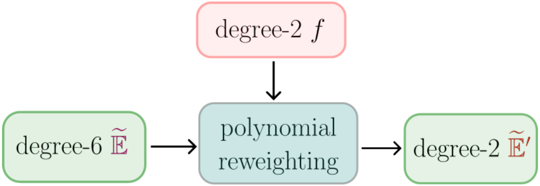
# New SDP Roundings and Certifiable Approximation for Cubic Optimization




$$(\widetilde{\mathbb{E}} f^2)^{\frac{1}{2}} \leq 3|\widetilde{\mathbb{E}}' f|$$

Tim Hsieh

CMU

Pravesh Kothari

CMU

*Lucas Pesenti*

Bocconi University

Luca Trevisan

Bocconi University

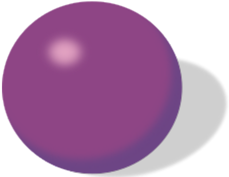
# Quadratic polynomial optimization

**Question:** Given an arbitrary homogeneous degree-2 multilinear polynomial

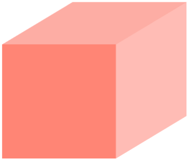
$$p(\mathbf{x}) := p(x_1, \ldots, x_n) = \sum_{1 \leqslant i < j \leqslant n} A_{ij} x_i x_j \,,$$

can we approximate efficiently $\max p(\mathbf{x})$?

$$\max_{\|\mathbf{x}\|_2 = 1} p(\mathbf{x})$$



$$\max_{\|\mathbf{x}\|_\infty = 1} p(\mathbf{x})$$



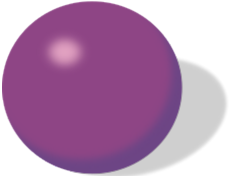
# Quadratic polynomial optimization

**Question:** Given an arbitrary homogeneous degree-2 multilinear polynomial

$$p(\mathbf{x}) := p(x_1, \ldots, x_n) = \sum_{1 \leqslant i < j \leqslant n} A_{ij} x_i x_j \,,$$

can we approximate efficiently $\max p(\mathbf{x})$?

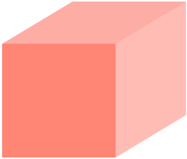
$\max_{\|\mathbf{x}\|_2=1} p(\mathbf{x})$



Largest eigenvalue of

$$\mathbf{A} = \frac{1}{2} \begin{bmatrix} 0 & A_{12} & \ldots & A_{1n} \\ A_{12} & 0 & \ldots & A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \ldots & 0 \end{bmatrix}$$

$\max_{\|\mathbf{x}\|_\infty=1} p(\mathbf{x})$



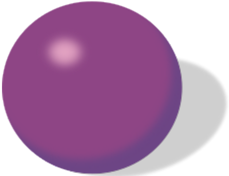
# Quadratic polynomial optimization

**Question:** Given an arbitrary homogeneous degree-2 multilinear polynomial

$$p(\mathbf{x}) := p(x_1, \ldots, x_n) = \sum_{1 \leqslant i < j \leqslant n} A_{ij} x_i x_j ,$$

can we approximate efficiently $\max p(\mathbf{x})$?

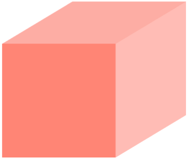
$\max_{\|\mathbf{x}\|_2 = 1} p(\mathbf{x})$



Largest eigenvalue of

$$\mathbf{A} = \frac{1}{2} \begin{bmatrix} 0 & A_{12} & \ldots & A_{1n} \\ A_{12} & 0 & \ldots & A_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ A_{1n} & A_{2n} & \ldots & 0 \end{bmatrix}$$

$\max_{\|\mathbf{x}\|_\infty = 1} p(\mathbf{x})$



[Charikar-Wirth'04] $O(\log n)$-approximation

[Grothendieck'53, ..., Alon-Naor'04]
$O(1)$-approximation when $p(\mathbf{x}, \mathbf{y}) = \sum_{ij} A_{ij} x_i y_j$

Based on rounding the basic SDP relaxation

# Our main results

**Question:** Given an arbitrary homogeneous degree-3 multilinear polynomial

$$p(\mathbf{x}) := p(x_1, \ldots, x_n) = \sum_{1 \leqslant i < j < k \leqslant n} A_{ijk} x_i x_j x_k \,,$$

can we approximate efficiently $\max p(\mathbf{x})$?

**Ex:** "multiplicative" approximation of Max-3XOR

# Our main results

**Question:** Given an arbitrary homogeneous degree-3 multilinear polynomial

$$p(\mathbf{x}) := p(x_1, \ldots, x_n) = \sum_{1 \leqslant i < j < k \leqslant n} A_{ijk} x_i x_j x_k \,,$$

can we approximate efficiently $\max p(\mathbf{x})$?

**Ex:** "multiplicative" approximation of Max-3XOR

---

Over both $S^{n-1}$ and $\{-1, 1\}^n$:

**Thm 1:** [HKPT'24] $O(\sqrt{n})$-approximation in time $n^{O(1)}$ via a "canonical" SDP relaxation

**Thm 2:** [HKPT'24] $O(\sqrt{n/\log n})$-approximation in time $n^{O(1)}$ via a "pruned" SDP relaxation

# Our main results

**Question:** Given an arbitrary homogeneous degree-3 multilinear polynomial

$$p(\mathbf{x}) := p(x_1, \dots, x_n) = \sum_{1 \leqslant i < j < k \leqslant n} A_{ijk} x_i x_j x_k ,$$

can we approximate efficiently $\max p(\mathbf{x})$?

**Ex:** "multiplicative" approximation of Max-3XOR

---

Over both $S^{n-1}$ and $\{-1, 1\}^n$:

**Thm 1:** [HKPT'24] $O(\sqrt{n/\mathbf{k}})$-approximation in time $n^{O(\mathbf{k})}$ via a "canonical" SDP relaxation

**Thm 2:** [HKPT'24] $O(\sqrt{n/\mathbf{k}})$-approximation in time $2^{O(\mathbf{k})} n^{O(1)}$ via a "pruned" SDP relaxation

# Our main results

**Question:** Given an arbitrary homogeneous degree-3 multilinear polynomial

$$p(\mathbf{x}) := p(x_1, \ldots, x_n) = \sum_{1 \leqslant i < j < k \leqslant n} A_{ijk} x_i x_j x_k \,,$$

can we approximate efficiently $\max p(\mathbf{x})$?

**Ex:** "multiplicative" approximation of Max-3XOR

---

Over both $S^{n-1}$ and $\{-1, 1\}^n$:

**Thm 1:** [HKPT'24] $O(\sqrt{n/\mathbf{k}})$-approximation in time $n^{O(\mathbf{k})}$ via a "canonical" SDP relaxation

$\longrightarrow$ generalizes [Bhattiprolu-Gosh-Guruswami-Lee-Tulsiani'17]

**Thm 2:** [HKPT'24] $O(\sqrt{n/\mathbf{k}})$-approximation in time $2^{O(\mathbf{k})} n^{O(1)}$ via a "pruned" SDP relaxation

$\longrightarrow$ matches [Khot-Naor'07] for $\mathbf{k} = \log n$

$\longrightarrow$ also provides a *certifiable upper bound* on $\max p(\mathbf{x})$ (the SDP dual)

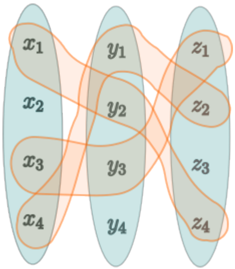
# Decoupling interlude

Why *cubic* optimization?

[Khot-Naor'07] Let $\mathbf{A} = (A_{ijk})$ be a symmetric 3-tensor with zero "diagonal",

$$p(\mathbf{x}) := \sum_{ijk} A_{ijk} x_i x_j x_k \,,$$

$$\tilde{p}(\mathbf{x}, \mathbf{y}, \mathbf{z}) := \sum_{ijk} A_{ijk} x_i y_j z_k \,.$$




*Decoupled polynomial*

# Decoupling interlude

Why *cubic* optimization?

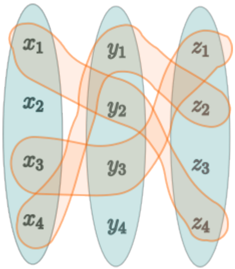
[Khot-Naor'07] Let $\mathbf{A} = (A_{ijk})$ be a symmetric 3-tensor with zero "diagonal",

$$p(\mathbf{x}) := \sum_{ijk} A_{ijk} x_i x_j x_k \,,$$

$$\tilde{p}(\mathbf{x}, \mathbf{y}, \mathbf{z}) := \sum_{ijk} A_{ijk} x_i y_j z_k \,.$$

Then

$$\max_{\mathbf{x} \in \{-1,1\}^n} p(\mathbf{x}) \asymp \max_{\mathbf{x},\mathbf{y},\mathbf{z} \in \{-1,1\}^n} \tilde{p}(\mathbf{x}, \mathbf{y}, \mathbf{z}) \,.$$



*Decoupled polynomial*

# Decoupling interlude

Why *cubic* optimization?

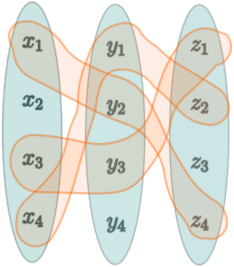
[Khot-Naor'07] Let $\mathbf{A} = (A_{ijk})$ be a symmetric 3-tensor with zero "diagonal",

$$p(\mathbf{x}) := \sum_{ijk} A_{ijk} x_i x_j x_k \,,$$

$$\tilde{p}(\mathbf{x}, \mathbf{y}, \mathbf{z}) := \sum_{ijk} A_{ijk} x_i y_j z_k \,.$$

Then

$$\max_{\mathbf{x} \in \{-1,1\}^n} p(\mathbf{x}) \asymp \max_{\mathbf{x},\mathbf{y},\mathbf{z} \in \{-1,1\}^n} \tilde{p}(\mathbf{x}, \mathbf{y}, \mathbf{z}) \,.$$



*Decoupled polynomial*

Only holds for *odd-degree* (multilinear) polynomials

# Result 1: the canonical sum-of-squares relaxation

**Thm 1:** Let $p(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{ijk} A_{ijk} x_i y_j z_k$. Rounding the degree-6 sum-of-squares SDP relaxation yields a $O(\sqrt{n})$-approximation to $\max_{\mathbf{x},\mathbf{y},\mathbf{z} \in \{-1,1\}^n} p(\mathbf{x}, \mathbf{y}, \mathbf{z})$.

# Result 1: the canonical sum-of-squares relaxation

**Thm 1:** Let $p(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{ijk} A_{ijk} x_i y_j z_k$. Rounding the degree-6 sum-of-squares SDP relaxation yields a $O(\sqrt{n})$-approximation to $\max_{\mathbf{x},\mathbf{y},\mathbf{z} \in \{-1,1\}^n} p(\mathbf{x}, \mathbf{y}, \mathbf{z})$.

**The SDP relaxation:**

$$\max_{\widetilde{\mathbb{E}}} \widetilde{\mathbb{E}}\left[p(\mathbf{x}, \mathbf{y}, \mathbf{z})\right]$$

over all degree-6 pseudo-expectations $\widetilde{\mathbb{E}}$.
(Solvable in polynomial time)

# Result 1: the canonical sum-of-squares relaxation

**Thm 1:** Let $p(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{ijk} A_{ijk} x_i y_j z_k$. Rounding the degree-6 sum-of-squares SDP relaxation yields a $O(\sqrt{n})$-approximation to $\max_{\mathbf{x},\mathbf{y},\mathbf{z} \in \{-1,1\}^n} p(\mathbf{x}, \mathbf{y}, \mathbf{z})$.

**The SDP relaxation:**

$$\max_{\widetilde{\mathbb{E}}} \widetilde{\mathbb{E}} \left[ p(\mathbf{x}, \mathbf{y}, \mathbf{z}) \right]$$

over all degree-6 pseudo-expectations $\widetilde{\mathbb{E}}$.
(Solvable in polynomial time)

**Def:** A degree-6 *pseudo-expectation* $\widetilde{\mathbb{E}}$ is a linear map

$$\widetilde{\mathbb{E}} : \{\text{polynomials of degree} \leqslant 6 \text{ in } \mathbf{x}, \mathbf{y}, \mathbf{z}\} \to \mathbb{R}$$

satisfying "consistency checks" for the degree $\leqslant 6$ moments of a distribution over $\mathbf{x}, \mathbf{y}, \mathbf{z} \in \{-1, 1\}^n$.

# Result 1: the canonical sum-of-squares relaxation

**Thm 1:** Let $p(\mathbf{x},\mathbf{y},\mathbf{z}) = \sum_{ijk} A_{ijk} x_i y_j z_k$. Rounding the degree-6 sum-of-squares SDP relaxation yields a $O(\sqrt{n})$-approximation to $\max_{\mathbf{x},\mathbf{y},\mathbf{z}\in\{-1,1\}^n} p(\mathbf{x},\mathbf{y},\mathbf{z})$.

**The SDP relaxation:**

$$\max_{\widetilde{\mathbb{E}}} \widetilde{\mathbb{E}}\left[p(\mathbf{x},\mathbf{y},\mathbf{z})\right]$$

over all degree-6 pseudo-expectations $\widetilde{\mathbb{E}}$.
(Solvable in polynomial time)

**Def:** A degree-6 *pseudo-expectation* $\widetilde{\mathbb{E}}$ is a linear map

$$\widetilde{\mathbb{E}} : \{\text{polynomials of degree} \leqslant 6 \text{ in } \mathbf{x},\mathbf{y},\mathbf{z}\} \to \mathbb{R}$$

satisfying "consistency checks" for the degree $\leqslant 6$ moments of a distribution over $\mathbf{x},\mathbf{y},\mathbf{z} \in \{-1,1\}^n$.

**Ex:** if $f$ is a polynomial in $\mathbf{x},\mathbf{y},\mathbf{z}$ of degree $\leqslant 3$, $|\widetilde{\mathbb{E}} f| \leqslant (\widetilde{\mathbb{E}} f^2)^{\frac{1}{2}}$

# Result 1: new rounding via polynomial reweightings

**Input:** degree-6 $\widetilde{\mathbb{E}}$ maximizing $\widetilde{\mathbb{E}}\,[p(\mathbf{x}, \mathbf{y}, \mathbf{z})]$

1. Draw at random $\overline{\mathbf{x}} \sim \{-1, 1\}^n$
2. Use quadratic optimization rounding on $\widetilde{\mathbb{E}}$ to get $\overline{\mathbf{y}}, \overline{\mathbf{z}} \in \{-1, 1\}^n$
3. Output $(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{\mathbf{z}})$

# Result 1: new rounding via polynomial reweightings

**Input:** degree-6 $\widetilde{\mathbb{E}}$ maximizing $\widetilde{\mathbb{E}}\left[p(\mathbf{x}, \mathbf{y}, \mathbf{z})\right]$

1. Draw at random $\overline{\mathbf{x}} \sim \{-1, 1\}^n$
2. Use quadratic optimization rounding on $\widetilde{\mathbb{E}}$ to get $\overline{\mathbf{y}}, \overline{\mathbf{z}} \in \{-1, 1\}^n$
3. Output $(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{\mathbf{z}})$

$\widetilde{\mathbb{E}}\left[p(\overline{\mathbf{x}}, \mathbf{y}, \mathbf{z})\right]$ could be much smaller than $\widetilde{\mathbb{E}}\left[p(\mathbf{x}, \mathbf{y}, \mathbf{z})\right]$ for a typical $\overline{\mathbf{x}}$

# Result 1: new rounding via polynomial reweightings

**Input:** degree-6 $\widetilde{\mathbb{E}}$ maximizing $\widetilde{\mathbb{E}}\left[p(\mathbf{x}, \mathbf{y}, \mathbf{z})\right]$

1. Draw at random $\overline{\mathbf{x}} \sim \{-1, 1\}^n$
2. Use quadratic optimization rounding on $\widetilde{\mathbb{E}}$ to get $\overline{\mathbf{y}}, \overline{\mathbf{z}} \in \{-1, 1\}^n$
3. Output $(\overline{\mathbf{x}}, \overline{\mathbf{y}}, \overline{\mathbf{z}})$

$\widetilde{\mathbb{E}}\left[p(\overline{\mathbf{x}}, \mathbf{y}, \mathbf{z})\right]$ could be much smaller than $\widetilde{\mathbb{E}}\left[p(\mathbf{x}, \mathbf{y}, \mathbf{z})\right]$ for a typical $\overline{\mathbf{x}}$

Move the mass of $\widetilde{\mathbb{E}}$ towards where $\left|p(\overline{\mathbf{x}}, \mathbf{y}, \mathbf{z})\right|$ is large

# Result 1: new rounding via polynomial reweightings

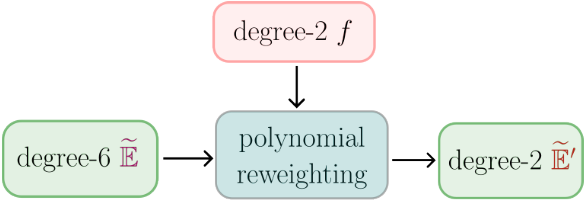
**Input:** degree-6 $\widetilde{\mathbb{E}}$ maximizing $\widetilde{\mathbb{E}}\,[p(\mathbf{x},\mathbf{y},\mathbf{z})]$

1. Draw at random $\overline{\mathbf{x}} \sim \{-1,1\}^n$
2. Use quadratic optimization rounding on $\widetilde{\mathbb{E}}$ to get $\overline{\mathbf{y}},\overline{\mathbf{z}} \in \{-1,1\}^n$
3. Output $(\overline{\mathbf{x}},\overline{\mathbf{y}},\overline{\mathbf{z}})$

$\widetilde{\mathbb{E}}\,[p(\overline{\mathbf{x}},\mathbf{y},\mathbf{z})]$ could be much smaller than $\widetilde{\mathbb{E}}\,[p(\mathbf{x},\mathbf{y},\mathbf{z})]$ for a typical $\overline{\mathbf{x}}$

Move the mass of $\widetilde{\mathbb{E}}$ towards where $|p(\overline{\mathbf{x}},\mathbf{y},\mathbf{z})|$ is large

**Key Lemma:**
[HKPT'24]




$$(\widetilde{\mathbb{E}} f^2)^{\frac{1}{2}} \leq 3|\widetilde{\mathbb{E}}' f|$$

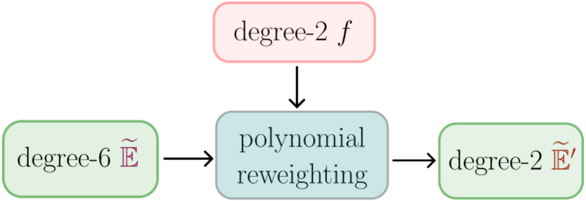
# Result 1: new rounding via polynomial reweightings

**Input:** degree-6 $\widetilde{\mathbb{E}}$ maximizing $\widetilde{\mathbb{E}}\,[p(\mathbf{x},\mathbf{y},\mathbf{z})]$

1. Draw at random $\bar{\mathbf{x}} \sim \{-1,1\}^n$
2. **Reweight $\widetilde{\mathbb{E}}$ into $\widetilde{\mathbb{E}}'$ based on $f = p(\bar{\mathbf{x}},\mathbf{y},\mathbf{z})$**
3. Use quadratic optimization rounding on $\widetilde{\mathbb{E}}'$ to get $\bar{\mathbf{y}},\bar{\mathbf{z}} \in \{-1,1\}^n$
4. Output $(\bar{\mathbf{x}},\bar{\mathbf{y}},\bar{\mathbf{z}})$

Move the mass of $\widetilde{\mathbb{E}}$ towards where $|p(\bar{\mathbf{x}},\mathbf{y},\mathbf{z})|$ is large

**Key Lemma:** [HKPT'24]



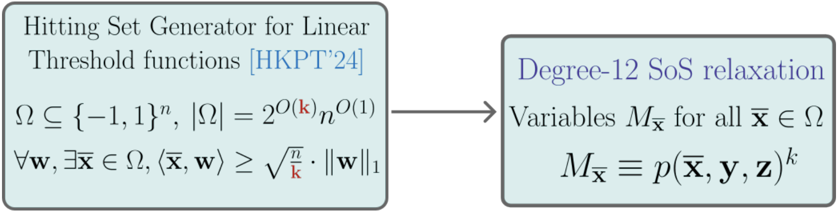

$$\left(\widetilde{\mathbb{E}} f^2\right)^{\frac{1}{2}} \le 3\left|\widetilde{\mathbb{E}}' f\right|$$

# Result 2: the pruned SDP relaxation

**Thm 2:** Let $p(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{ijk} A_{ijk} x_i y_j z_k$. There is an SDP relaxation of $\max p$ with $2^{O(\mathbf{k})} n^{O(1)}$ variables/constraints that achieves approximation $O(\sqrt{n/\mathbf{k}})$.

# Result 2: the pruned SDP relaxation

**Thm 2:** Let $p(\mathbf{x}, \mathbf{y}, \mathbf{z}) = \sum_{ijk} A_{ijk} x_i y_j z_k$. There is an SDP relaxation of $\max p$ with $2^{O(\mathbf{k})} n^{O(1)}$ variables/constraints that achieves approximation $O(\sqrt{n/\mathbf{k}})$.



# Application to Max-3SAT

**Problem:** Given a *satisfiable* 3SAT formula with $n$ variables, find an assignment satisfying as many clauses as possible.

# Application to Max-3SAT

**Problem:** Given a *satisfiable* 3SAT formula with $n$ variables, find an assignment satisfying as many clauses as possible.

- [Hastad'01] Satisfying a $\frac{7}{8} + \varepsilon$ fraction of the clauses is NP-hard.

# Application to Max-3SAT

**Problem:** Given a *satisfiable* 3SAT formula with $n$ variables, find an assignment satisfying as many clauses as possible.

- [Hastad'01] Satisfying a $\frac{7}{8} + \varepsilon$ fraction of the clauses is NP-hard.
- A random assignment satisfies a $\frac{7}{8} \pm \frac{1}{\sqrt{\#\text{clauses}}}$ fraction of the clauses.

# Application to Max-3SAT

**Problem:** Given a *satisfiable* 3SAT formula with $n$ variables, find an assignment satisfying as many clauses as possible.

- [Hastad'01] Satisfying a $\frac{7}{8} + \varepsilon$ fraction of the clauses is NP-hard.
- A random assignment satisfies a $\frac{7}{8} \pm \frac{1}{\sqrt{\#\text{clauses}}}$ fraction of the clauses.

[HKPT'24] $n^{O(1)}$-time algorithm satisfying a $\approx \frac{7}{8} + n^{-\frac{3}{4}}$ fraction of the clauses

# Application to Max-3SAT

**Problem:** Given a *satisfiable* 3SAT formula with $n$ variables, find an assignment satisfying as many clauses as possible.

- [Hastad'01] Satisfying a $\frac{7}{8} + \varepsilon$ fraction of the clauses is NP-hard.
- A random assignment satisfies a $\frac{7}{8} \pm \frac{1}{\sqrt{\#\text{clauses}}}$ fraction of the clauses.

[HKPT'24] $n^{O(1)}$-time algorithm satisfying a $\approx \frac{7}{8} + n^{-\frac{3}{4}}$ fraction of the clauses

✖ The advantage over $\frac{7}{8}$ is a *non-homogeneous* polynomial.

# Application to Max-3SAT

**Problem:** Given a *satisfiable* 3SAT formula with $n$ variables, find an assignment satisfying as many clauses as possible.

- [Hastad'01] Satisfying a $\frac{7}{8} + \varepsilon$ fraction of the clauses is NP-hard.
- A random assignment satisfies a $\frac{7}{8} \pm \frac{1}{\sqrt{\#\text{clauses}}}$ fraction of the clauses.

[HKPT'24] $n^{O(1)}$-time algorithm satisfying a $\approx \frac{7}{8} + n^{-\frac{3}{4}}$ fraction of the clauses

The advantage over $\frac{7}{8}$ is a *non-homogeneous* polynomial.

- Optimize separately the degree-1, 2, 3 parts.
- Add specific constraints to the SDP to avoid cancellations.

# Conclusion

**Problem:** maximize a homogeneous degree-3 polynomial over $S^{n-1}$ or $\{\pm 1\}^n$

- Rounding of the canonical SoS hierarchy via *polynomial reweightings*.
- Slightly improved time/approximation tradeoff via a *pruned SDP relaxation*.
- Improved approximation of satisfiable Max-3SAT instances by adding ad-hoc constraints.

# Conclusion

**Problem:** maximize a homogeneous degree-3 polynomial over $S^{n-1}$ or $\{\pm 1\}^n$

- Rounding of the canonical SoS hierarchy via *polynomial reweightings*.
- Slightly improved time/approximation tradeoff via a *pruned SDP relaxation*.
- Improved approximation of satisfiable Max-3SAT instances by adding ad-hoc constraints.

**Open problems:**

- Best known integrality gap for cubic optimization: $n^{1/4}$
- $\sqrt{n}$-approximation for *non-satisfiable* 3SAT instances?
- Applications of our rounding techniques to other problems?