Algorithms and Certificates for Refuting CSPs
“smoothed is no harder than random”

Pravesh Kothari
CMU

Venkat Guruswami
CMU

Peter Manohar
CMU

Refuting CSPs

Refutation Algorithm:

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

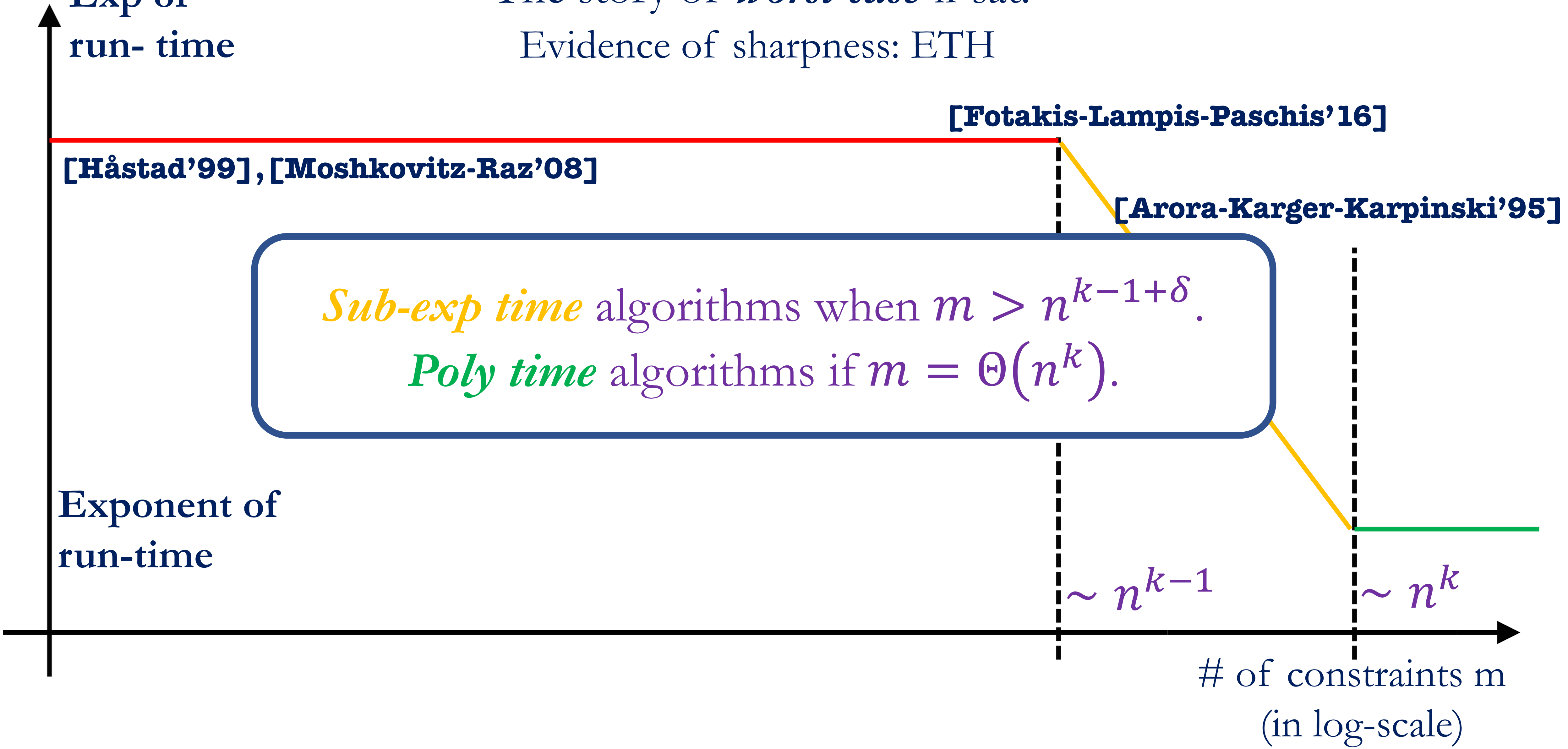
Goal: refute largest possible family of instances ϕ: $val(\phi) < 0.99$.

refutation = *certificate* that $val(\phi) \leq v$

A Tale of Two Worlds

The story of ***worst-case*** k-sat.

Evidence of sharpness: ETH



Sub-exp time algorithms when $m > n^{k-1+\delta}$.

Poly time algorithms if $m = \Theta(n^k)$.

A Tale of Two Worlds

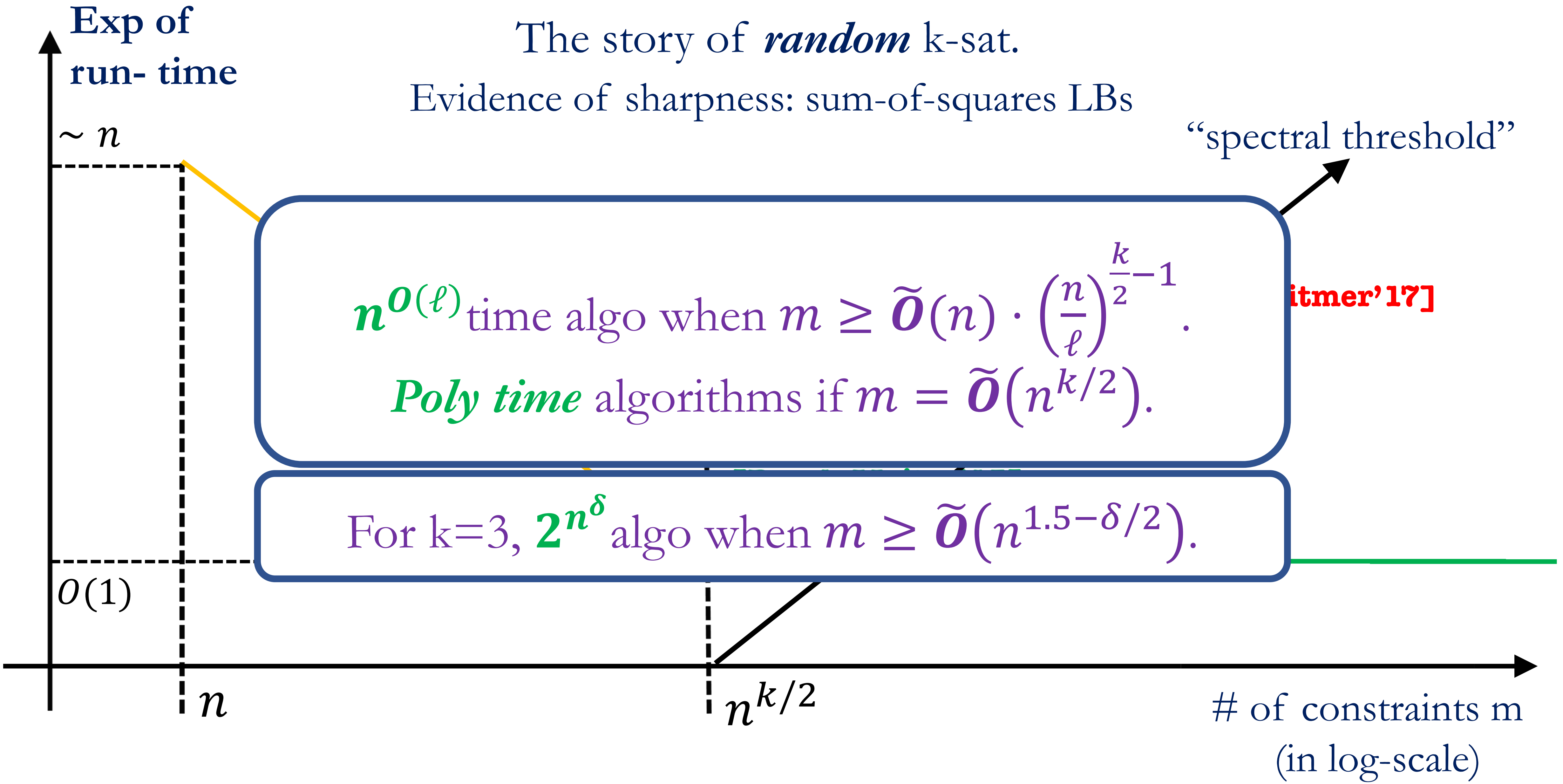
The story of ***random*** k-sat.

Evidence of sharpness: sum-of-squares LBs

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

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

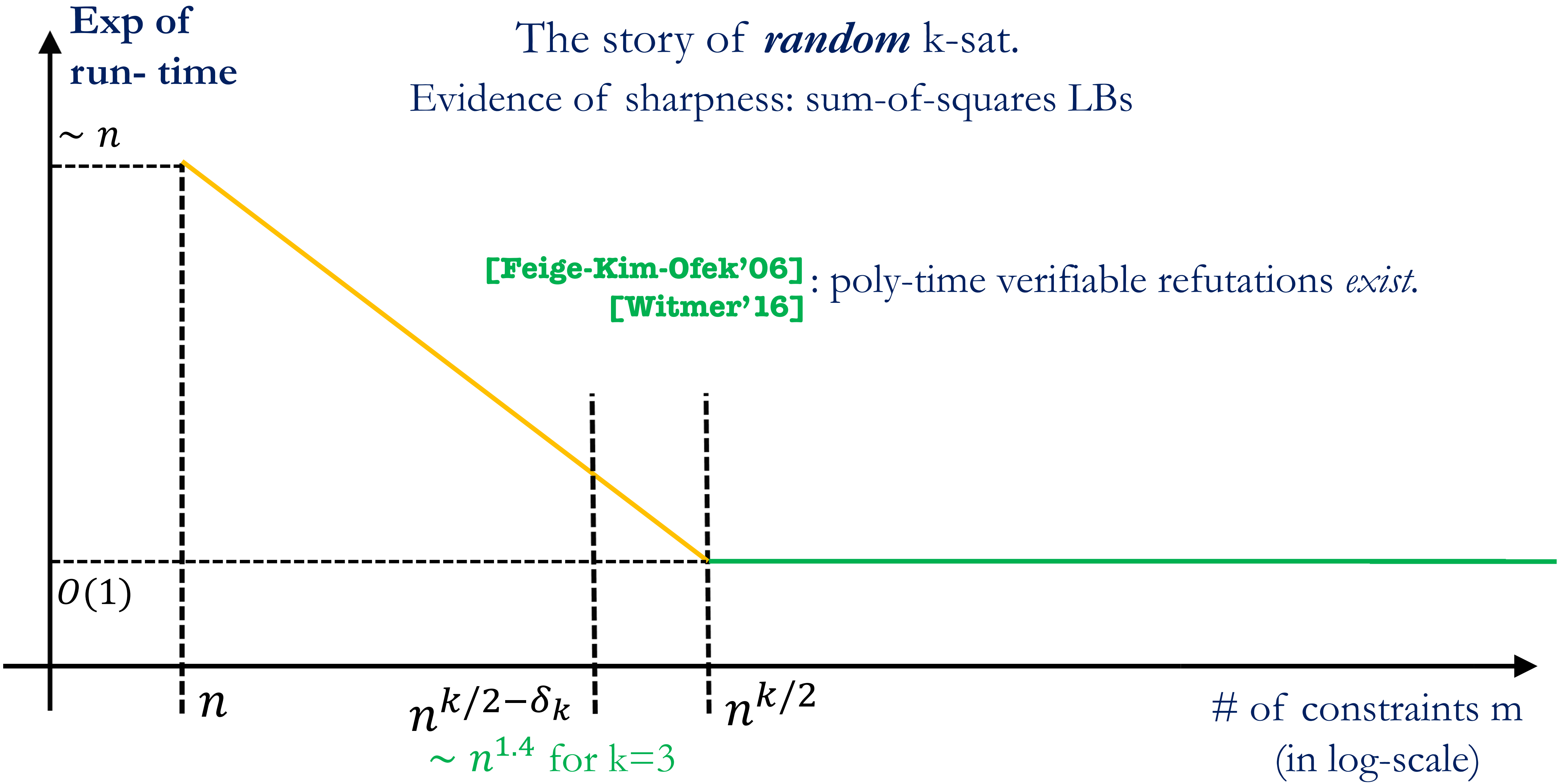
For k=3, $2^{n^{\delta}}$ algo when $m \geq \tilde{O}(n^{1.5-\delta/2})$.



A Tale of Two Worlds

The story of ***random*** k-sat.

Evidence of sharpness: sum-of-squares LBs



How does the complexity of k-sat interpolate between the two worlds?

Is worst-case world pessimistic?

Are random instances idealistic?

Do algorithms/certificates generalize beyond random?

Does the randomness of the clause structure matter?

Smoothed CSPs [Feige'07]

1: Generate worst-case instance ϕ of k-SAT.

2: Negate each literal with prob 0.01 independently to produce $\boldsymbol{\phi_s}$.

Fact: $val(\boldsymbol{\phi_s}) \leq 1 - 2^{-ck}$ whp.

- clause structure (i.e., instance hypergraph) is worst-case.
- only randomness in literals: via small random perturbation.

This Work: Algorithms

[Guruswami,K,Manohar'21]: The story of ***smoothed*** k-sat.

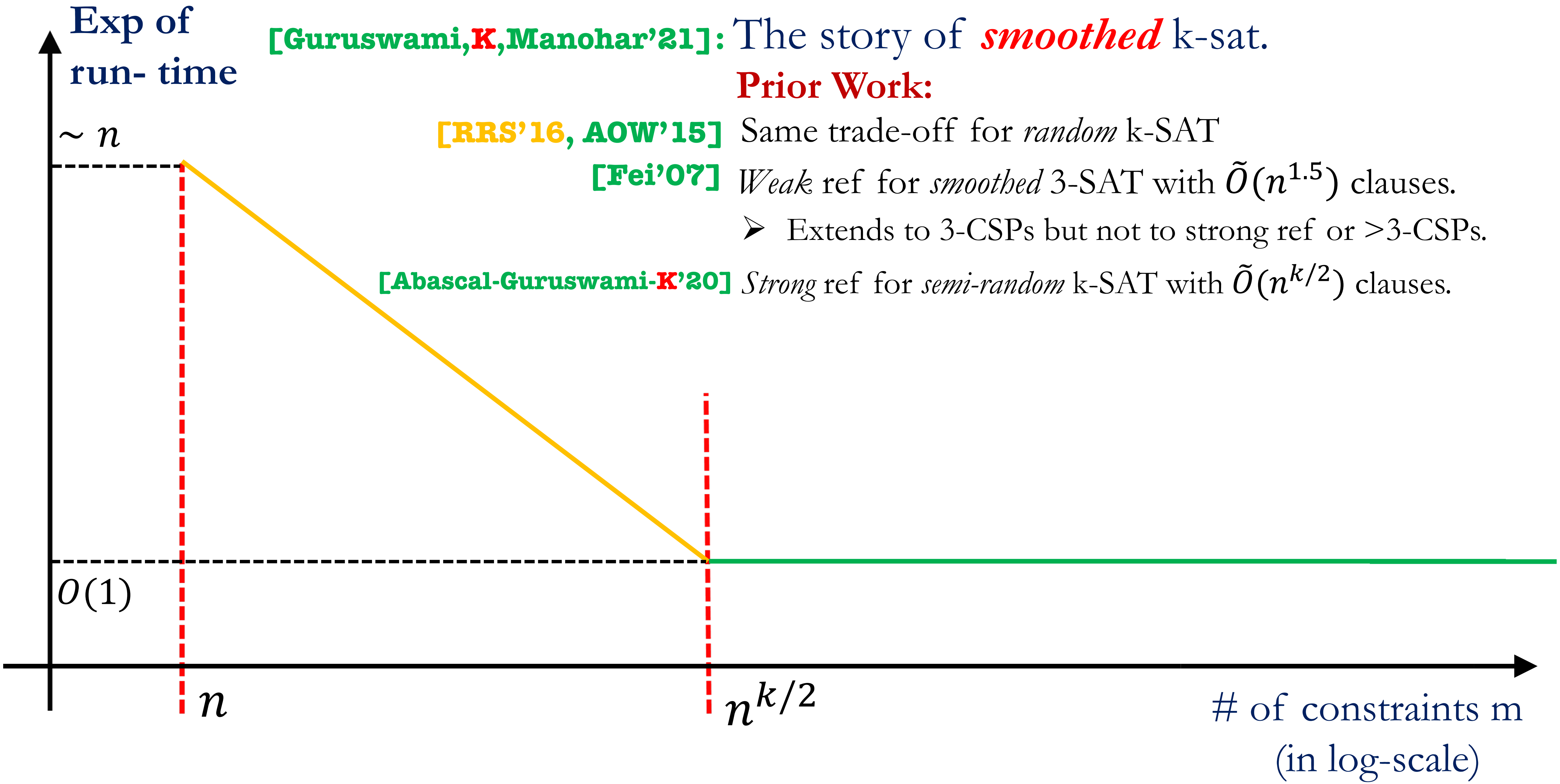
Prior Work:

[RRS'16, AOW'15] Same trade-off for *random* k-SAT

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

- Extends to 3-CSPs but not to strong ref or >3-CSPs.

[Abascal-Guruswami-K'20] *Strong* ref for *semi-random* k-SAT with $\tilde{O}(n^{k/2})$ clauses.



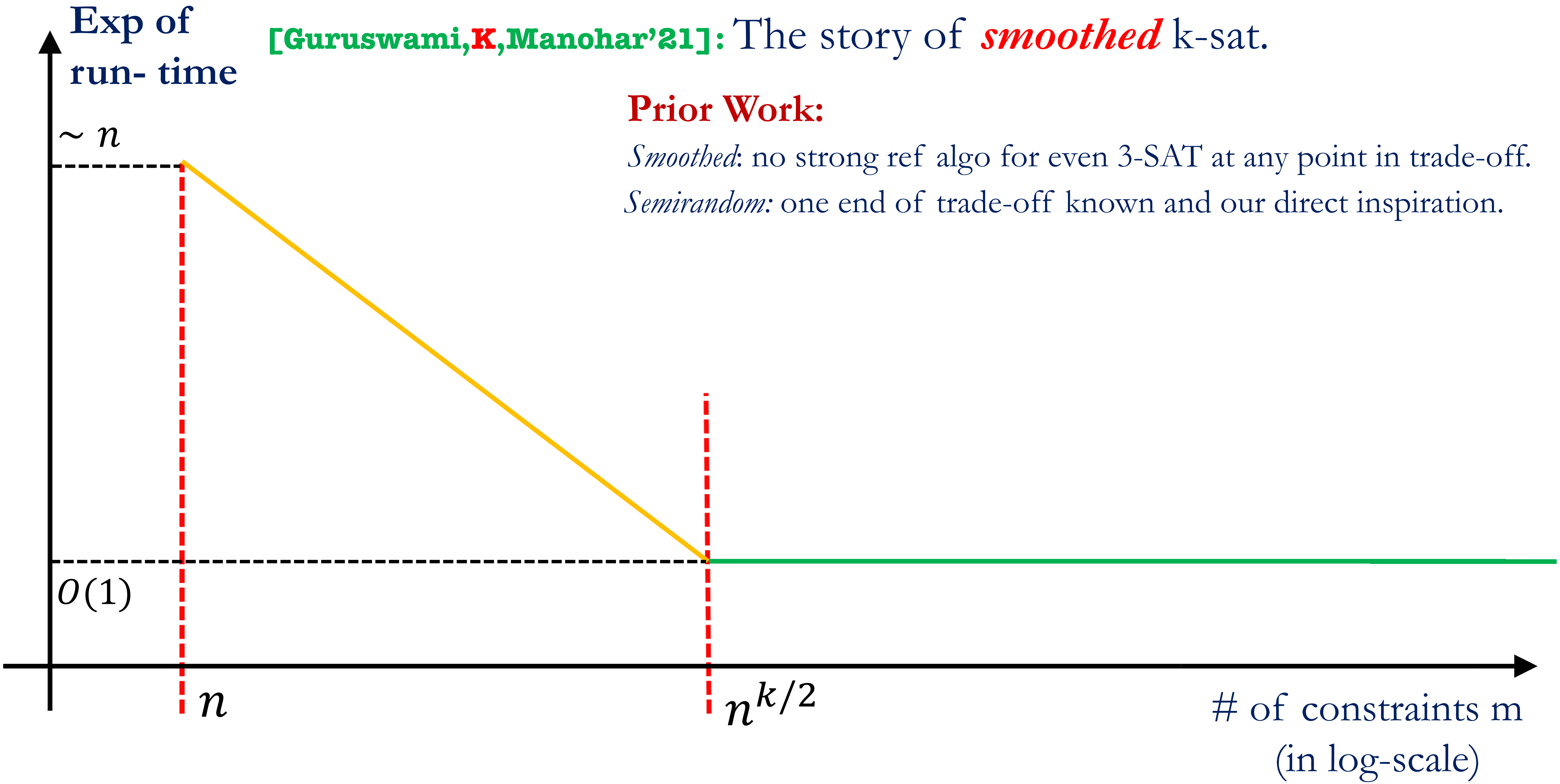
This Work: Algorithms

[Guruswami,K,Manohar'21]: The story of ***smoothed*** k-sat.

Prior Work:

Smoothed: no strong ref algo for even 3-SAT at any point in trade-off.

Semirandom: one end of trade-off known and our direct inspiration.

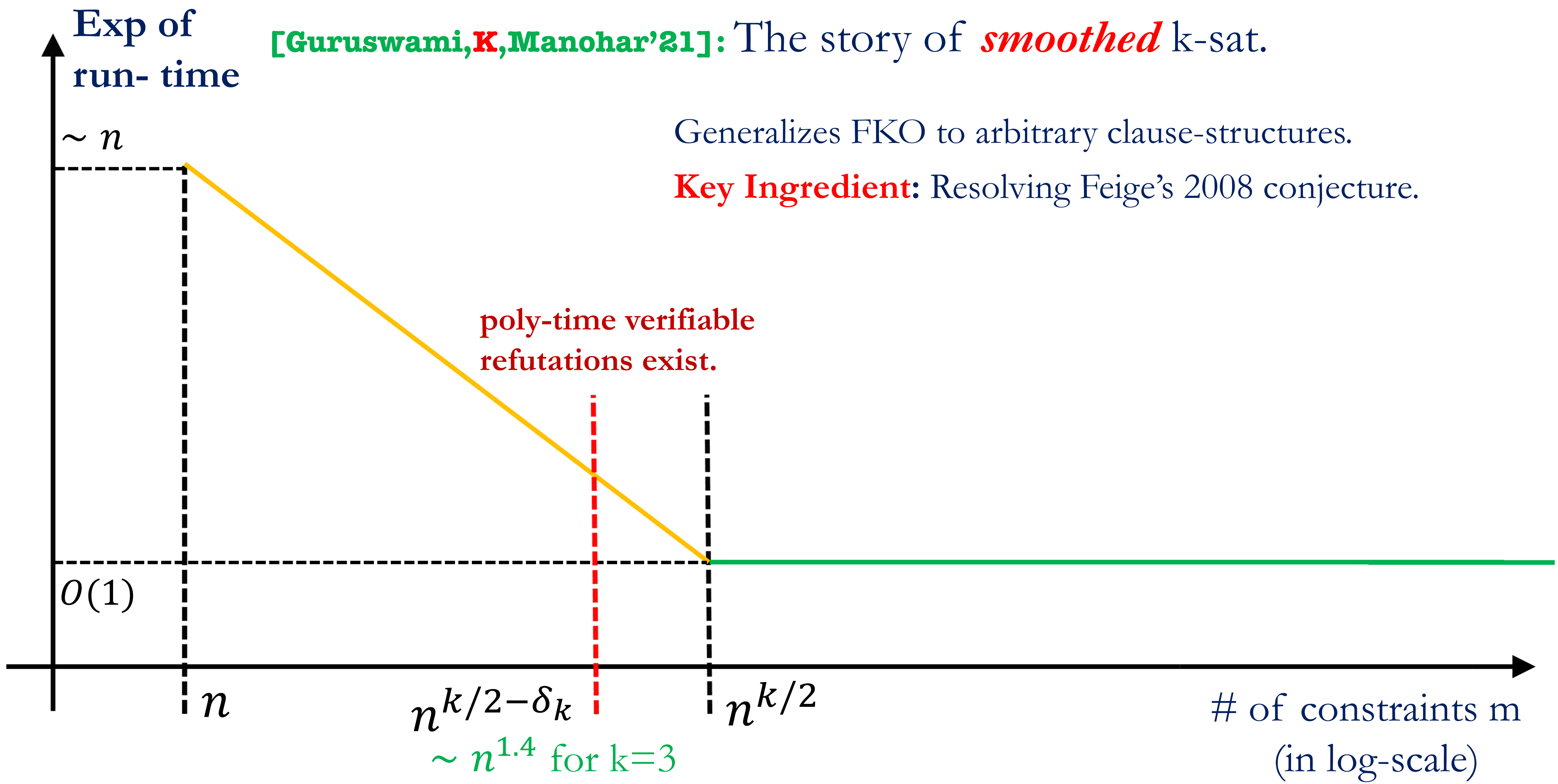


This Work: Certificates

[Guruswami,K,Manohar'21]: The story of ***smoothed*** k-sat.

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 tradeoff 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: Our Result

An extremal conjecture about girth of hypergraphs.

Feige's Conjecture (2008):

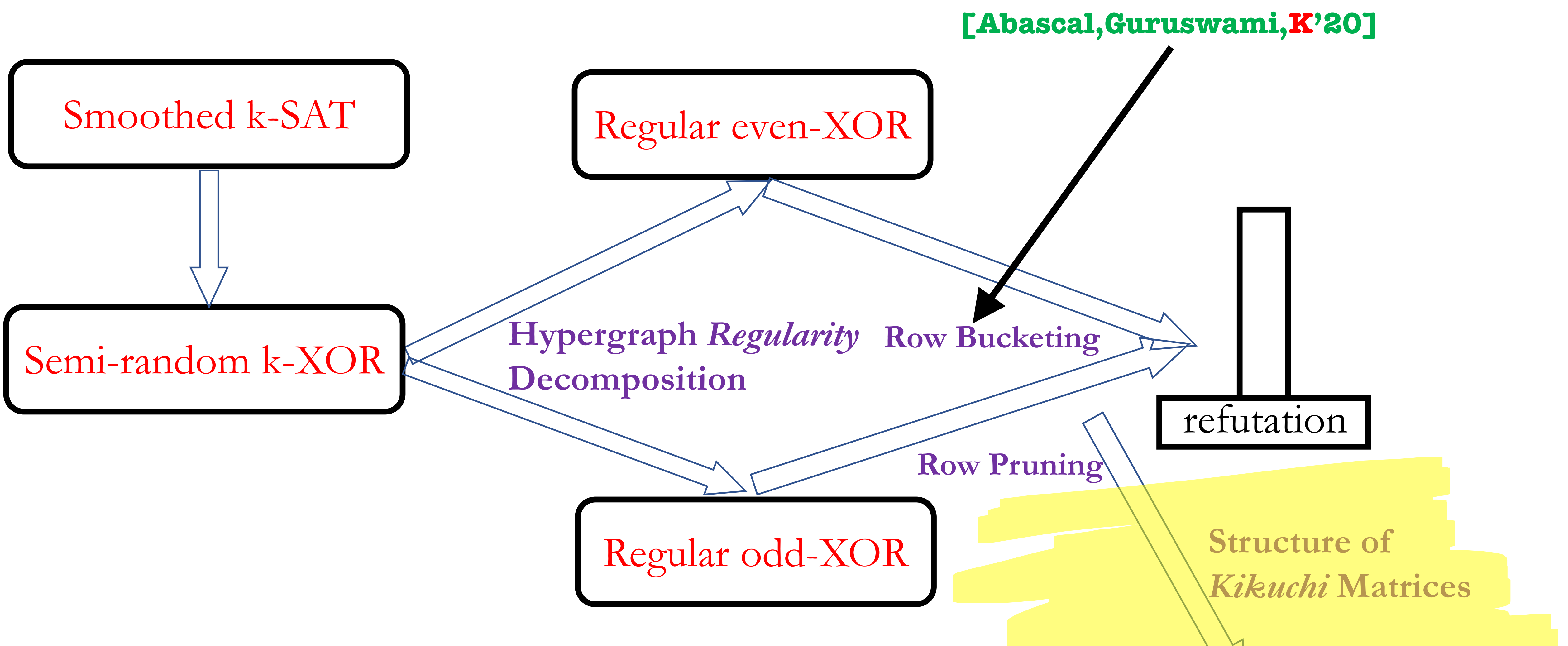
Every hypergraph with $m \geq n \cdot \left(\frac{n}{\ell}\right)^{\left(\frac{k}{2}-1\right)}$ hyperedges has a cycle of length $\leq \ell \log_2 n$.

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

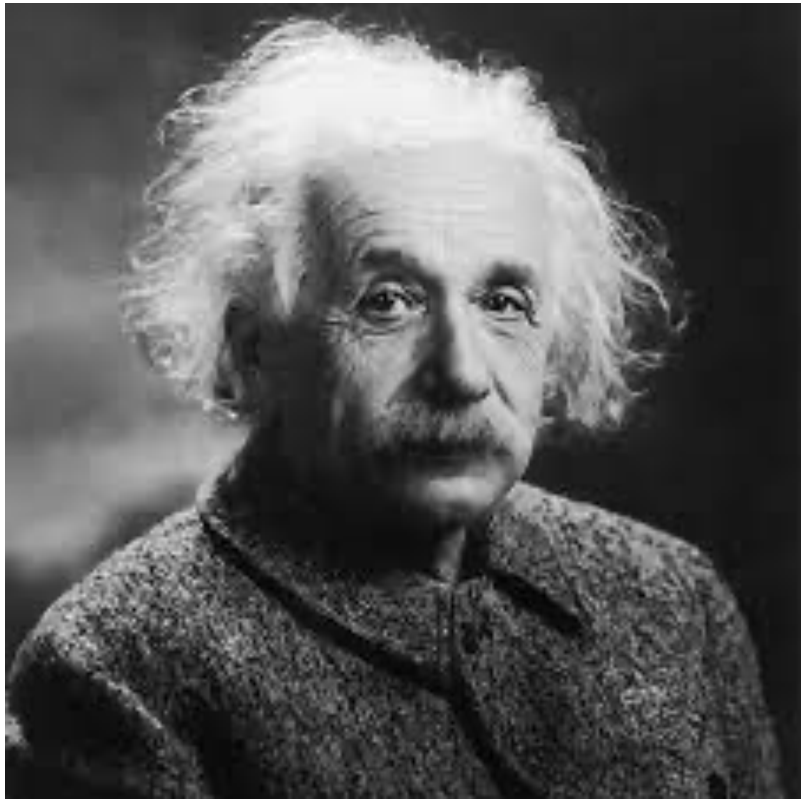
"Spectral double counting" : a conceptually simple connection between hypergraph cycles and *sub-exp size spectral refutations* **below** spectral threshold.

Time for some actual math!



[Wein, Alaoui,Moore'19]

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



"You've got to look at the *Kikuchi* matrices if you want to prove something about CSPs…or hypergraphs…or tensors…"

Tightly refuting *random* 4-XOR

Let's start with the case of $\ell = O(1)$.

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

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 $\phi(x) \le \epsilon$ for all $x \in \{\pm 1\}^n$

Tightly 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 some matrix! **[Goerdt,Krivilevich'01…]**

$\{k, \ell\}$

$A =$ $\{i, j\}$ — $\left[\; b_{\{i,j,k,\ell\}} \; \right]$

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

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

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

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

Tightly 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 $x \in \{\pm 1\}^n$

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*.

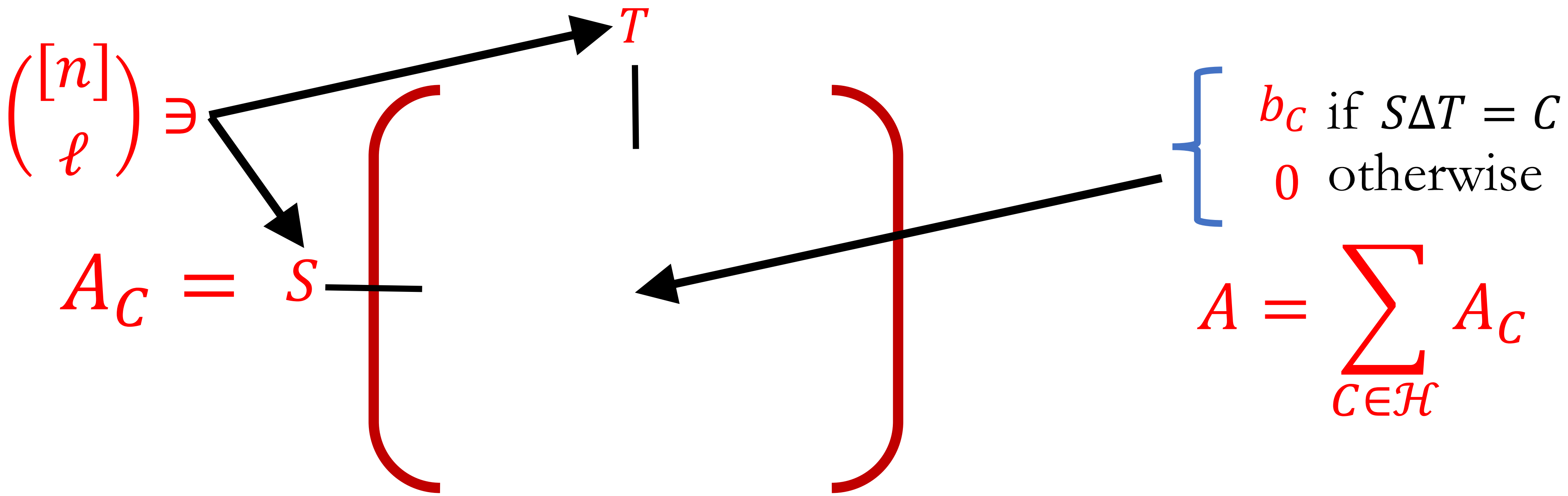
[Wein-Alaoui-Moore'19] Introduce ***Kikuchi* matrix** and significantly simplify **even-arity random** k-XOR refutation.

This is our starting point!

Tightly 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.



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

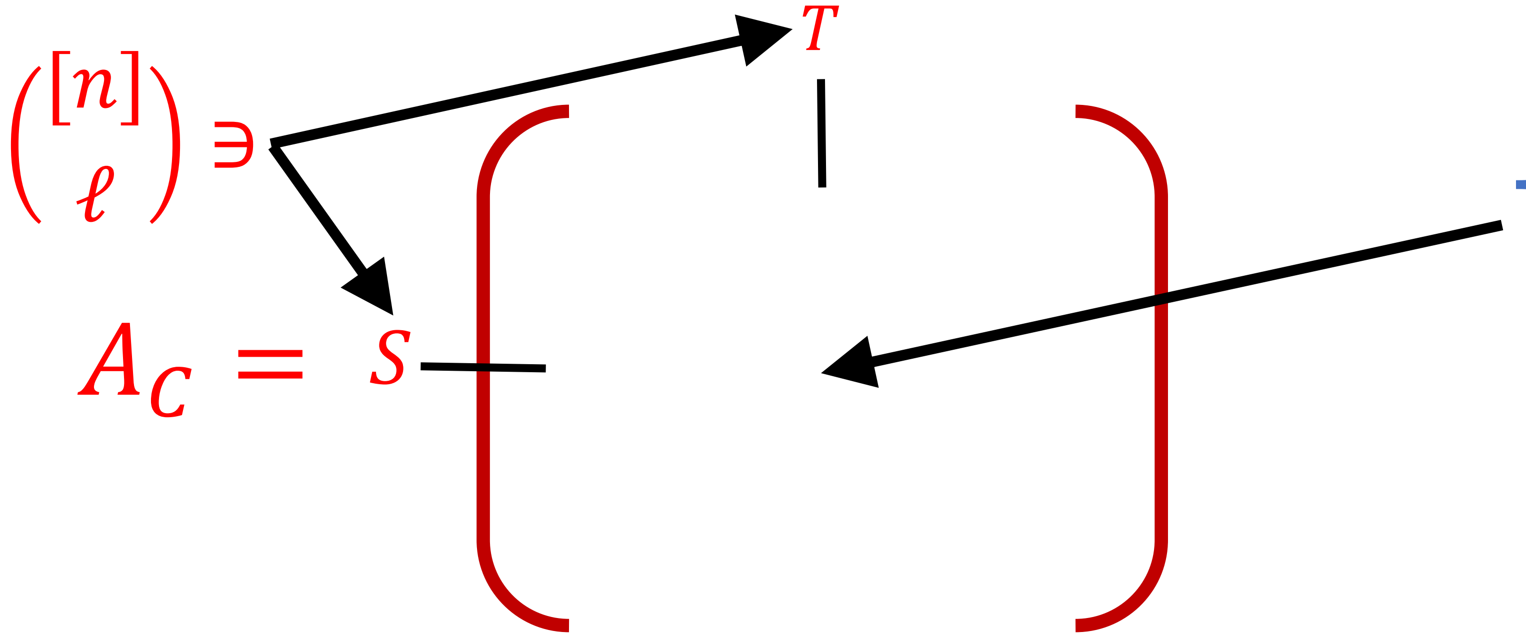
Then, $\phi(x) = \frac{1}{D_\ell}\left(x^{\odot \ell}\right)^{\top} \mathrm{A}\left(x^{\odot \ell}\right) = \frac{1}{D_\ell}\sum_{S,T} A(S,T) x_S x_T$

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

Tightly 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.



$$\begin{cases} b_C & \text{if } S\Delta T = C \\ 0 & \text{otherwise} \end{cases}$$

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

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 $\|A\|_2$?

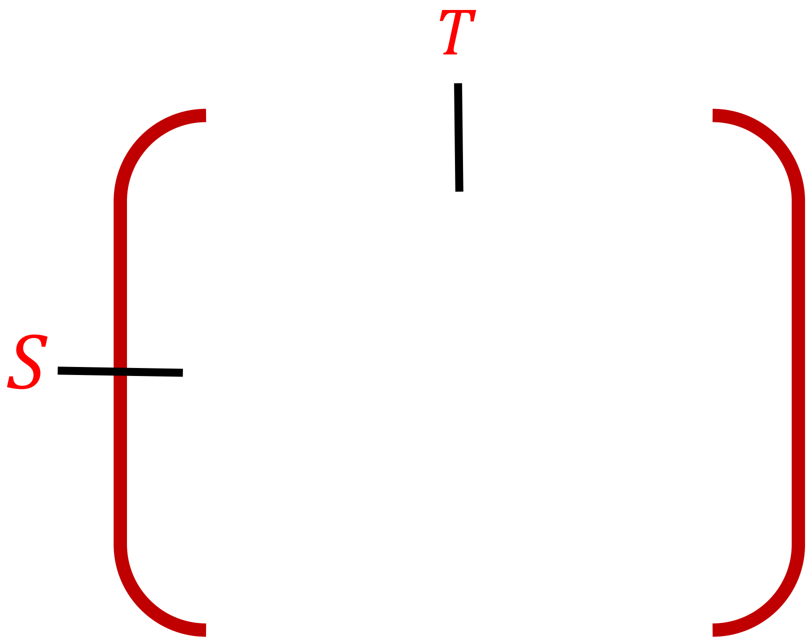
Tightly refuting *random* 4-XOR

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

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

independent, random matrices.

$A =$



Analysis: Apply matrix Chernoff inequality.

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

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 b_Cs**. 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 $||A||_2$** 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** $||\boldsymbol{A}||_2$ that matches the one when b_Cs are indep. random.

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

$$\boldsymbol{Tr}\left(\boldsymbol{A}^{2r}\right) = \sum_{(\boldsymbol{S}_1, \boldsymbol{S}_2, \ldots, \boldsymbol{S}_{2r})} \boldsymbol{A}(\boldsymbol{S}_1, \boldsymbol{S}_2)\boldsymbol{A}(\boldsymbol{S}_2, \boldsymbol{S}_3) \cdots \boldsymbol{A}(\boldsymbol{S}_{2r}, \boldsymbol{S}_1)$$

"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 $S_1, C_1, C_2, \ldots, C_{2r}$.

$$S_1 \oplus S_2 = C_1$$
$$S_2 \oplus S_3 = C_2$$
$$\ldots$$
$$S_{2r} \oplus S_1 = C_{2r}$$

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**.

□

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*.

This work:

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.

Main take-away: Kikuchi matrices are beautiful and can solve all life's problems.

Thank you.