinct, must be a cycle of length $2r$ in $\mathcal{H}$.
So, can happen only if each C_i occurs an even number of times.
$\Leftrightarrow$ the corresponding walk is **even returning**.

$\square$

What about *semi-random* instances?

Goal: Certify that $\phi(x) = \frac{1}{m}\sum_{C\in\mathcal{H}} b_C x_C \leq \epsilon$ for all $x \in \{\pm 1\}^n$

$\mathcal{H}$ arbitrary (worst-case), b_Cs indep. random.

Spectral norm of A is too large and cannot work.

Obs: "Offending" quadratic forms are on *sparse* vectors.
While we only care about "flat" vectors.

"Row bucketing" allows bounding flat quadratic forms of semirandom matrices.

[Abascal,Guruswami,K'20]

What about *odd-arity* instances?

Goal: Certify that $\phi(x) = \frac{1}{m}\sum_{C\in\mathcal{H}} b_C x_C \leq \epsilon$ for all $x \in \{\pm 1\}^n$

$\mathcal{H}$ arbitrary (worst-case), b_Cs indep. random.

Define an appropriate Kikuchi matrix.
Spectral norm of A is too large and cannot work *even for random 3-XOR!*.

Idea: "Row Pruning" – removing some appropriate rows enough for random case.

More generally, works for hypergraphs with *small spread*.

Hypergraph Regularity Decomposition:

Decompose a k-uniform hypergraph into k'-uniform hypergraphs for $k' \leq k$ + "error" such that each non-error piece has *small spread*.

What about *semi-random* instances?

Goal: Certify that $\phi(x) = \frac{1}{m}\sum_{C\in\mathcal{H}} b_C x_C \leq \epsilon$ for all $x \in \{\pm 1\}^n$

Let's now see how to improve this to a full trade-off for 3-XOR….

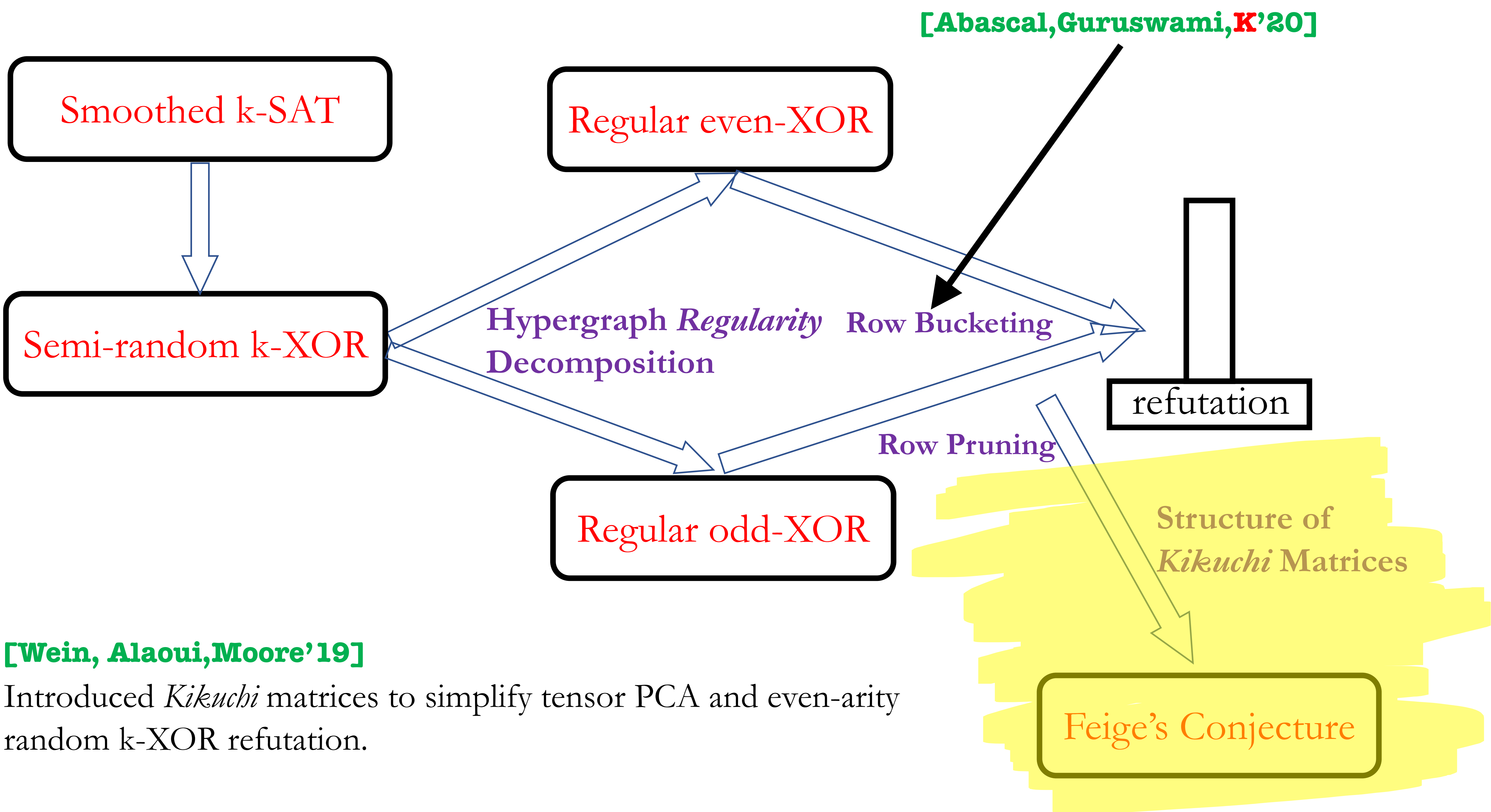
Issue: Natural matrices are rectangular. So spectral norm works only if $m \gg n^2$.

Idea: "Cauchy-Schwarz trick"

$$\phi(x)^2 = \left(\frac{1}{m}\sum_{i} x_i \sum_{C\ni i} b_C x_{C\setminus i}\right)^2 \leq \frac{n}{m^2}\sum_{C,C'\ni i} b_C b_{C'} \underbrace{x_{C\setminus i}\, x_{C'\setminus i}}_{\text{4-xor clause}}$$

Now use the square matrix for 4-XOR.

Issue: Significantly less randomness in matrix A than in case of 4-XOR.
Analysis via trace-moment method. Crucially uses randomness of $\mathcal{H}$.



[Wein, Alaoui,Moore'19]

Introduced *Kikuchi* matrices to simplify tensor PCA and even-arity random k-XOR refutation.