Computing Sparse Fourier Sum of Squares on Finite Abelian Groups

Lihong Zhi

Chinese Academy of Sciences

Fundamental Algorithms and Algorithmic Complexity, September 25-29, 2021

Joint work with Jianting Yang (CNRS@CREATE) and Ke Ye (CAS)

Outline

Sparse Fourier SOS

Sparse FSOS of Integer Valued Functions

Lower Bounds of Functions on Finite Abelian Groups

Outline

Sparse Fourier SOS

Sparse FSOS of Integer Valued Functions

Lower Bounds of Functions on Finite Abelian Groups

Group Theory

- $\mathbb{Z}_n = \{0, 1, 2, ..., n-1\}$ is the cyclic group of order n, with modular addition.
- If G is a finite abelian group, then

$$G \simeq \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_k}.$$

for some integers $n_1, \ldots, n_k$.

Motivation and Problem Statement

- For $f : G \mapsto \mathbb{C}$

$$f \geq 0 \iff f \text{ has a Fourier SOS, i.e. } f = \sum_{i \in I} |g_i|^2.$$

- Applications: MAXSAT, MAXCUT.

MAXSAT Problem

- $C_2^n = \{-1,1\}^n \cong \mathbb{Z}_2^n$
- For $c = x_1 \vee \cdots \vee x_s \vee \neg x_{s+1} \vee \cdots \vee \neg x_n$, define its characteristic function:

$$f_c(y) = \frac{1}{2^{s+r}} \prod_{i=1}^{s} (1+y_i) \cdot \prod_{i=s+1}^{n} (1-y_i), \quad y \in C_2^n.$$

$$f_c(y) = \begin{cases} 1, & \text{if } y_1 = \cdots = y_s = 1(\text{false}), y_{s+1} = \cdots = y_n = -1(\text{true}) \\ 0, & \text{otherwise}. \end{cases}$$

- For the CNF formula $\phi = \bigwedge_{i=1}^{m} c_i$ in n variables, we define its characteristic function by

$$f_\phi(y) = \sum_{i=1}^{m} f_{c_i}(y), \quad y \in C_2^n.$$

- $f_\phi(y) = \#\{\text{falsified clauses in } \phi \text{ with assignment } y\}$.

Representation of Finite Abelian Group

Character

A nonzero complex function χ on finite abelian group G is called a character of G if it satisfies:

$$\chi(xy) = \chi(x)\chi(y), \quad \forall x, y \in G.$$

$\widehat{G}$: the set of all characters of G, also called Fourier basis, $\widehat{G} \simeq G$.

- Any $f : G \to \mathbb{C}$ has a unique form $f(x) = \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi(x)$.
- Support of f is $\mathrm{supp}(f) = \{\chi : \widehat{f}(\chi) \neq 0\}$.

Representation of Finite Abelian Group

Character

A nonzero complex function χ on finite abelian group G is called a character of G if it satisfies:

$$\chi(xy) = \chi(x)\chi(y), \quad \forall x, y \in G.$$

$\widehat{G}$: the set of all characters of G, also called Fourier basis, $\widehat{G} \simeq G$.

- Any $f : G \to \mathbb{C}$ has a unique form $f(x) = \sum_{\chi \in \widehat{G}} \widehat{f}(\chi)\chi(x)$.
- Support of f is $\mathrm{supp}(f) = \{\chi : \widehat{f}(\chi) \neq 0\}$.

Example [Fawzi,Saunderson,Parrilo'16]

$f : \mathbb{Z}_6 \to \mathbb{C}, \; f(x) = 1 - \frac{1}{2}(\chi_1(x) + \chi_5(x)), \quad \chi_k(x) = e^{\frac{2ik\pi x}{6}}, \; k = 0, 1, 2, \cdots, 5.$

Gram Matrix and FSOS

- G : finite abelian group.
- $f : G \to \mathbb{C}$ map.
- $f \geq 0 \iff \exists Q = (Q_{\chi,\chi'})_{\chi,\chi' \in \widehat{G}} \in \mathbb{C}^{\widehat{G} \times \widehat{G}}$, s.t.

$$Q \succeq 0, \quad \sum_{\chi' \in \widehat{G}} Q_{\chi',\chi'\chi} = \widehat{f}(\chi), \quad \forall \chi \in \widehat{G}$$

$$\iff Q = M^* M, \; M = (M_{j,\chi})_{1 \leq j \leq r, \chi \in \widehat{G}} \in \mathbb{C}^{r \times \widehat{G}}$$

$$f = \sum_{j=1}^{r} |\sum_{\chi \in \widehat{G}} M_{j,\chi} \chi|^2 = \sum_{j=1}^{r} |g_j|^2.$$

- Q : Gram matrix of f, not unique.
- Sparsity of FSOS $f = \sum_{j=1}^{r} |g_j|^2$ is $\left|\bigcup_{j=1}^{r} \mathrm{supp}(g_j)\right|$.

Example of Gram Matrix

Example [Fawzi,Saunderson,Parrilo'16]

$f : \mathbb{Z}_6 \to \mathbb{C},\ f(x) = 1 - \frac{1}{2}(\chi_1(x) + \chi_5(x)),\ \chi_k(x) = e^{\frac{2ik\pi x}{6}}.k = 0, 1, \ldots, 5.$

- Gram matrix of f:

$$\frac{1}{6} \cdot \begin{array}{c} \\ \chi_0 \\ \chi_1 \\ \chi_2 \\ \chi_3 \\ \chi_4 \\ \chi_5 \end{array} \begin{array}{c} \begin{array}{cccccc} \chi_0 & \chi_1 & \chi_2 & \chi_3 & \chi_4 & \chi_5 \end{array} \\ \begin{bmatrix} 1 & -1/2 & 0 & 0 & 0 & -1/2 \\ -1/2 & 1 & -1/2 & 0 & 0 & 0 \\ 0 & -1/2 & 1 & -1/2 & 0 & 0 \\ 0 & 0 & -1/2 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & -1/2 & 1 & -1/2 \\ -1/2 & 0 & 0 & 0 & -1/2 & 1 \end{bmatrix} \end{array}.$$

FSOS Sparsity

Gram matrix $Q \Rightarrow$ FSOS (by Cholesky decomposition).

$$
\begin{aligned}
f = \quad & \frac{1}{6}\left|\frac{\chi_5}{2}+\frac{\chi_1}{2}-1\right|^2+\frac{1}{6}\left|\frac{\sqrt{3}\,\chi_5}{6}+\frac{\sqrt{3}\,\chi_2}{3}-\frac{\sqrt{3}\,\chi_1}{2}\right|^2 \\
& +\frac{1}{6}\left|\frac{\sqrt{6}\,\chi_5}{12}+\frac{\sqrt{6}\,\chi_3}{4}-\frac{\sqrt{6}\,\chi_2}{3}\right|^2+\frac{1}{6}\left|\frac{\sqrt{10}\,\chi_5}{20}+\frac{\sqrt{10}\,\chi_4}{5}-\frac{\sqrt{10}\,\chi_3}{4}\right|^2 \\
& +\frac{1}{6}\left|\frac{\sqrt{15}\,\chi_4}{5}-\frac{\sqrt{15}\,\chi_5}{5}\right|^2
\end{aligned}
$$

Sparsity of FSOS of f is 6.

Example of FSOS

- $f : \mathbb{Z}_6 \to \mathbb{C},\ f(x) = 1 - \frac{1}{2}(\chi_1(x) + \chi_5(x))$, $\chi_k(x) = e^{\frac{2ik\pi x}{6}}$.
- f admits a FSOS with sparsity 4. [Fawzi,Saunderson,Parrilo'16]

Example of FSOS

- $f : \mathbb{Z}_6 \to \mathbb{C},\ f(x) = 1 - \frac{1}{2}(\chi_1(x) + \chi_5(x)),\ \chi_k(x) = e^{\frac{2ik\pi x}{6}}$.
- f admits a FSOS with sparsity 4. [Fawzi,Saunderson,Parrilo'16]
- f admits a FSOS with sparsity 2. [Yang,Ye,Zhi'22a]
  - Gram matrix of f:

$$
\begin{array}{c} \\ \chi_0 \\ \chi_1 \\ \chi_2 \\ \chi_3 \\ \chi_4 \\ \chi_5 \end{array}
\begin{array}{c}
\begin{array}{cccccc} \chi_0 & \chi_1 & \chi_2 & \chi_3 & \chi_4 & \chi_5 \end{array} \\
\begin{bmatrix}
1/2 & -1/2 & 0 & 0 & 0 & 0 \\
-1/2 & 1/2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\end{array}
$$

  - $f(x) = \frac{1}{2}\left|1 - \chi_1(x)\right|^2$.

Sparse FSOS

Finding a sparse FSOS is equivalent to solving the minimization problems:

$$\min_{f=\sum_i g_i \overline{g_i}} \# \left(\bigcup_i \operatorname{supp}(g_i) \right)$$

$$\Updownarrow$$

$$\min_{Q\text{: Gram matrix}} \| \operatorname{diag}(Q) \|_{\ell^0}.$$

- $\operatorname{diag}(Q)$: diagonal elements of Q.

Convex Relaxation

- ℓ^0-norm optimization $\xrightarrow{\text{relaxation}}$ ℓ^1-norm optimization.

Convex Relaxation

- ℓ^0-norm optimization $\xrightarrow{\text{relaxation}}$ ℓ^1-norm optimization.
- Group case: $\|\operatorname{diag}(Q)\|_{\ell^1} = \text{constant}$:

$$\|\operatorname{diag}(Q)\|_{\ell^1} = \sum_{\chi \in \widehat{G}} Q(\chi, \chi) = \sum_{\chi \in \widehat{G}} Q(\chi, \chi\chi_0) = \widehat{f}(\chi_0).$$

$\chi_0(x) = 1, \forall x \in G$ is the character of the trivial representation of G.

Convex Relaxation

- ℓ^0-norm optimization $\xrightarrow{\text{relaxation}}$ ℓ^1-norm optimization.
- Group case: $\|\operatorname{diag}(Q)\|_{\ell^1} =$ constant:

$$\|\operatorname{diag}(Q)\|_{\ell^1} = \sum_{\chi \in \widehat{G}} Q(\chi,\chi) = \sum_{\chi \in \widehat{G}} Q(\chi,\chi\chi_0) = \widehat{f}(\chi_0).$$

$\chi_0(x) = 1, \forall x \in G$ is the character of the trivial representation of G.

Theorem [Yang,Ye,Zhi'22a]

$$\min_{Q:\ \text{Gram matrix}} \|\operatorname{diag}(Q)\|_{\ell^0} = \min_{\substack{Q(\chi_0,\chi_0)\neq 0\\ Q:\ \text{Gram matrix}}} \#\{\chi \neq \chi_0 : Q(\chi,\chi) \neq 0\} + 1,$$

Convex Relaxation

The convex relaxation problem for

$$\min_{\substack{Q(\chi_0,\chi_0)\neq 0\\ Q:\ \text{Gram matrix}}} \#\{\chi \neq \chi_0 : Q(\chi,\chi) \neq 0\}$$

can be formulated as an SDP problem:

$$\begin{aligned} \min_{Q\in\mathbb{C}^{\widehat{G}\times\widehat{G}}} \quad & \operatorname{trace}(Q) - Q(\chi_0,\chi_0) \\ \text{s.t.} \quad & Q \succeq 0, \quad \sum_{\chi'\in\widehat{G}} Q_{\chi',\chi'\chi} = \widehat{f}(\chi), \quad \forall \chi \in \widehat{G} \end{aligned}$$

Convex Relaxation

The convex relaxation problem for

$$\min_{\substack{Q(\chi_0,\chi_0)\neq 0\\ Q:\ \text{Gram matrix}}} \#\{\chi\neq\chi_0 : Q(\chi,\chi)\neq 0\}$$

can be formulated as an SDP problem:

$$\begin{aligned}\min_{Q\in\mathbb{C}^{\widehat{G}\times\widehat{G}}}\quad & \operatorname{trace}(Q)-Q(\chi_0,\chi_0)\\ \text{s.t.}\quad & Q\succeq 0,\quad \sum_{\chi'\in\widehat{G}} Q_{\chi',\chi'\chi}=\widehat{f}(\chi),\quad \forall\chi\in\widehat{G}\end{aligned}$$

- Complexity of SDP: $\geq O(|G|^4)$.
- Closed form solution[Yang,Ye,Zhi'22a]: $O(|G|\log(|G|))$ (FFT).

Solution to the Relaxed Problem

Theorem [Yang,Ye,Zhi'22a]

Let $f : G \mapsto \mathbb{R}$, $f \geq 0$, h be its square root:

$$h(x) = \sqrt{f(x)} = \sum_{\chi \in \widehat{G}} a_\chi \chi(x), \quad x \in G,$$

and

$$Q_0(\chi, \chi') = \overline{a_\chi} a_{\chi'}.$$

Then Q_0 is a solution of the convex relaxation problem.

Solution to the Relaxed Problem

Theorem [Yang,Ye,Zhi'22a]

Let $f : G \mapsto \mathbb{R}$, $f \geq 0$, h be its square root:

$$h(x) = \sqrt{f(x)} = \sum_{\chi \in \widehat{G}} a_\chi \chi(x), \quad x \in G,$$

and

$$Q_0(\chi, \chi') = \overline{a_\chi} a_{\chi'}.$$

Then Q_0 is a solution of the convex relaxation problem.

- Closed form solution.
- Complexity of FFT: $O(|G| \log(|G|))$

Example of FSOS (again)

- $f : \mathbb{Z}_6 \to \mathbb{C},\ f(x) = 1 - \frac{1}{2}(\chi_1(x) + \chi_5(x)) = 1 - \cos(\frac{2\pi x}{6}),\ \chi_k(x) = e^{\frac{2ik\pi x}{6}}$.
- By the fast Fourier transform, we have

$$\sqrt{f} = \left(\frac{\sqrt{2}}{3} + \frac{\sqrt{6}}{6}\right) - \left(\frac{\sqrt{2}}{12} + \frac{\sqrt{6}}{12}\right)(\chi + \chi^{-1}) + \left(\frac{\sqrt{2}}{12} - \frac{\sqrt{6}}{12}\right)(\chi^2 + \chi^{-2}) + \left(\frac{\sqrt{6}}{6} - \frac{\sqrt{2}}{3}\right)\chi^3.$$

- We obtain a rank one Gram matrix $Q_0 = u^* u$,

$$u = \left[\frac{\sqrt{2}}{3} + \frac{\sqrt{6}}{6} \quad -\frac{\sqrt{2}}{12} - \frac{\sqrt{6}}{12} \quad \frac{\sqrt{2}}{12} - \frac{\sqrt{6}}{12} \quad \frac{\sqrt{6}}{6} - \frac{\sqrt{2}}{3} \quad \frac{\sqrt{2}}{12} - \frac{\sqrt{6}}{12} \quad -\frac{\sqrt{2}}{12} - \frac{\sqrt{6}}{12}\right]$$

- Q_0 is not sparse, its large diagonal elements are useful for FSOS!

Error Bound by Square Root

Theorem [Yang,Ye,Zhi'22a]

Let $f: G \mapsto \mathbb{R}$, $f \geq 0$. Let S be a subset of $\widehat{G}$ s.t. $\mathrm{supp}(f) \subseteq S$, $S = S^{-1}$.

- Define h as the truncation of $\sqrt{f}$ at S, i.e.

$$\widehat{h}(\chi) = \begin{cases} \widehat{\sqrt{f}}(\chi), & \text{if } \chi \in S, \\ 0, & \text{if } \chi \notin S. \end{cases}$$

- The function $f + M$ has an FSOS with support S for

$$M := 2\|\widehat{\sqrt{f}} - \widehat{h}\|_{\ell^1} \cdot \|\widehat{h}\|_{\ell^1} + \|\widehat{\sqrt{f}} - \widehat{h}\|_{\ell^1}^2.$$

Main Steps of Our Algorithm

- Compute Q_0 by FFT, sort diagonal elements of Q_0 by $\sigma \in \mathfrak{S}_{|G|}$

$$Q_0(\chi_{\sigma(1)}, \chi_{\sigma(1)}) \geq Q_0(\chi_{\sigma(2)}, \chi_{\sigma(2)}) \geq \cdots \geq Q_0(\chi_{\sigma(|G|)}, \chi_{\sigma(|G|)}).$$

Main Steps of Our Algorithm

- Compute Q_0 by FFT, sort diagonal elements of Q_0 by $\sigma \in \mathfrak{S}_{|G|}$

$$Q_0(\chi_{\sigma(1)}, \chi_{\sigma(1)}) \geq Q_0(\chi_{\sigma(2)}, \chi_{\sigma(2)}) \geq \cdots \geq Q_0(\chi_{\sigma(|G|)}, \chi_{\sigma(|G|)}).$$

- Find the minimal k s.t $S_k \neq \emptyset$ via an SDP solver, where

$$S_k = \{Q \in \mathbb{C}^{\widehat{G} \times \widehat{G}} : \text{Gram matrix of } f, \text{ with supports } \chi_{\sigma(1)}, \ldots, \chi_{\sigma(k)}\}.$$

Main Steps of Our Algorithm

- Compute Q_0 by FFT, sort diagonal elements of Q_0 by $\sigma \in \mathfrak{S}_{|G|}$

$$Q_0(\chi_{\sigma(1)}, \chi_{\sigma(1)}) \geq Q_0(\chi_{\sigma(2)}, \chi_{\sigma(2)}) \geq \cdots \geq Q_0(\chi_{\sigma(|G|)}, \chi_{\sigma(|G|)}).$$

- Find the minimal k s.t $S_k \neq \emptyset$ via an SDP solver, where

$$S_k = \{Q \in \mathbb{C}^{\widehat{G} \times \widehat{G}} : \text{Gram matrix of } f, \text{ with supports } \chi_{\sigma(1)}, \dots, \chi_{\sigma(k)}\}.$$

Theorem [Yang,Ye,Zhi'22a]

The total complexity of our algorithm with input $f : G \mapsto \mathbb{R}$ is at most

$$\mathrm{O}(|G| \log |G| + \mathrm{poly}(t)),$$

which is quasi-linear in $|G|$, and polynomial in the FSOS sparsity t.

Numerical Experiment: Bounded Degree Case

group	FSOS sparsity	time(s)	bound
$\mathbb{Z}_{10000}$	16.7	1.49	648
$\mathbb{Z}_{20000}$	18.6	2.42	720
$\mathbb{Z}_{30000}$	19	2.82	792
$\mathbb{Z}_{40000}$	17.8	3.03	792
$\mathbb{Z}_{50000}$	18.8	3.38	864
$\mathbb{Z}_{60000}$	19	3.89	864

Table: Bounded degree (not greater than 25)

Theoretical bounds [Fawzi,Saunderson,Parrilo'16]: $3d\log_2(N/d)$.

Numerical Experiment: Product Groups

group	FSOS sparsity	time(s)
$\mathbb{Z}_{500} \times \mathbb{Z}_{500}$	51.4	24.3
$\mathbb{Z}_{1000} \times \mathbb{Z}_{1000}$	50.8	63.3
$\mathbb{Z}_{1500} \times \mathbb{Z}_{1500}$	49.2	123.3
$\mathbb{Z}_{2000} \times \mathbb{Z}_{2000}$	49.6	208.4
$\mathbb{Z}_{2500} \times \mathbb{Z}_{2500}$	50.6	318.6
$\mathbb{Z}_{3000} \times \mathbb{Z}_{3000}$	50.2	457.6
$\mathbb{Z}_{3500} \times \mathbb{Z}_{3500}$	49.2	632.8
$\mathbb{Z}_{4000} \times \mathbb{Z}_{4000}$	49.8	831.7
$\mathbb{Z}_{4500} \times \mathbb{Z}_{4500}$	48.2	1066.0
$\mathbb{Z}_{5000} \times \mathbb{Z}_{5000}$	50.6	1325.1

Table: Bounded FSOS support (at least 10).

URL: github.com/jty-AMSS/FSOS

Outline

Sparse Fourier SOS

Sparse FSOS of Integer Valued Functions

Lower Bounds of Functions on Finite Abelian Groups

Integer-valued Functions over Finite Abelian Groups

- Integer-valued function: $f : G \mapsto \mathbb{Z}$. $\max_{x \in G} |f(x)| = \mathrm{O}(\mathrm{polylog}(|G|))$.

Integer-valued Functions over Finite Abelian Groups

- Integer-valued function: $f : G \mapsto \mathbb{Z}$. $\max_{x \in G} |f(x)| = \mathrm{O}(\mathrm{polylog}(|G|))$.
- Examples: MAXSAT, MAXCUT

Integer-valued Functions over Finite Abelian Groups

- Integer-valued function: $f : G \mapsto \mathbb{Z}$. $\max_{x \in G} |f(x)| = \mathrm{O}(\mathrm{polylog}(|G|))$.
- Examples: MAXSAT, MAXCUT

Integer-valued function $f \geq L,\ L \in \mathbb{Z}$.

$$\Updownarrow$$

$$\max_{x \in G} \left| (f(x) - L) - \sum_{j \in J} |g_j(x)|^2 \right| < 1, \text{ for some } \{g_j\}_{j \in J}.$$

Integer-valued Functions over Finite Abelian Groups

- Integer-valued function: $f : G \mapsto \mathbb{Z}$. $\max_{x \in G} |f(x)| = \mathrm{O}(\mathrm{polylog}(|G|))$.
- Examples: MAXSAT, MAXCUT

Integer-valued function $f \geq L,\ L \in \mathbb{Z}$.

$$\Updownarrow$$

$$\max_{x \in G} \left| (f(x) - L) - \sum_{j \in J} |g_j(x)|^2 \right| < 1, \text{ for some } \{g_j\}_{j \in J}.$$

- We call $\{g_j\}_{j \in J}$ a certificate of $f \geq L$.

FSOS with Error

Theorem [Blekherman,Gouveia,Pfeiffer'16]

$$f(x_1,\ldots,x_n) = \left(\sum_{j=1}^{n} x_j - \left\lfloor \frac{n}{2} \right\rfloor\right)\left(\sum_{j=1}^{n} x_j - \left\lfloor \frac{n}{2} \right\rfloor - 1\right)$$

is non-negative on $\{0,1\}^n$, f has no polynomial or rational FSOS of degree less than $\frac{n-1}{2}$.

FSOS with Error

Theorem [Blekherman,Gouveia,Pfeiffer'16]

$$f(x_1,\ldots,x_n) = \left(\sum_{j=1}^{n} x_j - \left\lfloor \frac{n}{2} \right\rfloor\right)\left(\sum_{j=1}^{n} x_j - \left\lfloor \frac{n}{2} \right\rfloor - 1\right)$$

is non-negative on $\{0,1\}^n$, f has no polynomial or rational FSOS of degree less than $\frac{n-1}{2}$.

- FSOS certificate of degree 1 and sparsity $n+1$ [Yang,Ye,Zhi'22b]:

$$f \text{ integer-valued, } f(x_1,\ldots,x_n) = \left(\sum_{j=1}^{n} x_j - \left\lfloor \frac{n}{2} \right\rfloor - \frac{1}{2}\right)^2 - \frac{1}{4}.$$

$$\Longrightarrow \left| f(x) - \left(\sum_{j=1}^{n} x_j - \left\lfloor \frac{n}{2} \right\rfloor - \frac{1}{2}\right)^2 \right| = \frac{1}{4} < 1,$$

$$\Longrightarrow f \geq 0.$$

Integer-valued Functions over Finite Abelian Groups

- If $G = \mathbb{Z}_2^n$, complexity of FFT: $O(n \cdot 2^n)$.

Integer-valued Functions over Finite Abelian Groups

- If $G = \mathbb{Z}_2^n$, complexity of FFT: $O(n \cdot 2^n)$.
- Question: can we compute the FSOS faster (polynomial in n)?

Polynomial Approximation of $\sqrt{f(x)}$

Suppose

- $0 \le f(x) \le m$, $\|f\|_{\ell^\infty} = \max_{x \in G} |f(x)|$.
- polynomial $q : \mathbb{R} \mapsto \mathbb{R}$ satisfying $|q(x) - \sqrt{x}| < \varepsilon, \forall x \in [0, m]$.

Polynomial Approximation of $\sqrt{f(x)}$

Suppose

- $0 \le f(x) \le m$, $\|f\|_{\ell^\infty} = \max_{x \in G} |f(x)|$.
- polynomial $q : \mathbb{R} \mapsto \mathbb{R}$ satisfying $|q(x) - \sqrt{x}| < \varepsilon, \forall x \in [0, m]$.

Then we have

- $\|q(f(x)) - \sqrt{f(x)}\|_{\ell^\infty} < \varepsilon$, $\left\|q(f(x))^2 - f(x)\right\|_{\ell^\infty} \le 2\sqrt{m}\varepsilon + \varepsilon^2$.

Polynomial Approximation of $\sqrt{f(x)}$

Suppose

- $0 \le f(x) \le m$, $\|f\|_{\ell^\infty} = \max_{x \in G} |f(x)|$.
- polynomial $q : \mathbb{R} \mapsto \mathbb{R}$ satisfying $|q(x) - \sqrt{x}| < \varepsilon, \forall x \in [0, m]$.

Then we have

- $\|q(f(x)) - \sqrt{f(x)}\|_{\ell^\infty} < \varepsilon$, $\left\|q(f(x))^2 - f(x)\right\|_{\ell^\infty} \le 2\sqrt{m}\varepsilon + \varepsilon^2$.
- error of Fourier coefficients satisfies:

$$\left|\widehat{q(f)}(\chi) - \widehat{\sqrt{f}}(\chi)\right| < \varepsilon,\ \forall \chi \in \widehat{G}.$$

Polynomial Approximation of $\sqrt{f(x)}$

Suppose

- $0 \le f(x) \le m$, $\|f\|_{\ell^\infty} = \max_{x\in G} |f(x)|$.
- polynomial $q : \mathbb{R} \mapsto \mathbb{R}$ satisfying $|q(x) - \sqrt{x}| < \varepsilon, \forall x \in [0, m]$.

Then we have

- $\|q(f(x)) - \sqrt{f(x)}\|_{\ell^\infty} < \varepsilon$, $\left\|q(f(x))^2 - f(x)\right\|_{\ell^\infty} \le 2\sqrt{m}\varepsilon + \varepsilon^2$.
- error of Fourier coefficients satisfies:

$$\left|\widehat{q(f)}(\chi) - \widehat{\sqrt{f}}(\chi)\right| < \varepsilon,\ \forall \chi \in \widehat{G}.$$

Remark: sparsity of f is $s \Rightarrow$ sparsity of $q(f) \le s^{\deg q}$.

Low Degree Polynomial Approximation

Theorem [Stoer,Bulirsch'02]

Let $a < b$ be two real numbers. For each $p \in C^{d+1}([a,b])$, we have

$$\max_{t\in[a,b]} |p(t) - q(t)| \leq \left(\frac{b-a}{2}\right)^{d+1} \frac{\max_{y\in[a,b]} |p^{(d+1)}(t)|}{2^d (d+1)!},$$

q is the degree d Chebyshev interpolation polynomial for p on $[a,b]$.

Low Degree Polynomial FSOS Certificate

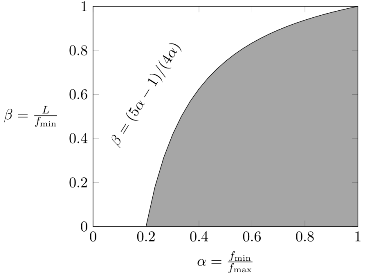
Applying the above result to $p(t) := \sqrt{t}$ on $[\alpha, 1]$ for some $\alpha > 0$, we have

Theorem [Yang,Ye,Zhi'22b]

Let $f : G \mapsto \mathbb{Z}$, if $f_{\max} < (5-4\beta) f_{\min}$, $\beta \in [0,1)$, there exists a rank-one FSOS certificate g, $\| f - L - |g|^2 \|_{\ell^\infty} < 1$ for $f \geq L := \lceil \beta f_{\min} \rceil$ s.t.

$$\deg(g) \leq \deg(f) \left\lfloor \frac{3 + \log(f_{\max} - L)}{2 + \log(f_{\min} - L) - \log(f_{\max} - f_{\min})} \right\rfloor$$

- $\deg(g)$ is bounded by $O(\log(f_{\max}))$.



Impossibility Theorem

- It is impossible to approximate the square root function exponentially and uniformly [Varga'87].

Theorem [Yang,Ye,Zhi'22b]

Let $0 < \varepsilon < 1/6$ and p be a polynomial such that $\left|p(i) - \sqrt{i}\right| \leq \varepsilon$ for each integer $0 \leq i \leq m$, then

$$\deg p \geq \sqrt{\frac{(1-2\varepsilon)m}{1+\sqrt{m}}} = O(m^{\frac{1}{4}}).$$

Low Degree Rational Function Approximation

Let E_d be the approximation error of rational functions $r(t)$ to $\sqrt{t}$ on $[0,1]$:

$$E_d := \inf_{r \in R_d} \left\{ \max_{t\in[0,1]} |\sqrt{t} - r(t)| \right\}.$$

[Vyacheslavov'75,Stahl'03]

There exists some constant $C > 0$ s.t. for each $d \in \mathbb{N}$, it holds that

$$\frac{1}{3} e^{-2\pi\sqrt{\frac{d}{2}}} \leq E_d \leq C e^{-2\pi\sqrt{\frac{d}{2}}}.$$

Lower Degree Rational FSOS Certificate

Theorem [Yang,Ye,Zhi'22b]

Let $f : G \mapsto \mathbb{Z}$, $L \leq f_{\min}$, there exists a rational FSOS certificate (g,h) for $f \geq L$,

$$\|f - L - \frac{|g|^2}{|h|^2}\|_{\ell^\infty} < 1,$$

with

$$\deg g = \deg h \leq \deg(f) \left\lceil \frac{(\log(f_{\max} - L) + c)^2}{2\pi^2} \right\rceil.$$

for some constant $c > 0$.

Lower Degree Rational FSOS Certificate

Theorem [Yang,Ye,Zhi'22b]

Let $f : G \mapsto \mathbb{Z}$, $L \leq f_{\min}$, there exists a rational FSOS certificate (g,h) for $f \geq L$,

$$\|f - L - \frac{|g|^2}{|h|^2}\|_{\ell^\infty} < 1,$$

with

$$\deg g = \deg h \leq \deg(f) \left\lceil \frac{(\log(f_{\max} - L) + c)^2}{2\pi^2} \right\rceil.$$

for some constant $c > 0$.

- Rational SOS: $O(\log^2(f_{\max}))$
- Polynomial SOS : $O((f_{\max})^{\frac{1}{4}})$

Validation by ℓ^1-norm

- $\| f - L - \frac{\sum_{j\in J} |g_j|^2}{\sum_{i\in I} |h_i|^2} \|_{\ell^\infty} < 1$.

Validation by ℓ^1-norm

- $\|f - L - \frac{\sum_{j\in J}|g_j|^2}{\sum_{i\in I}|h_i|^2}\|_{\ell^\infty} < 1$.

Theorem [Yang,Ye,Zhi'22b]

Suppose $h_i, g_j : G \mapsto \mathbb{R}$ satisfy

- Gram matrix V of $\sum_{i\in I}|h_i|^2$, $V \succeq \mathrm{Id}$.
- error function $e(y) = \sum_{j\in J}|g_j(y)|^2 - (\sum_{i\in I}|h_i(y)|^2)\cdot(f(y) - L)$.
- $\|\widehat{e}\|_{\ell^1} \leq \frac{1}{2}$, sum of absolute value of Fourier coefficients is less than $\frac{1}{2}$.

Then $\|f - L - \frac{\sum_{j\in J}|g_j|^2}{\sum_{i\in I}|h_i|^2}\|_{\ell^\infty} \leq \frac{1}{2}$.

Validation by ℓ^1-norm

- $\|f - L - \frac{\sum_{j\in J}|g_j|^2}{\sum_{i\in I}|h_i|^2}\|_{\ell^\infty} < 1$.

Theorem [Yang,Ye,Zhi'22b]

Suppose $h_i, g_j : G \mapsto \mathbb{R}$ satisfy

- Gram matrix V of $\sum_{i\in I}|h_i|^2$, $V \succeq \mathrm{Id}$.
- error function $e(y) = \sum_{j\in J}|g_j(y)|^2 - (\sum_{i\in I}|h_i(y)|^2)\cdot(f(y) - L)$.
- $\|\widehat{e}\|_{\ell^1} \le \frac{1}{2}$, sum of absolute value of Fourier coefficients is less than $\frac{1}{2}$.

Then $\|f - L - \frac{\sum_{j\in J}|g_j|^2}{\sum_{i\in I}|h_i|^2}\|_{\ell^\infty} \le \frac{1}{2}$.

The condition $\|\widehat{e}\|_{\ell^1} \le \frac{1}{2}$ can be verified by solving an SDP problem.

Validation by Sampling on $G = C_2^n$

- $\|f - L - \frac{\sum_{j\in J} g_j^2}{\sum_{i\in I} h_i^2}\|_{\ell^\infty} < 1$.

Validation by Sampling on $G = C_2^n$

- $\|f - L - \frac{\sum_{j\in J} g_j^2}{\sum_{i\in I} h_i^2}\|_{\ell^\infty} < 1$.
- Naive pointwise verification: $O(2^n)$.

Validation by Sampling on $G = C_2^n$

- $\|f - L - \frac{\sum_{j\in J} g_j^2}{\sum_{i\in I} h_i^2}\|_{\ell^\infty} < 1$.
- Naive pointwise verification: $O(2^n)$.

Theorem [Yang,Ye,Zhi'22b]

Let $f:\Gamma_2^n \mapsto \mathbb{Z}$, $(\{g_j\}_{j\in J}, \{h_i\}_{i\in I})$ is a pair of families of functions on Γ_2^n s.t.

(i) $\sum_{i\in I} |h_i|^2 \geq 1$;

(ii) $d := 2\max_{j\in J, i\in I}\{\deg g_j, \deg h_i\}$,$\deg(f) + d \leq n$ ($d = O(\log^2 n)$).

(iii) $\left|\sum_{j\in J} |g_j(y)|^2 - \left(\sum_{i\in I} |h_i(y)|^2\right)(f(y) - L)\right| \leq n^{-2(\deg(f)+d)}$ for $y \in \Gamma_2^n$ s.t.

$$\sum_{i=1}^{n} y_i \leq 2\deg(f) - n + 2d.$$

Then $(\{g_j\}_{j\in J}, \{h_i\}_{i\in I})$ is a rational FSOS certificate for $f \geq L$.

- It is sufficient to check $\mathbf{n}^{\mathbf{O}(\log^2(\mathbf{n}))}$ many inequalities.

Outline

Sparse Fourier SOS

Sparse FSOS of Integer Valued Functions

Lower Bounds of Functions on Finite Abelian Groups

FSOS with Error

Theorem [Yang,Ye,Zhi'23]

- Let $f : G \mapsto \mathbb{R}$, G is a finite abelian group, $S \subseteq \widehat{G}$.
- Let λ be the minimal eigenvalue of Hermitian matrix $Q \in \mathbb{C}^{S\times S}$,

$$e := f - v_S^* Q v_S, \ v_S := (\chi)_{\chi \in S}.$$

We have

- f is bounded below by

$$\min_{x\in G} f(x) \geq -\|\widehat{e}\|_{\ell^1} + \lambda |S|.$$

FSOS with Error

Theorem [Yang,Ye,Zhi'23]

- Let $f : G \mapsto \mathbb{R}$, G is a finite abelian group, $S \subseteq \widehat{G}$.
- Let λ be the minimal eigenvalue of Hermitian matrix $Q \in \mathbb{C}^{S\times S}$,

$$e := f - v_S^* Q v_S, \; v_S := (\chi)_{\chi \in S}.$$

We have

- f is bounded below by

$$\min_{x\in G} f(x) \geq -\|\widehat{e}\|_{\ell^1} + \lambda |S|.$$

- $Q - \lambda \, \mathrm{Id} \succeq 0$ is a Gram matrix of the function $f - e - \lambda |S|$.

Difference between finite abelian group G and $\mathbb{R}$

Let $f(z_1, z_2) = \begin{bmatrix} z_1 & z_2 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = 2z_1z_2$.

- on $C_2^2 = \{-1, 1\}^2$, we have

$$f = |z_1 + z_2|^2 - (|z_1|^2 + |z_2|^2) \geq -(|z_1|^2 + |z_2|^2) \geq -2.$$

- on $\mathbb{R}^2$:

$$f = (z_1 + z_2)^2 - (z_1^2 + z_2^2) \geq \min_{(z_1, z_2) \in \mathbb{R}^2} -(z_1^2 + z_2^2) = -\infty$$

Lower Bound by FSOS

Given $S \subseteq \widehat{G}$, compute a lower bound of $f : G \to \mathbb{R}$ by FSOS

$$\max_{\mathrm{supp}(h_i) \subseteq S} \alpha, \quad \text{s.t. } f - \alpha = \sum_{i \in I} |h_i|^2.$$

It is equivalent to the SDP problem:

$$\begin{aligned} \max\nolimits_{Q \in \mathbb{C}^{S \times S}} \quad & \widehat{f}(\chi_0) - \mathrm{trace}(Q), \\ \text{s.t.} \quad & \sum_{\chi' \in \widehat{G}} Q_{\chi', \chi'\chi} = \widehat{f}(\chi), \ \chi \neq \chi_0 \in \widehat{G} \\ & Q \succeq 0 \end{aligned}$$

Computation of Lower Bounds

- Any Hermitian matrix $Q \in \mathbb{C}^{S\times S}$ gives a lower bound of f,
- We solve the unconstrained optimization problem

$$\max_{Q=Q^*\in\mathbb{C}^{S\times S}} \widehat{f}(\chi_0) - F(Q)$$

where

$$F : \mathbb{C}^{S\times S} \mapsto \mathbb{R},\ F(Q) = \mathrm{trace}(Q) + \|E(Q)\|_{\ell^1} - \lambda_{\mathsf{min}}(Q)|S|,$$

and $E(Q) = \widehat{e - e_0}$, $e := f - v_S^* Q v_S$.

Solving Unconstrained Optimization Problem

Advantages of minimizing F without constraints:

- The subgradient of F is given explicitly by

$$\partial F = \mathrm{Id} + (\partial E)^* \, \mathrm{sign}(E(Q)) - |S| u u^*,$$

  $\mathrm{sign}(x)$: sign function, u: unit eigenvector of Q w.r.t. $\lambda_{\min}(Q)$.
- Early termination;
- Adaptive to more SDP solvers;
- Size reduction.

Random examples

For group $G = C_2^{25}$, C_3^{15}, C_5^{10} randomly generate $f : G \to \mathbb{R}$ with sparsity at least 450 or 200, $f_{min} = 1$.

No	group	sp	Our Algorithm		TSSOS		CS-TSSOS	
			bound	time	bound	time	bound	time
1	C_2^{25}	451	1.00	1058.00	1.00	1027.03	1.00	1451.53
2	C_2^{25}	451	0.67	867.38	-8.72	853.47	-7.23	1483.55
3	C_2^{25}	451	0.75	773.06	1.00	1442.03	1.00	1846.23
4	C_2^{25}	451	0.99	906.15	1.00	1519.08	1.00	1831.33
5	C_2^{25}	451	0.02	718.21	1.00	1364.66	1.00	1710.47
6	C_3^{15}	203	0.97	327.58	–	–	1.00	7336.23
7	C_3^{15}	203	1.00	223.73	–	–	1.00	2876.34
8	C_3^{15}	203	1.00	212.43	–	–	1.00	1353.14
9	C_5^{10}	201	0.97	191.54	–	–	-	-
10	C_5^{10}	201	0.94	236.33	–	–	1.00	559.68
11	C_5^{10}	213	0.90	233.82	–	–	-	-

TSSOS, CS-TSSOS: [Wang,Victor,Lasserre'22].

MAX-2SAT benchmark with 120 variables

No	clause	min	Our Algorithm		TSSOS		CS-TSSOS	
			bound	time	bound	time	bound	time
1	1200	161	159.5	370	146.7	45	146.7	52
2	1200	159	156.7	327	143.1	49	143.1	55
3	1200	160	159.0	362	146.8	46	146.8	64
4	1300	180	177.5	450	162.4	52	162.4	73
5	1300	172	170.6	417	156.2	47	156.2	65
6	1300	173	171.6	432	158.8	44	158.8	58
7	1400	197	194.8	506	179.8	46	179.8	75
8	1400	191	189.3	499	174.3	51	174.3	87
9	1400	189	187.2	504	172.1	58	172.1	78

Table: Unweighted MAX-2SAT problems

TSSOS, CS-TSSOS run out of memory when the relaxation order ≥ 2.

Computation of Feasible Solutions

Rounding by Gram Matrix:

- Compute $v \in \mathbb{R}^S$, $(Q - \lambda_{\min}(Q)\,\mathrm{Id})v = 0,\ v(\chi_0) = 1$.
- Recover $g \in C_2^n$ by $\mathrm{sign}(v)$.

$f(z_1,z_2,z_3) = 4 + z_1 + z_2 + z_3 + z_1z_2z_3 \geq 0$ on C_2^3

- we obtain a Hermitian matrix by SDPNAL+:

$$Q = \begin{array}{c} \\ 1 \\ z_1 \\ z_2 \\ z_3 \\ z_1z_2z_3 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & z_1 & z_2 & z_3 & z_1z_2z_3 \end{array} \\ \begin{bmatrix} 1.9546 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.5000 & 0.4536 & 0.0000 & 0.0000 & 0.0000 \\ 0.5000 & 0.0000 & 0.4536 & 0.0000 & 0.0000 \\ 0.5000 & 0.0000 & 0.0000 & 0.4536 & 0.0000 \\ 0.5000 & 0.0000 & 0.0000 & 0.0000 & 0.4536 \end{bmatrix} \end{array},$$

whose eigenvalues are $-0.0462, 0.4536, 0.4536, 0.4536, 2.4544.$

$f(z_1,z_2,z_3) = 4 + z_1 + z_2 + z_3 + z_1z_2z_3 \geq 0$ on C_2^3

- we obtain a Hermitian matrix by SDPNAL+:

$$Q = \begin{array}{c} \\ 1 \\ z_1 \\ z_2 \\ z_3 \\ z_1z_2z_3 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & z_1 & z_2 & z_3 & z_1z_2z_3 \end{array} \\ \begin{bmatrix} 1.9546 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.5000 & 0.4536 & 0.0000 & 0.0000 & 0.0000 \\ 0.5000 & 0.0000 & 0.4536 & 0.0000 & 0.0000 \\ 0.5000 & 0.0000 & 0.0000 & 0.4536 & 0.0000 \\ 0.5000 & 0.0000 & 0.0000 & 0.0000 & 0.4536 \end{bmatrix} \end{array},$$

  whose eigenvalues are $-0.0462, 0.4536, 0.4536, 0.4536, 2.4544$.

- The normalized null vector of $Q + 0.0462\,\mathrm{Id}$ is

$$v = \begin{array}{c} \begin{array}{ccccc} 1 & z_1 & z_2 & z_3 & z_1z_2z_3 \end{array} \\ \begin{bmatrix} 1 & -1.000447 & -1.000447 & -1.000447 & -1.000447 \end{bmatrix} \end{array}.$$

$f(z_1, z_2, z_3) = 4 + z_1 + z_2 + z_3 + z_1 z_2 z_3 \geq 0$ on C_2^3

- we obtain a Hermitian matrix by SDPNAL+:

$$Q = \begin{array}{c} \\ 1 \\ z_1 \\ z_2 \\ z_3 \\ z_1z_2z_3 \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & z_1 & z_2 & z_3 & z_1z_2z_3 \end{array} \\ \begin{bmatrix} 1.9546 & 0.5000 & 0.5000 & 0.5000 & 0.5000 \\ 0.5000 & 0.4536 & 0.0000 & 0.0000 & 0.0000 \\ 0.5000 & 0.0000 & 0.4536 & 0.0000 & 0.0000 \\ 0.5000 & 0.0000 & 0.0000 & 0.4536 & 0.0000 \\ 0.5000 & 0.0000 & 0.0000 & 0.0000 & 0.4536 \end{bmatrix} \end{array},$$

  whose eigenvalues are $-0.0462, 0.4536, 0.4536, 0.4536, 2.4544.$

- The normalized null vector of $Q + 0.0462\,\mathrm{Id}$ is

$$v = \begin{array}{c} \begin{array}{ccccc} 1 & z_1 & z_2 & z_3 & z_1z_2z_3 \end{array} \\ \begin{bmatrix} 1 & -1.000447 & -1.000447 & -1.000447 & -1.000447 \end{bmatrix} \end{array}.$$

- $f_{min} = 0$ is achieved at $z_1 = z_2 = z_3 = -1$ (rounding elements of v).

MAX-2SAT benchmark with 120 variables.

No	clause	min	Gram	ρ_i^N	$2^{-(i-1)}$
1	1200	161	162	225	227
2	1200	159	159	215	194
3	1200	160	160	162	160
4	1300	180	180	226	243
5	1300	172	173	225	230
6	1300	173	173	245	253
7	1400	197	198	234	270
8	1400	191	192	255	246
9	1400	189	189	227	231

Table: rounding on MAX-2SAT benchmarks

- "Gram" : our rounding methods
- ρ_i^N and $2^{-(i-1)}$: rounding methods in [Maaren,Norden,Heule'08].

References

- Hamza Fawzi, James Saunderson, and Pablo A Parrilo. Sparse sums of squares on finite abelian groups and improved semidefinite lifts. Mathematical Programming, 160(1-2):149-191, 2016.
- Jianting Yang, Ke Ye, and Lihong Zhi. Computing sparse Fourier sum of squares on finite abelian groups in quasi-linear time. arXiv preprint arXiv:2201.03912, 2022.
- Jianting Yang, Ke Ye, and Lihong Zhi. Short certificates for MAX-SAT via Fourier sum of squares. arXiv preprint arXiv:2207.08076, 2022.
- Jianting Yang, Ke Ye, and Lihong Zhi. Lower bounds of functions on finite abelian groups. COCOON 2023.

References

- Grigoriy Blekherman, Joao Gouveia, and James Pfeiffer. Sums of squares on the hypercube. Mathematische Zeitschrift, 284(1):41-54, 2016.
- van Maaren, H., van Norden, L., Heule, M.: Sums of squares based approximation algorithms for max-sat. Discrete Applied Mathematics 156(10), 1754-1779 (2008).
- Jie Wang, Victor Magron, and Jean-Bernard Lasserre. 2021. TSSOS: A Moment-SOS hierarchy that exploits term sparsity. SIAM Journal on Optimization 31, 1 (2021), 30-58.
- Jie Wang, Victor Magron, J. B. Lasserre, and Ngoc Hoang Anh Mai. 2022. CS-TSSOS: Correlative and Term Sparsity for Large-Scale Polynomial Optimization. ACM Trans Math. Softw. 48, 4 (2022), 1-26.

Thank you for your attention!