Applied and Computational Harmonic Analysis 73 (2024) 101686

Contents lists available at ScienceDirect

Applied and Computational Harmonic Analysis

journal homepage: www.elsevier.com/locate/acha

Full Length Article

Computing sparse Fourier sum of squares on finite abelian groups in quasi-linear time ☆

Jianting Yang [a], Ke Ye [b], Lihong Zhi [b,*]

[a] CNRS@CREATE, 1 CREATE Way, Singapore, 138602, Singapore, Singapore
[b] Key Lab of Mathematics Mechanization, AMSS, University of Chinese Academy of Sciences, No.55 Zhongguancun East Road, Beijing, 100190, Beijing, China

ARTICLE INFO

Communicated by Vera Kurkova

Keywords:
Abelian group
Chordal graph
Convex optimization
Fast Fourier transform
Fourier sum of squares
Graph theory
Quasi-linear algorithm
Semidefinite programming
Sparse Gram matrices

ABSTRACT

The problem of verifying the nonnegativity of a function on a finite abelian group is a long-standing challenging problem. The basic representation theory of finite groups indicates that a function f on a finite abelian group G can be written as a linear combination of characters of irreducible representations of G by $f(x) = \sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\chi(x)$, where $\widehat{G}$ is the dual group of G consisting of all characters of G and $\widehat{f}(\chi)$ is the *Fourier coefficient* of f at $\chi \in \widehat{G}$. In this paper, we show that by performing the fast (inverse) Fourier transform, we are able to compute a sparse Fourier sum of squares (FSOS) certificate of f on a finite abelian group G with complexity that is quasi-linear in the order of G and polynomial in the FSOS sparsity of f. Moreover, for a nonnegative function f on a finite abelian group G and a subset $S \subseteq \widehat{G}$, we give a lower bound of the constant M such that $f + M$ admits an FSOS supported on S. We demonstrate the efficiency of the proposed algorithm by numerical experiments on various abelian groups of orders up to 10^7. As applications, we also solve some combinatorial optimization problems and the sum of Hermitian squares (SOHS) problem by sparse FSOS.

1. Introduction

Let X be a finite set and let F be a nonnegative real-valued function on X. This paper concerns the sum of squares (SOS) sparsity of F in the sense of Fourier support [1]. More precisely, if we equip X with an abelian group structure, i.e., we choose a finite abelian group G together with a bijection $\varphi : G \mapsto X$, then $f := F\circ\varphi$ can be naturally identified with an element in the group algebra $\mathbb{C}[G]$. The basic representation theory [2, Chapter 1] of finite groups indicates that the function f can be written as a linear combination of *characters* of irreducible representations of G:

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

☆ Ke Ye and Lihong Zhi are supported by the National Key R&D Program of China 2023YFA1009401 and the National Natural Science Foundation of China 12071467.

* Corresponding author.
E-mail address: lzhi@mmrc.iss.ac.cn (L. Zhi).

https://doi.org/10.1016/j.acha.2024.101686
Received 5 November 2023; Received in revised form 29 May 2024; Accepted 2 July 2024

Available online 10 July 2024

where $\widehat{G}$ is the dual group of G consisting of all characters of G and $\widehat{f}(\chi)$ is the *Fourier coefficient* of f at $\chi \in \widehat{G}$. A subset $S \subseteq \widehat{G}$ containing those $\chi \in \widehat{G}$ such that $\widehat{f}(\chi) \neq 0$ is called a *Fourier support* of f. An *Fourier sum of squares (FSOS) certificate* of a nonnegative function f on G is a finite family $\{g_i\}_{i\in I}$ of functions on G such that

$$f = \sum_{i\in I} |g_i|^2.$$

Clearly, an FSOS certificate of f indeed proves that f is a nonnegative function on G. With the aim of reducing computational complexity, this paper focuses on computing a sparse FSOS certificate. Here, an FSOS certificate $\{g_i\}_{i\in I}$ of f is sparse if there is a small subset $S \subseteq \widehat{G}$ on which each g_i is supported.

In [1,3], based on graph theory, the authors provide interesting theoretical bounds on the sparsity of FSOS. In this paper, we focus on the computational aspect of FSOS. Our main contributions are as follows:

1. We first formulate the problem of computing sparse FSOS as the optimization problem (5). Next we prove in Theorem 3.5 that the square root of f provides a closed-form solution to a properly formulated convex relaxation of (5). Based on that, we design Algorithm 1 to compute a sparse FSOS of a nonnegative function in quasi-linear time. Numerical experiments are presented in Tables 1, 2, 3 to demonstrate the efficiency of our algorithm.
2. In Theorem 3.5, we only select terms of large magnitude in $\sqrt{f}$ to compute a sparse FSOS. We expound the reasons why this heuristic term selection strategy works very well in practice in Theorem 5.4. Furthermore, we show in Proposition 3.6 that the terms selected by our method remain unchanged under the group isomorphism.
3. Applications of FSOS to combinatorial optimization problems and the sum of Hermitian squares (SOHS) problem are presented. Remarkably, we show that one can prove the pigeon-hole principle by an FSOS certificate of sparsity $O(n^2)$ in Proposition 6.2. As a comparison, any resolution refutation requires exponentially many inference steps to prove the pigeon-hole principle [4, Theorem 16, Corollary 18]. In Theorem 6.4, we give sufficient conditions for the existence of a lifting of FSOS on a finite abelian group to SOHS on unit circle in complex plane. Moreover, Example 6.6 indicates that such a lifting may provide a much simpler certificate for the nonnegativity of polynomials on cubes.

We remark that although the graph theoretic approach in [1] can be turned into an algorithm to compute sparse FSOS, it is inefficient since it only uses the Fourier support of the given function. However, more information about FSOS sparsity can be acquired by exploring the structure of coefficients of that function and this observation eventually leads to our Algorithm 1.

The rest of the paper is organized as follows: In Section 3, we establish a proper formulation of the convex relaxation of FSOS sparsity minimization problem. In Section 4, we give an algorithm to compute a sparse FSOS of a given nonnegative function. Numerical experiments are provided to demonstrate the correctness and efficiency of our algorithm. In section 5, we present an error analysis to validate the term selection strategy in the algorithm. In Section 6, we discuss applications of FSOS to combinatorial optimization problems and the SOHS problem.

2. Preliminaries

In this section, we recall some basic definitions and results in representation theory [5,2] and graph theory [6–8].

2.1. Group theory and representation theory

Let G be a finite abelian group. A nonzero complex valued function χ on G is called a *character* of G if it satisfies:

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

The set $\widehat{G}$ of all characters of G is called the *dual group* of G. A subset $S \subseteq \widehat{G}$ is called symmetric if $\chi \in S$ implies $\chi^{-1} \in S$. It is straightforward to verify that $\widehat{G}$ is a finite abelian group, with the group operation given by pointwise multiplication. Since G is a finite abelian group, all irreducible representations of G are one dimensional. Hence we may identify $\widehat{G}$ with the set of all irreducible representations of G.

The fundamental theorem [9] of finite abelian groups implies that

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

For any positive integer d, we also have

$$\widehat{\mathbb{Z}_d} = \left\{ \chi_l(x) := \exp\left(\frac{2i\pi l x}{d}\right), 0 \le l \le d-1 \right\}.$$

Moreover, if G_1, G_2 are finite abelian groups, then $\widehat{G_1 \times G_2} = \widehat{G_1} \times \widehat{G_2}$. Therefore we can regard each $\chi \in \widehat{G}$ as

$$\chi(x_1, \dots, x_k) = \prod_{j=1}^{k} \exp\left(\frac{2i\pi l_j x_j}{n_j}\right), \quad (x_1, \dots, x_k) \in \mathbb{Z}_{n_1} \times \cdots \times \mathbb{Z}_{n_k},$$

for some $0 \le l_j \le n_j - 1, 1 \le j \le k$. Accordingly, χ^{-1} is identified with

$$\chi^{-1}(x_1,\ldots,x_k)=\prod_{j=1}^{k}\exp\left(\frac{-2i\pi l_j x_j}{n_j}\right),\quad (x_1,\ldots,x_k)\in\mathbb{Z}_{n_1}\times\cdots\times\mathbb{Z}_{n_k}.$$

In particular, $\widehat{\mathbb{Z}_2^n}$ consists of square-free monomials in n variables. We have the following theorem for functions on finite abelian groups.

Theorem 2.1. *[2, Chapter 1] Let G be a finite abelian group. Any function $f:G\to\mathbb{C}$ can be uniquely written as a linear combination of elements in $\widehat{G}$, i.e., there is a unique $\widehat{f}:\widehat{G}\to\mathbb{C}$ such that*

$$f=\sum_{\chi\in\widehat{G}}\widehat{f}(\chi)\chi. \tag{1}$$

The unique expansion of f in (1) is called the *Fourier expansion* of f. We define the *support* of f by $\operatorname{supp}(f):=\{\chi:\widehat{f}(\chi)\neq 0\}$. The cardinality of $\operatorname{supp}(f)$ is called the *sparsity of f*.

2.2. Fourier sum of squares of functions on finite abelian groups

In this subsection, we briefly summarize the theory of Fourier sum of squares (FSOS) developed in [1,3]. The definition of FSOS is as follows:

Definition 2.2. [1] Let f be a nonnegative function on finite abelian group G, i.e. $f(x)\geq 0$ for all $x\in G$, then an FSOS representation of f is in form of

$$f=\sum_{i\in I}|g_i|^2. \tag{2}$$

Here $\{g_i\}_{i\in I}$ is a finite family of functions on G, which is called an FSOS certificate of f. Moreover, we say $\{g_i\}_{i\in I}$ is a sparse FSOS certificate of f if $\bigcup_{i\in I}\operatorname{supp}(g_i)$ has small cardinality.

Clearly, an FSOS representation of f provides a certificate of nonnegativity of f, making it of significantly valuable in both mathematics and computer science. The close relationship between the FSOS representation and semidefinite programming problem is highlighted in [1]. The following theorem, stated in [1], elucidates this connection.

Proposition 2.3. *[1, Proposition 1] Let f be a real-valued function on finite abelian group G, then f has an FSOS representation if and only if there exists a Hermitian positive semidefinite matrix $Q\in\mathbb{C}^{|G|\times|G|}$, with rows and columns indexed by $\widehat{G}$, such that*

$$\widehat{f}(\chi)=\sum_{\chi=\chi'^{-1}\chi''}Q(\chi',\chi''),\qquad \chi\in\widehat{G}, \tag{3}$$

where $Q(\chi',\chi'')$ is the element of Q indexed by χ' and χ''. A Hermitian positive semidefinite matrix $Q\in\mathbb{C}^{|G|\times|G|}$ that satisfies the aforementioned conditions is called a Gram matrix *of f.*

Moreover, for any finite abelian group G and any nonnegative function f on G, [1] directly provides the following Gram matrix:

Proposition 2.4. *[1, Proposition 3] Let f be a nonnegative function on finite abelian group G, define the Hermitian matrix $Q\in\mathbb{C}^{|G|\times|G|}$, $Q(\chi,\chi')=\frac{1}{|G|}\widehat{f}(\chi^{-1}\chi')$, then Q is a Gram matrix of f.*

For ease of reference, we record below two known results about upper bound of the FSOS sparsity of nonnegative functions on special groups.

Theorem 2.5. *[1, Theorem 3] Let N,d be positive integers such that d divides N. Then there exists $T\subseteq\widehat{\mathbb{Z}}_N$ with $|T|\leq 3d\log_2(N/d)$ such that any nonnegative function on $\mathbb{Z}_N$ of degree at most d has an FSOS with support T.*

Theorem 2.6. *[3, Theorem 3.2] Every degree r nonnegative polynomial on $\mathbb{Z}_2^n$ has an FSOS certificate of degree $\lceil (n+r-1)/2\rceil$.*

3. The problem of computing sparse FSOS and its convex relaxation

We formulate the problem of computing the sparsest FSOS as the optimization problem (8) and present its convex relaxation (9). We prove in Theorem 3.5 that the point-wise square root of the given function provides a solution to (9), which can be computed in quasi-linear time in $|G|$.

3.1. The FSOS sparsity minimization problem

We recall that Proposition 2.4 gives a Gram matrix

$$Q(\chi,\chi') = \frac{1}{|G|}\widehat{f}(\chi^{-1}\chi'), \quad \chi,\chi' \in \widehat{G}. \tag{4}$$

It is very efficient to compute Q via (4), but one usually gets a dense Gram matrix, as the next example illustrates.

Example 3.1. The following function is considered in [1, Example 4]:

$$f : \mathbb{Z}_6 \to \mathbb{C}, \quad f(x) = 1 - \frac{1}{2}(\chi(x) + \chi^{-1}(x)),$$

where $\chi(x) = \exp\left(\frac{i\pi x}{3}\right)$. The Gram matrix of f found in [1] is:

$$\frac{1}{6}\cdot \begin{array}{c} \\ \chi^0 \\ \chi^1 \\ \chi^2 \\ \chi^3 \\ \chi^4 \\ \chi^5 \end{array} \begin{array}{c} \begin{array}{cccccc} \chi^0 & \chi^1 & \chi^2 & \chi^3 & \chi^4 & \chi^5 \end{array} \\ \begin{bmatrix} 1 & -1/2 & 0 & 0 & 0 & -1/2 \\ -1/2 & 1 & -1/2 & 0 & 0 & 0 \\ 0 & -1/2 & 1 & -1/2 & 0 & 0 \\ 0 & 0 & -1/2 & 1 & -1/2 & 0 \\ 0 & 0 & 0 & -1/2 & 1 & -1/2 \\ -1/2 & 0 & 0 & 0 & -1/2 & 1 \end{bmatrix} \end{array}.$$

By this Gram matrix and techniques from graph theory, one can only get an FSOS with sparsity at least four. However, it turns out that f has another sparse Gram matrix:

$$\frac{1}{6}\cdot \begin{array}{c} \\ \chi^0 \\ \chi^1 \\ \chi^2 \\ \chi^3 \\ \chi^4 \\ \chi^5 \end{array} \begin{array}{c} \begin{array}{cccccc} \chi^0 & \chi^1 & \chi^2 & \chi^3 & \chi^4 & \chi^5 \end{array} \\ \begin{bmatrix} 3 & -3 & 0 & 0 & 0 & 0 \\ -3 & 3 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \end{array},$$

from which one can obtain an FSOS of f with sparsity two: $f(x) = \frac{1}{2}|1-\chi(x)|^2$.

By Proposition 2.3, there is a 1-1 correspondence between (sparse) FSOS of f and (sparse) Gram matrices. Therefore, the problem of computing sparse FSOS can be formulated as the following optimization problem:

$$\min_{Q \text{ is a Gram matrix}} \#\left(\{\chi : \exists \chi', Q(\chi,\chi') \neq 0\} \cup \{\chi' : \exists \chi, Q(\chi,\chi') \neq 0\}\right). \tag{5}$$

Here, we label columns and rows of Q by characters of the group and $Q(\chi,\chi')$ denotes the element of Q labeled by χ and χ'. We denote by $\#S$ the cardinality of a set S.

Lemma 3.2. *Let* $\operatorname{diag}(Q)$ *be the vector consisting of diagonal elements of* Q. *Problem (5) is equivalent to the problem of minimizing the* ℓ^0*-norm of* $\operatorname{diag}(Q)$*, i.e.*

$$\min_{Q\text{: Gram matrix}} \|\operatorname{diag}(Q)\|_0. \tag{6}$$

Proof. As $Q = Q^*$, we have $Q(\chi,\chi') \neq 0$ if and only if $Q(\chi',\chi) \neq 0$. This implies

$$\#\{\chi' : \exists \chi, Q(\chi,\chi') \neq 0\} = \#\{\chi : \exists \chi', Q(\chi,\chi') \neq 0\}.$$

Moreover, Q is positive semidefinite, hence for $\chi,\chi' \in \widehat{G}$, $Q(\chi,\chi') \neq 0$ implies $Q(\chi,\chi) > 0$ and $Q(\chi',\chi') > 0$. Therefore, we have

$$\#\{\chi : \exists \chi', Q(\chi',\chi) \neq 0\} = \#\{\chi : Q(\chi,\chi) \neq 0\},$$

which implies the equivalence between (5) and (6). □

3.2. A convex relaxation

Although the problem of minimizing the ℓ^0-norm is notoriously difficult, one can relax the problem by replacing the ℓ^0-norm by the ℓ^1-norm, which is a widely used method in practice [10,11]. Our method of solving (6) is motivated by this popular method.

However, simply replacing $\|\operatorname{diag}(Q)\|_0$ by $\|\operatorname{diag}(Q)\|_1$ in (6) makes no sense in our situation. Let $f = \sum_{\chi\in\widehat{G}} \widehat{f}(\chi)\chi$ be a nonnegative function on a finite abelian group G, let Q be a Gram matrix of f. Then, for any $\chi' \in \widehat{G}$, we have

$$\sum_{\chi\in\widehat{G}} Q(\chi,\chi\chi') = \widehat{f}(\chi').$$

In particular, we have

$$\sum_{\chi\in\widehat{G}} Q(\chi,\chi) = \sum_{\chi\in\widehat{G}} Q(\chi,\chi\chi_0) = \widehat{f}(\chi_0).$$

Here χ_0 is the identity element in $\widehat{G}$, i.e., $\chi_0(x) = 1$ for all $x \in G$. As Q is positive semidefinite, we have

$$\|\operatorname{diag}(Q)\|_1 = \sum_{\chi\in\widehat{G}} Q(\chi,\chi) = \widehat{f}(\chi_0). \tag{7}$$

The above calculation indicates the following

Lemma 3.3. *Given a finite abelian group G and a nonnegative function f on G, $\|\operatorname{diag}(Q)\|_1$ is a constant value for any Gram matrix of f.*

As a consequence of Lemma 3.3, if we replace $\|\operatorname{diag}(Q)\|_0$ by $\|\operatorname{diag}(Q)\|_1$ directly in (6), then any Gram matrix of f serves as an optimizer of the relaxed problem. Therefore, a more delicate consideration is necessary to alleviate (6). To that end, we observe that if Q is a Gram matrix of f, then for any fixed character $\chi \in \widehat{G}$, the matrix Q_1 defined by

$$Q_1(\chi',\chi'') := Q(\chi\chi',\chi\chi''), \quad \chi',\chi'' \in \widehat{G},$$

is also a Gram matrix of f. Hence it is sufficient to find a sparse Gram matrix Q such that $Q(\chi_0,\chi_0) \neq 0$ and the minimum of $\|\operatorname{diag}(Q)\|_0$ in (6) remains unchanged if we impose this extra condition. Thus we obtain

$$\min_{Q:\ \text{Gram matrix}} \|\operatorname{diag}(Q)\|_0 = \min_{Q:\ \text{Gram matrix}} \#\{\chi \neq \chi_0 : Q(\chi,\chi) \neq 0\} + 1, \tag{8}$$

which is equivalent to (6) in the sense that they have the same optimal value. The convex relaxation problem for the right-hand side optimization problem in (8) is

$$\min_{Q:\ \text{Gram matrix}} \sum_{\chi\neq\chi_0} |Q(\chi,\chi)|,$$

which can be formulated as the following semidefinite programming problem:

$$\begin{aligned} \min_{Q\in\mathbb{C}^{\widehat{G}\times\widehat{G}}} \quad & \langle Q, A\rangle \\ \text{s.t.} \quad & \langle Q, B_\chi\rangle = \widehat{f}(\chi), \quad \chi\in\widehat{G}, \\ & Q \succeq 0, \end{aligned} \tag{9}$$

where

$$A(\chi,\chi') = \begin{cases} 1, & \text{if } \chi = \chi' \neq \chi_0 \\ 0, & \text{otherwise.} \end{cases}, \quad B_\chi(\chi',\chi'') = \begin{cases} 1, & \text{if } \chi'^{-1}\chi'' = \chi \\ 0, & \text{otherwise.} \end{cases}$$

It is clear that

$$\langle Q, A\rangle = \sum_{\chi\neq\chi_0} Q(\chi,\chi).$$

Since $Q \succeq 0$, we have $Q(\chi,\chi) \geq 0$ for each $\chi \in \widehat{G}$ and this implies

$$\sum_{\chi\neq\chi_0} Q(\chi,\chi) = \sum_{\chi\neq\chi_0} |Q(\chi,\chi)|.$$

The conditions $\langle Q, B_\chi\rangle = \widehat{f}(\chi)$ are equivalent to (3), which ensures that Q is indeed a Gram matrix of f. Now we arrive at the following result about the relation between (9) and (8).

Proposition 3.4. *The minimization problem* (9) *is a convex relaxation of* (8). *Moreover,* (9) *is equivalent to*

$$\begin{aligned} \max_{Q\in\mathbb{C}^{\widehat{G}\times\widehat{G}}} \quad & Q(\chi_0,\chi_0) \\ \text{s.t.} \quad & \langle Q, B_\chi\rangle = \widehat{f}(\chi), \quad \forall \chi\in\widehat{G}, \\ & Q\succeq 0, \end{aligned} \tag{10}$$

Here χ_0 *is the identity element in* $\widehat{G}$.

Proof. The first part of the proposition can be verified easily by a direct computation and the second claim follows immediately by recalling the fact that $\sum_{\chi\in\widehat{G}} Q(\chi,\chi)$ is equal to the constant $\widehat{f}(\chi_0)$. □

3.3. A closed form solution to the convex relaxation

By the simple fact that a nonnegative function on a finite abelian group has a square root, we are able to prove that solving (9) is equivalent to computing the square root of f.

Theorem 3.5. *Let* f *be a nonnegative, nonzero function on a finite abelian group* G *and let* h *be its square root defined by*

$$h(x)=\sqrt{f(x)}, \quad x\in G.$$

Suppose that $h=\sum_{\chi\in\widehat{G}} a_\chi\chi$ *is the Fourier expansion of* h *and* Q_0 *is the matrix defined by*

$$Q_0(\chi,\chi')=\overline{a}_\chi a_{\chi'}. \tag{11}$$

Then Q_0 *is a solution of (9).*

Proof. First we prove that Q_0 is a feasible solution of (9). By definition, we observe that $Q_0\succeq 0$ and $\operatorname{rank}(Q_0)=1$. Since for each $x\in G$,

$$f(x)=\left(\sum_{\chi\in\widehat{G}} a_\chi\chi(x)\right)\overline{\left(\sum_{\chi\in\widehat{G}} a_\chi\chi(x)\right)}=\sum_{\chi,\chi'\in\widehat{G}} Q_0(\chi,\chi')\chi^{-1}(x)\chi'(x),$$

Q_0 is a Gram matrix of f of rank 1.

Next, we prove that Q_0 is an optimal solution of (9). To achieve this, we recall from Proposition 3.4 that (9) is equivalent to (10). Thus it suffices to prove that Q_0 an optimal solution of (10).

Claim. *If* $\widetilde{Q}$ *is a feasible solution of* (10) *of rank bigger than one, then*

$$\widetilde{Q}(\chi_0,\chi_0)\le Q_0(\chi_0,\chi_0).$$

We assume for now that the claim holds true. The proof is completed by showing that $\widehat{Q}(\chi_0,\chi_0)\le Q_0(\chi_0,\chi_0)$ for any rank one Gram matrix $\widehat{Q}$. To see this, we write $\widehat{Q}=u^*u$ and denote by g the function corresponding to u. Then we have

$$\widehat{Q}(\chi_0,\chi_0)=\frac{1}{|G|^2}\left|\sum_{x\in G} g(x)\right|^2\le\frac{1}{|G|^2}\left(\sum_{x\in G}|g(x)|\right)^2=\frac{1}{|G|^2}\left(\sum_{x\in G}h(x)\right)^2=Q_0(\chi_0,\chi_0).$$

Here the penultimate equality follows from $|g|^2=f=h^2$.

It is left to prove the claim. We suppose that $\operatorname{rank}(\widetilde{Q})=r\ge 2$. We recall that a spectral decomposition

$$\widetilde{Q}=\sum_{i=1}^r u_i^*u_i$$

leads to an SOS decomposition $f=\sum_{i=1}^r |g_i|^2$, where g_i is the function corresponding to u_i defined by

$$g_i=\sum_{\chi\in\widehat{G}} u_i(\chi)\chi, \quad i=1,\dots,r.$$

In particular, we have

$$u_i(\chi_0)=\langle g_i,\chi_0\rangle=\frac{1}{|G|}\sum_{x\in G} g_i(x),$$

$$\widetilde{Q}(\chi_0,\chi_0)=\sum_{i=1}^{r}|u_i(\chi_0)|^2=\frac{1}{|G|^2}\sum_{i=1}^{r}\left|\sum_{x\in G}g_i(x)\right|^2.$$

By the relation $h^2=f=\sum_{i=1}^{r}|g_i|^2$ and Cauchy-Schwartz inequality, we obtain

$$\begin{aligned}\sum_{i=1}^{r}\left|\sum_{x\in G}g_i(x)\right|^2&=\sum_{i=1}^{r}\sum_{x\in G}|g_i(x)|^2+\sum_{i=1}^{r}\sum_{x\neq y}g_i(x)\overline{g_i}(y)\\&=\left(\sum_{x\in G}h(x)^2\right)+\sum_{x\neq y}\sum_{i=1}^{r}g_i(x)\overline{g_i}(y)\\&\le\left(\sum_{x\in G}h(x)^2\right)+\sum_{x\neq y}\sqrt{\left(\sum_{i=1}^{r}|g_i(x)|^2\right)\left(\sum_{i=1}^{r}|g_i(y)|^2\right)}\\&=\left(\sum_{x\in G}h(x)^2\right)+\sum_{x\neq y}h(x)h(y)\\&=\left|\sum_{x\in G}h(x)\right|^2.\end{aligned}$$

This implies that $\widetilde{Q}(\chi_0,\chi_0)\le Q_0(\chi_0,\chi_0)$ and completes the proof of the claim. □

An important consequence of Theorem 3.5 is that the optimization problem (9) can be solved in $O(|G|\log(|G|))$ time via the fast Fourier transform (FFT) and the inverse fast Fourier transform (iFFT). As an illustrative example, for xqfunction $f(x)=1-\cos(\frac{2\pi x}{6})$ discussed in Example 3.1. By the Fourier transform, we have

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

According to Lemma 3.5, the rank one matrix $Q_0=u^*u$ is a solution to (9), where

$$u=\left[\begin{array}{cccccc}\frac{\sqrt{2}}{3}+\frac{\sqrt{6}}{6} & -\frac{\sqrt{2}}{12}-\frac{\sqrt{6}}{12} & \frac{\sqrt{2}}{12}-\frac{\sqrt{6}}{12} & \frac{\sqrt{6}}{6}-\frac{\sqrt{2}}{3} & \frac{\sqrt{2}}{12}-\frac{\sqrt{6}}{12} & -\frac{\sqrt{2}}{12}-\frac{\sqrt{6}}{12}\end{array}\right].$$

However, it is obvious that the matrix Q_0 is not sparse and actually this is also the case in general.

3.4. Square-root-based basis selection method for sparse FSOS

Although Theorem 3.5 already provides a solution Q_0 of (9), it is not sparse. To obtain a sparse solution of (9), we use the magnitude of the diagonal elements of Q_0 as a heuristic guidance: *if the term $Q_0(\chi,\chi)$ is small, then we search for a Gram matrix $\widetilde{Q}$ such that $\widetilde{Q}(\chi,\chi)=0$*. Our heuristic guidance is inspired by the idea of truncation, which is widely used in scientific computing [12–14] and numerical analysis [15,16]. However, the sparse matrix $\widetilde{Q}$ we search for is an exact Gram matrix of f without error, it is indeed a sparse Gram matrix. We remark that although there is no guarantee on the existence of $\widetilde{Q}$, Theorem 5.4 supplies a rationale for this heuristic guidance. Moreover, our numerical experiments in Section 4 illustrate both the effectiveness and efficiency of the heuristic guidance.

Suppose Q is a solution of (9). We observe that there exists a permutation $\sigma\in\mathfrak{S}_{|G|}$ such that

$$Q(\chi_{\sigma(1)},\chi_{\sigma(1)})\ge Q(\chi_{\sigma(2)},\chi_{\sigma(2)})\ge Q(\chi_{\sigma(3)},\chi_{\sigma(3)})\ge\cdots\ge Q(\chi_{\sigma(|G|)},\chi_{\sigma(|G|)}).\tag{12}$$

Without loss of generality, we may simply assume $\chi_k=\chi_{\sigma(k)}$ in the sequel.

For each $|G|\ge k\ge 1$, we define

$$\mathcal{S}_k=\{Q\in\mathbb{C}^{\widehat{G}\times\widehat{G}}: Q \text{ is Gram matrix of } f,\text{ with } Q(\chi_i,\chi_i)=0 \text{ for all } k\le i\le|G|\}.\tag{13}$$

By definition, each $Q\in\mathcal{S}_k$ is a Gram matrix of sparsity at most $(k-1)$. According to the heuristic guidance, we choose k such that $\mathcal{S}_k$ is nonempty via binary search. The problem of checking the non-emptiness of the convex set $\mathcal{S}_k$ can be solved via an SDP solver in polynomial time in k for fixed accuracy.

We end this subsection by the following proposition on the invariance of the square-root-based basis selection method under a group isomorphism.

Proposition 3.6. *Let G_1,G_2 be isomorphic finite abelian groups and let $\phi: G_2\to G_1$ be an isomorphism. For any nonnegative function f on G_1 and $\chi\in\widehat{G_1}$, we have*

$$\chi\circ\phi\in\widehat{G_2},\quad \widehat{\sqrt{f}}(\chi)=\widehat{\sqrt{f\circ\phi}}(\chi\circ\phi).$$

Moreover, f and $f\circ\phi$ share the same matrix Q_0 given in (11).

Proof. It is clear that $\chi\circ\phi\in\widehat{G_2}$, and

$$\widehat{\sqrt{f\circ\phi}}(\chi\circ\phi)=\frac{1}{|G|}\sum_{x\in G_2}\sqrt{f\circ\phi(x)}\cdot\overline{\chi\circ\phi(x)}=\frac{1}{|G|}\sum_{x\in G_1}\sqrt{f(x)}\cdot\overline{\chi(x)}=\widehat{\sqrt{f}}(\chi).$$

The moreover part follows from the fact that Q_0 is determined by the square root of the function. □

As a comparison, we consider the following example. For any positive integer n, the group isomorphism $\mathbb{Z}_{2\cdot 3^n}\simeq\mathbb{Z}_2\times\mathbb{Z}_{3^n}$ leads to the following isomorphism of rings:

$$\mathbb{C}[x,y]/\langle 1-x^{3^n},1-y^2\rangle\cong\mathbb{C}[\widehat{\mathbb{Z}_2\times\mathbb{Z}_{3^n}}]\cong\mathbb{C}[\widehat{\mathbb{Z}_{2\cdot 3^n}}]\cong\mathbb{C}[x]/\langle 1-x^{2\cdot 3^n}\rangle.$$

This isomorphism indicates that from the perspective of polynomial optimization, checking the nonnegativity of function

$$f(x)=1-\frac{1}{2}(\chi(x)+\chi^{-1}(x)),\quad x\in\mathbb{Z}_{2\cdot 3^n},$$

where $\chi(x)=\exp(x\pi i/3^n)$ is equivalent to checking the nonnegativity of the optimal value of either one of the following two polynomial optimization problems:

$$\begin{aligned}\min_{x\in\mathbb{C}}\quad & 1-\frac{1}{2}x-\frac{1}{2}x^{2\cdot 3^n-1}\\ \text{s.t.}\quad & x^{2\cdot 3^n}=1\end{aligned}$$

and

$$\begin{aligned}\min_{x,y\in\mathbb{C}}\quad & 1-\frac{1}{2}xy^{\frac{3^n+1}{2}}-\frac{1}{2}xy^{\frac{3^n-1}{2}}\\ \text{s.t.}\quad & x^2=y^{3^n}=1.\end{aligned}$$

Although these two problems are essentially same, the bases obtained by degree-based selection method are different. In fact, for any n, the objective function of the former problem always admits a degree-one FSOS

$$1-\frac{1}{2}x-\frac{1}{2}x^{2\cdot 3^n-1}=\frac{1}{2}|1-x|^2=\frac{1}{2}(1-x)^*(1-x),$$

where $x^*=x^{-1}=x^{2\cdot 3^n-1}$, whereas the objective function of the latter problem has no degree-one FSOS if $n>1$.

4. An algorithm for sparse FSOS and numerical experiments

In this section, we present Algorithm 1 to compute a sparse FSOS of a given nonnegative function on a finite abelian group, and then we test Algorithm 1 by different classes of numerical experiments. These experiments are done in Matlab R2016b with CVX package [17,18] and SDPT3 solver [19] on a desktop computer with Intel Core i9-10900X CPU (3.7 GHz). Codes for these examples can be found in https://github.com/jty-AMSS/FSOS. Our numerical examples indicate the following four features of Algorithm 1:

(i) For nonnegative functions with a given degree bound, the cardinality of sparse FSOS is much smaller than the theoretical upper bound for FSOS sparsity given in Theorem 2.5.
(ii) Algorithm 1 can deal with nonnegative functions whose minima are close to zero.
(iii) Algorithm 1 has quasi-linear complexity in the cardinality of the group. In particular, it can deal with nonnegative functions with sparse FSOS on groups of order up to 10^7.
(iv) Algorithm 1 works for arbitrary finite abelian groups. Due to the page limit, we only present examples for $\mathbb{Z}_N$ and $\mathbb{Z}_N\times\mathbb{Z}_N$, but interested readers can find codes and examples for $\mathbb{Z}_2^N$ on https://github.com/jty-AMSS/FSOS.

With all the preparations above, we are ready to present Algorithm 1 which computes a sparse FSOS of a nonnegative function on a finite abelian group. Below we supply some details of this algorithm.

In Step 1, we first perform an inverse Fast Fourier transform (iFFT) on f, then compute the pointwise square root, and finally apply the Fast Fourier transform (FFT) to obtain the Fourier coefficients of $\sqrt{f}$, which can be done in $O(|G|\log(|G|))$ time.

In Step 2, we sort the absolute values of Fourier coefficients instead of the diagonal elements of Q_0 since $Q_0(\chi,\chi)=|a_\chi|^2$ holds for all $\chi\in\hat{G}$. This step can be completed in $O(|G|\log(|G|))$ time.

In Step 3, since f has no FSOS with sparsity less than $\sqrt{s}$, we set the initial value of k equal to $\lceil\sqrt{s}\rceil$.

In Steps 4 and step 5, checking whether S_k is empty and the selection of Q can be done by any SDP solver. Moreover, Steps 4 and 5 respectively require at most $1+2\log(k_{\min}/\sqrt{s})$ and $\log(3k_{\min}/2)$ SDP computations, where each SDP problem is of size at most $2k_{\min}$. In Theorem 5.4, we will show that it is reasonable to add a small perturbation δ to f to obtain a sparse FSOS.

As a consequence, we obtain the following proposition about the complexity of Algorithm 1.

Algorithm 1 Sparse FSOS of a nonnegative function on a finite abelian group.

Input nonnegative function f, finite abelian group G, a small positive number δ
Output sparse FSOS certificate of f on G.
1: compute the Fourier coefficient of square root of f, i.e. compute $\{a_\chi\}_{\chi\in\widehat{G}}$ such that $\sqrt{f}=\sum_{\chi\in\widehat{G}} a_\chi \chi$.
2: Sort $\{a_\chi\}$ in descending order of their absolute values and get $|a_{\chi_1}| \geq |a_{\chi_2}| \geq \cdots \geq |a_{\chi_{|G|}}|$.
3: set $k := \lceil \sqrt{s} \rceil$, where s is the sparsity of f.
4: Test whether S_k in (13) is empty for $f+\delta$. If it is empty, then let $k := 2k$ and recalculate S_k until it is not empty.
5: Find the minimum $k_{\min} \in [1/2k, k]$ by the bisection method such that $S_{k_{\min}}$ is not empty, select $Q \in S_{k_{\min}}$.
6: Compute Cholesky factorization $Q = H^* H$, with the columns of H indexed by $\{\chi_j\}_{j=1}^{k_{\min}}$.
7: **return** $f = \sum_{i=1}^{r} |\sum_{j=1}^{k_{\min}} H(i,\chi_j)\chi_j|^2$.

Table 1
First experiment: bounded degree and bounded minimum.

group	FSOS sparsity	time (s)	bounds [1]
$\mathbb{Z}_{10000}$	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

Proposition 4.1. *The total complexity of Algorithm 1 is at most*

$$\mathrm{O}\left(|G|\log(|G|) + \log(k_{\min})\,\mathrm{SDP}(2k_{\min})\right).$$

Here $\mathrm{SDP}(k)$ *denotes the complexity of solving the SDP problem of size* $k \times k$*.*

According to [20], $\mathrm{SDP}(k) = \mathrm{O}(k^6)$ for fixed accuracy, thus the total complexity is quasi-linear in the order of G and polynomial in $k_{\min}$. Moreover, if the FSOS sparsity $k_{\min} = \mathrm{O}(|G|^{1/6})$, then the complexity of Algorithm 1 applying to f is $\mathrm{O}(|G|\log(|G|))$.

4.1. Experiments on $\mathbb{Z}_N$

We present numerical results on randomly generated functions on groups of the form $\mathbb{Z}_N$.

4.1.1. The first experiment on $\mathbb{Z}_N$

In the first experiment, we randomly pick 10 functions $h_i, i = 1, \ldots, 10$, satisfying

(i) $\deg(h_i) \leq 24$.
(ii) $0 < \min_{x\in G} h_i(x) < 1$.
(iii) $\mathrm{Re}(\widehat{h}_i(\chi)), \mathrm{Im}(\widehat{h}_i(\chi)) \in [-10, 10]$ for all $\chi \in \widehat{\mathbb{Z}_N}$.

For each $1 \leq i \leq 10$, we apply Algorithm 1 to h_i separately. We record results in Table 1 in which the second column shows the average cardinality of the sparsity of FSOS found by Algorithm 1 for h_i, the third column is the average time cost of Algorithm 1 and the last column is the theoretical upper bound $3d\log_2(N/d)$ for the FSOS sparsity given by Theorem 2.5. Since the theoretical bound given by Theorem 2.5 (last column) is for arbitrary nonnegative functions, it is much larger than our computed FSOS sparsity (second column).

4.1.2. The second experiment on $\mathbb{Z}_N$

In the second experiment, we randomly pick 10 functions $h_i, i = 1, \ldots, 10$ satisfying

(i) $|\mathrm{supp}(h_i)| \leq 25$.
(ii) $0 < \min_{x\in G}(h_i(x)) < 1$.
(iii) $\mathrm{Re}(\widehat{h}_i(\chi)), \mathrm{Im}(\widehat{h}_i(\chi)) \in [-1, 1]$ for all $\chi \neq \chi_0$.

For $1 \leq i \leq 10$, we execute Algorithm 1 separately for each h_i and record the sparsity and running time, the results are summarized in Table 2. Here the second column shows the average sparsity of FSOS certificates found by Algorithm 1, the third column is the average time cost of Algorithm 1. Notice that in the first experiment, we impose a degree bound to compare the computed FSOS sparsity and its theoretical upper bound. In this experiment, we impose a bound on the cardinality of the support of h_i to test the performance of Algorithm 1 on nonnegative functions whose minimum values are close to zero.

Table 2
Second experiment: bounded support and bounded minimum.

group	FSOS sparsity	time (s)
$\mathbb{Z}_{10000}$	110.4	224.3
$\mathbb{Z}_{20000}$	144.4	1074.6
$\mathbb{Z}_{30000}$	168.2	2273.7
$\mathbb{Z}_{40000}$	204.8	3869.0
$\mathbb{Z}_{50000}$	195.8	3314.3
$\mathbb{Z}_{60000}$	219.7	4594.6

Table 3
Third experiment: bounded FSOS support.

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

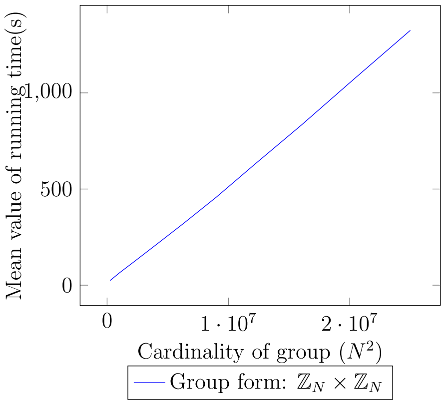

Fig. 1. Time complexity.

4.2. *Experiment on* $\mathbb{Z}_N \times \mathbb{Z}_N$

We carry out an experiment on groups of the form $\mathbb{Z}_N \times \mathbb{Z}_N$.

In this experiment, we randomly choose a subset $T \subseteq \widehat{\mathbb{Z}_N \times \mathbb{Z}_N}$ with $|T| = 10$ and randomly choose 10 real-valued functions g_j, $j = 1, \dots, 10$ on $\mathbb{Z}_N \times \mathbb{Z}_N$ satisfying:

(i) $\operatorname{supp}(g_j) \subseteq T$.
(ii) $\operatorname{Re}(\widehat{g}_j(\chi)), \operatorname{Im}(\widehat{g}_j(\chi)) \in [-10, 10]$ for each $\chi \in T$.

We apply Algorithm 1 to find a sparse FSOS certificate of $f = \sum_{j=1}^{10} |g_j|^2$. For each value of N, we repeat the experiment 10 times, and record the sparsity and the running time. In Table 3, we record numerical results for different values of N. The second column shows the mean sparsity of FSOS certificates found by Algorithm 1. The third column is the mean time cost for each example. In Fig. 1, we plot the running time cost of Algorithm 1 versus the group size N^2. It is notable that:

(i) The FSOS certificate found by Algorithm 1 always has cardinality around 50. This is because the FSOS sparsity of f is at most 10, which can be seen from its construction.
(ii) Algorithm 1 can work for groups of size up to $2.5 \cdot 10^7$.
(iii) Fig. 1 roughly fits $N \log(N)$ since in these examples, the factor k appeared in Proposition 4.1 can be regarded as a constant.

5. Bounds on FSOS sparsities

In this section we focus on bounding the FSOS sparsity of a nonnegative function. In Subsection 5.1, we prove that for functions with dominating constant terms, their FSOS sparsities are bounded by their Fourier sparsities. In Subsection 5.2, we prove that a suitable perturbation of the given function admits an FSOS supported on the Fourier support of the original function. As a consequence, we obtain an error analysis of step 4 in Algorithm 1 which provides a rationale for the square-root-based basis selection method we proposed in Subsection 3.4.

5.1. FSOS sparsities of functions with dominating constant terms

Let f be a real-valued function on G. We observe that for any $\chi \in \widehat{G}$, we have

$$\widehat{f}(\chi^{-1}) = \overline{\widehat{f}(\chi)}. \tag{14}$$

Indeed, we recall that $\widehat{f}(\chi)$ and $\widehat{f}(\chi^{-1})$ can be computed by

$$\widehat{f}(\chi) = \frac{1}{|G|}\sum_{x\in G}\overline{\chi(x)}f(x), \quad \widehat{f}(\chi^{-1}) = \frac{1}{|G|}\sum_{x\in G}\overline{\chi^{-1}(x)}f(x). \tag{15}$$

Since G is abelian, we have $\chi^{-1} = \overline{\chi}$. This together with (15) and the assumption that f is real-valued implies (14). In the following we bound the FSOS sparsity of a nonnegative function with dominating constant term.

Proposition 5.1. *Let G be a finite abelian group and let S be a symmetric subset of $\widehat{G}$, $f = \sum_{\chi\in S}\widehat{f}(\chi)\chi$ be a real-valued function on G such that $\widehat{f}(\chi_0) \ge \sum_{\chi\neq\chi_0}|\widehat{f}(\chi)|$, then the FSOS sparsity of f is at most $|S|$.*

Proof. It is sufficient to prove that f has a Gram matrix whose number of nonzero rows is at most $|S|$. Let $S = \{\chi_0, \chi_1, \dots, \chi_k\}$. Here χ_0 is the trivial character. We define $Q \in \mathbb{C}^{\widehat{G}\times\widehat{G}}$ by:

$$Q(\chi,\chi') = \begin{cases} \widehat{f}(\chi_0) - \frac{1}{2}\sum_{\chi\neq\chi_0}|\widehat{f}(\chi)|, \text{ if } \chi = \chi' = \chi_0, \\ \frac{1}{2}\widehat{f}(\chi'), \text{ if } \chi = \chi_0 \neq \chi', \\ \frac{1}{2}\widehat{f}(\chi^{-1}), \text{ if } \chi' = \chi_0 \neq \chi, \\ \frac{1}{2}|\widehat{f}(\chi)|, \text{ if } \chi = \chi' \neq \chi_0, \\ 0, \text{ otherwise.} \end{cases}$$

Obviously nonzero elements of Q are on rows and columns labeled by characters in S. It is straightforward to verify by definition that Q is a Gram matrix of f. Moreover, Q is an arrowhead matrix, i.e., nonzero elements of Q either lie in the diagonal or in the first row or first column.

It is left to prove $Q \succeq 0$. It is sufficient to verify the nonnegativity of each principal minor of Q. Since Q is an arrowhead matrix, a principal submatrix of Q not involving elements in the first row and column is a diagonal matrix. Thus it is straightforward to verify that the determinant of such a principal submatrix is nonnegative.

If a principal submatrix of Q contains the first row and column of Q, then we may compute its determinant directly. Indeed, the determinant of $Q(S,S)$ is

$$\det(Q(S,S)) = \left(Q(\chi_0,\chi_0) - \sum_{\chi\neq\chi_0\in S}\frac{Q(\chi_0,\chi)Q(\chi,\chi_0)}{Q(\chi,\chi)}\right)\prod_{\chi\neq\chi_0\in S}Q(\chi,\chi),$$

where $Q(S,S)$ denotes the principal submatrix of Q obtained by selecting rows and columns labeled by characters in $S \subseteq \widehat{G}$. For each $\chi \in \widehat{G}$ we have

$$Q(\chi_0,\chi) = \frac{1}{2}\widehat{f}(\chi) = \frac{1}{2}\overline{\widehat{f}(\chi^{-1})} = \overline{Q(\chi,\chi_0)}.$$

Here the second equality follows from (14). Since f has a dominating constant term, we have

$$Q(\chi_0,\chi_0) - \sum_{\chi\neq\chi_0\in S}\frac{Q(\chi_0,\chi)Q(\chi,\chi_0)}{Q(\chi,\chi)} = \widehat{f}(\chi_0) - \sum_{\chi\neq\chi_0}|\widehat{f}(\chi)| \ge 0.$$

The nonnegativity of other principal minors of Q can be proved similarly and this completes the proof. □

Remark 5.2. We recall that Theorem 2.5 supplies an upper bound $3d\log_2(N/d)$ for a nonnegative function f on $\mathbb{Z}_N$, where d is the degree of f. But according to Proposition 5.1, a degree d nonnegative function with dominating $\widehat{f}(\chi)$ on $\mathbb{Z}_N$ has FSOS sparsity at most $(2d+1)$, which is much smaller than $3d\log_2(N/d)$ for large N.

5.2. FSOS support of a perturbation

Let G be a finite abelian group and let f be a function on G. We define the ℓ^1 norm of $\widehat{f}$ as

$$\|\widehat{f}\|_{\ell^1} := \sum_{\chi\in\widehat{G}} |\widehat{f}(\chi)|. \tag{16}$$

We begin with the following lemma, which can be regard as an improvement of Lemma 1 in [21] with finite abelian group constraints.

Lemma 5.3. *Let G be a finite abelian group and let f be a real-valued function on G. If S is a subset of $\widehat{G}$ such that*

$$\operatorname{supp}(f)\subseteq S\cdot S^{-1} := \left\{\chi'\chi^{-1} : \chi,\chi'\in S\right\},$$

then $f+\|\widehat{f}\|_{\ell^1}$ admits an FSOS with support S.

Proof. It is clear that $S\cdot S^{-1}$ is a symmetric set. Let S_0 be the set of symmetric elements in $S\cdot S^{-1}$,

$$S_0 := \left\{\chi\in S\cdot S^{-1} : \chi^{-1}=\chi\right\},$$

then we can decompose the set $S\cdot S^{-1}\backslash S_0 := \left\{\chi\in S\cdot S^{-1} : \chi\notin S_0\right\}$ into two disjoint sets:

$$S\cdot S^{-1}\backslash S_0 = S_1\cup S_{-1},\quad S_{-1}=S_1^{-1}=\left\{\chi^{-1} : \chi\in S_1\right\},$$

and there are no mutually inverse elements in the set S_1. Since f is a real-valued function, (14) implies that $\widehat{f}(\chi^{-1})=\overline{\widehat{f}(\chi)}$ for any $\chi\in\widehat{G}$. Hence f can be expressed in the following form:

$$f=\sum_{\chi\in S_1}\widehat{f}(\chi)\chi+\sum_{\chi\in S_1}\overline{\widehat{f}(\chi)}\chi^{-1}+\sum_{\chi\in S_0}\widehat{f}(\chi)\chi,$$

and

$$\|\widehat{f}\|_{\ell^1}=2\sum_{\chi\in S_1}|\widehat{f}(\chi)|+\sum_{\chi\in S_0}|\overline{\widehat{f}(\chi)}|.$$

Since S_1 and S_0 are subsets of $S\cdot S^{-1}$, we have that for any $\chi\in S_0\cup S_1$, there exists $\chi',\chi''\in S$ such that $\chi=\chi''\chi'^{-1}$. Then we have

$$\|f\|_{\ell^1}+f=\sum_{\chi\in S_1}\left(2|\widehat{f}(\chi)|+\widehat{f}(\chi)\chi+\overline{\widehat{f}(\chi)}\chi^{-1}\right)+\sum_{\chi\in S_0}\left(|\widehat{f}(\chi)|+\widehat{f}(\chi)\chi\right).$$

We show below that $\|f\|_{\ell^1}+f$ has an FSOS with support S:

- For $\chi=\chi''\cdot\chi'^{-1}\in S_1$, $2|\widehat{f}(\chi)|+\widehat{f}(\chi)\chi+\overline{\widehat{f}(\chi)}\chi^{-1}=|\widehat{f}(\chi)|\left|\chi'+\frac{\widehat{f}(\chi)}{|\widehat{f}(\chi)|}\chi''\right|^2$.
- For $\chi=\chi''\cdot\chi'^{-1}\in S_0$, since $\overline{f}=f$, $\widehat{f}(\chi)=\overline{\widehat{f}(\chi)}\in\mathbb{R}$, we have

$$|\widehat{f}(\chi)|+\widehat{f}(\chi)\chi=\frac{|\widehat{f}(\chi)|}{2}|1+\chi|^2=\frac{|\widehat{f}(\chi)|}{2}\left|\chi''+\frac{\widehat{f}(\chi)}{|\widehat{f}(\chi)|}\chi'\right|^2.\quad\square$$

Based on Lemma 5.3, we present the following theorem to show that a perturbation of f has an FSOS supported in a given support.

Theorem 5.4. *Let G be a finite abelian group and let $f\neq 0$ be a nonnegative function on G. Assume that S is a subset of $\widehat{G}$ such that $\operatorname{supp}(f)\subseteq S$ and $S=S^{-1}$. We define h by*

$$\widehat{h}(\chi)=\begin{cases}\widehat{\sqrt{f}}(\chi), & \text{if } \chi\in S,\\ 0, & \text{if } \chi\notin S.\end{cases} \tag{17}$$

Then $f+M$ has an FSOS supported in S for

$$M := 2\|\widehat{\sqrt{f}}-\widehat{h}\|_{\ell^1}\cdot\|\widehat{h}\|_{\ell^1}+\|\widehat{\sqrt{f}}-\widehat{h}\|_{\ell^1}^2. \tag{18}$$

Before we proceed to the proof of the theorem, we remark that Theorem 5.4 actually provides a rationale of the square-root-based basis selection method proposed in Subsection 3.4. Indeed, we can regard the function h in Theorem 5.4 as the truncation of $\sqrt{f}$

by S. Suppose S consists of characters with large coefficients in the Fourier expansion of $\sqrt{f}$. Then $\sqrt{f} - h$ only contains terms of $\sqrt{f}$ with small coefficients. Thus $\|\widehat{\sqrt{f}} - \hat{h}\|_{\ell^1}$ and M defined in (18) are small. As a consequence, Theorem 5.4 implies that a small perturbation $f + M$ admits an FSOS supported in S.

Proof. Define $g := f - |h|^2$, since $\operatorname{supp}(f) \subseteq S$, $f \neq 0$, $\hat{f}(\chi_0) = \frac{1}{|G|}\sum_{x\in G} f(x) > 0$. Hence we have

$$\chi_0 \in S, \quad \operatorname{supp}(f) \subseteq S \subseteq S \cdot S.$$

Moreover, we have

$$\operatorname{supp}(h^2) \subseteq S \cdot S := \{\chi \cdot \chi' : \chi, \chi' \in S\} = S \cdot S^{-1}.$$

Since $S^{-1} = S$ and $\sqrt{f}$ is real-valued, we have $\overline{\hat{h}(\chi)} = \hat{h}(\chi^{-1})$ holds for all $\chi \in \widehat{G}$, thus h is also a real-valued function and $|h|^2 = h^2$. We denote $S' := \widehat{G} \setminus S$. According to (17), we have

$$g := f - |h|^2 = \left(\sum_{\chi'\in S'} \widehat{\sqrt{f}}(\chi')\chi'\right)^2 + 2\cdot\left(\sum_{\chi\in S}\sum_{\chi'\in S'} \widehat{\sqrt{f}}(\chi)\widehat{\sqrt{f}}(\chi')\chi\chi'\right).$$

We have

$$\|\hat{g}\|_{\ell^1} \leq \left(\sum_{\chi'\in S'} \left|\widehat{\sqrt{f}}(\chi')\right|\right)^2 + 2\left(\sum_{\chi\in S}\sum_{\chi'\in S'} \left|\widehat{\sqrt{f}}(\chi)\widehat{\sqrt{f}}(\chi')\right|\right). \tag{19}$$

Let

$$M := 2\|\widehat{\sqrt{f}} - \hat{h}\|_{\ell^1} \cdot \|\hat{h}\|_{\ell^1} + \|\widehat{\sqrt{f}} - \hat{h}\|_{\ell^1}^2 = 2\left(\sum_{\chi\in S}\sum_{\chi'\in S'} \left|\widehat{\sqrt{f}}(\chi)\widehat{\sqrt{f}}(\chi')\right|\right) + \left(\sum_{\chi'\in S'} \left|\widehat{\sqrt{f}}(\chi')\right|\right)^2. \tag{20}$$

By (19) and (20), we have

$$\|\hat{g}\|_{\ell^1} \leq M.$$

By Lemma 5.3, $M + g$ has an FSOS with support S. Hence $f + M = |h|^2 + (M + g)$ has an FSOS with support S. □

The following corollary shows that for a given s and function f, an appropriate perturbation of f admits an FSOS of sparsity at most s.

Corollary 5.5. *Let G be a finite abelian group and let $f \neq 0$ be a nonnegative function on G. Then for any positive integer $s > 1 + |\operatorname{supp}(f)|$, $f + M_{s'}$ admits an FSOS with sparsity at most s, where $s' = s - 1 - |\operatorname{supp}(f)|$ and*

$$M_{s'} := \|\widehat{\sqrt{f}}\|_{\ell^1}^2 \left(3 - 4\frac{s'}{|G|} + \frac{s'^2}{|G|^2}\right).$$

Proof. Since $\sqrt{f}$ is nonnegative, $|\widehat{\sqrt{f}}(\chi)| = |\widehat{\sqrt{f}}(\chi^{-1})|$ for all $\chi \in \widehat{G}$. Without loss of generality, we can arrange $\widehat{G}$ in descending order of their absolute values in $\widehat{\sqrt{f}}$, with any two mutually inverse elements always adjacent i.e. $|\widehat{\sqrt{f}}(\chi_1)| \geq |\widehat{\sqrt{f}}(\chi_2)| \geq \cdots \geq |\widehat{\sqrt{f}}(\chi_{|G|})|$, and for any integer $i > 0$, χ_i^{-1} must be one of the three characters χ_{i-1}, χ_i or χ_{i+1}. Then for any integer $k > 0$, the sum of the first k largest coefficients satisfies

$$\sum_{i=1}^{k} |\widehat{\sqrt{f}}(\chi_i)| \geq k\frac{\|\widehat{\sqrt{f}}\|_{\ell^1}}{|G|}.$$

In this case, either the set $\{\chi_1, \chi_2, \cdots \chi_s\}$ is symmetric, or the set $\{\chi_1, \chi_2, \cdots \chi_s, \chi_{s+1}\}$ is symmetric. Let S' be the symmetric set among these two sets, $S = \operatorname{supp}(f) \cup S'$, and let h be the truncation of $\sqrt{f}$ at set S defined in (17). Then we have

$$\|\widehat{\sqrt{f}} - \hat{h}\|_{\ell^1} \leq \|\widehat{\sqrt{f}}\|_{\ell^1} - s'\frac{\|\widehat{\sqrt{f}}\|_{\ell^1}}{|G|}.$$

According to (20), we have

$$M \leq 2\left(\|\widehat{\sqrt{f}}\|_{\ell^1}^2 - s'\frac{\|\widehat{\sqrt{f}}\|_{\ell^1}^2}{|G|}\right) + \left(\|\widehat{\sqrt{f}}\|_{\ell^1} - s'\frac{\|\widehat{\sqrt{f}}\|_{\ell^1}}{|G|}\right)^2 = \|\widehat{\sqrt{f}}\|_{\ell^1}^2 \left(3 - 4\frac{s'}{|G|} + \frac{s'^2}{|G|^2}\right) = M_{s'}.$$

By Theorem 5.4. we can conclude that $f + M_{s'}$ admits an FSOS with sparsity at most s. □

We can estimate $\|\widehat{\sqrt{f}}\|_{\ell^1}$ for $0 \le f \le 1$. The Fourier coefficient $\hat{f}(\chi_0)$ equals to $\sum_{\chi\in\hat{G}} |\widehat{\sqrt{f}}(\chi)|^2$ as

$$\hat{f}(\chi_0) = \frac{1}{|G|}\sum_{x\in G} f(x).$$

Hence, we have $0 \le \sum_{\chi\in\hat{G}} |\widehat{\sqrt{f}}(\chi)|^2 \le 1$ and $\|\widehat{\sqrt{f}}\|_{\ell^1} = \sum_{\chi\in\hat{G}} |\widehat{\sqrt{f}}(\chi)| \le \sqrt{|G|}$.

We remark that Theorem 5.4 and Corollary 5.5 only depend on the coefficients of $\sqrt{f}$ and the cardinality of G, regardless of the degree of f and the structure of G. Therefore, this result can be regarded as a complement to results in [22] for functions with high degree.

6. Applications of FSOS

We conclude this paper by a discussion on applications of FSOS in combinatorial optimization problems and sum of Hermitian squares of polynomials on tori.

6.1. Combinatorial optimization

Combinatorial optimization is a very natural resource of applications of FSOS [23–28].

6.1.1. Certificate problem for MAX-SAT

It has been proved in [29] that FSOS supply short certificates for MAX-SAT, MIN-SAT and UNSAT problems. In addition, MAX-2SAT and MAX-3SAT problems can be solved by optimizing polynomials on $\mathbb{Z}_2^n = \{-1, 1\}^n$. In order to reduce the size of the related SDP problems, some choices of monomial bases are proposed [30]. These monomial bases perform well on some benchmark problems but poorly on others. As an example, we consider the weighted MAX-2SAT problem corresponding to the function $g : \mathbb{Z}_2^{10} \to \mathbb{Z}$:

$$\begin{aligned} g = \ & 50450 + 234x_3 - 1386x_2 - 1389x_1 + 502x_4 + 3056x_5 - 4692x_6 - 2142x_7 - 1312x_8 \\ & -4645x_9 + 3787x_{10} - 3399x_1x_2 - 1140x_1x_3 - 282x_2x_3 - 2413x_1x_5 - 884x_2x_4 \\ & -2212x_1x_6 + 3457x_2x_5 + 4462x_3x_4 - 2002x_5x_{10} + 2057x_3x_9 + 4097x_1x_7 + 1707x_2x_6 \\ & +3419x_1x_8 - 4102x_2x_7 - 976x_3x_6 - 2403x_4x_5 - 1245x_1x_9 - 3786x_2x_8 - 1122x_6x_7 \\ & +1014x_3x_7 + 3139x_4x_6 + 483x_1x_{10} + 4417x_2x_9 - 854x_3x_8 - 2037x_5x_6 - 1678x_2x_{10} \\ & +667x_6x_8 - 491x_1x_4 - 981x_4x_8 + 4848x_5x_7 + 4085x_3x_{10} + 1129x_4x_9 - 4936x_5x_8 \\ & -2628x_4x_{10} + 2787x_5x_9 - 936x_3x_5 + 640x_6x_9 + 1874x_7x_8 - 707x_6x_{10} + 778x_7x_9 \\ & +3813x_7x_{10} - 2764x_8x_9 + 3038x_8x_{10} + 2170x_9x_{10} + 6x_4x_7. \end{aligned}$$

The monomial basis suggested by [30] consists of all monomials of degree at most 2, i.e.

$$M_{ap} = \{1\} \cup \{x_i : i = 1, \dots, 10\} \cup \{x_i x_j : 1 \le i < j \le 10\}.$$

It can be checked that there exists no SOS supported in M_{ap}. However, our algorithm produces an FSOS of g of sparsity $52 < |M_{ap}| = 56$. More examples can be found in [31,32].

6.1.2. The pigeon-hole principle

In the following, we consider a more remarkable application in combinatorial optimization: the proof complexity of the pigeon-hole principle. First of all, we recall that the pigeon-hole principle says that $n+1$ pigeons cannot be placed into n holes unless a hole contains more than one pigeon. For each positive integer n, we define a conjunctive normal form (CNF) formula in $(n+1)n$ variables $\{p_{ij}\}_{i=1,j=1}^{n+1,n}$:

$$\overline{\mathrm{PHP}_n^{n+1}} := \bigwedge_{i=1}^{n+1} (p_{i1} \vee \cdots \vee p_{in}) \wedge \bigwedge_{1\le i<k\le n+1, 1\le j\le n} (\neg p_{ij} \vee \neg p_{kj}).$$

Then the pigeon-hole principle is equivalent to the statement that $\overline{\mathrm{PHP}_n^{n+1}}$ is unsatisfiable for all $n \in \mathbb{N}$. The standard technique to prove the unsatisfiability of a CNF formula is the resolution refutation [33,4]. According to the theorem that follows, proving the pigeon-hole principle by resolution refutation is difficult.

Theorem 6.1. *[4, Theorem 16, Corollary 18] For sufficiently large n, any resolution refutation of* $\overline{\mathrm{PHP}_n^{n+1}}$ *requires* $2^{n/20}$ *inference steps.*

It turns out that, however, we are able to prove the pigeon-hole principle by an FSOS certificate of sparsity $O(n^2)$. To this end, we define:

$$p_n : \mathbb{Z}_n^{n+1} \mapsto \mathbb{C}, \quad p_n(x_1, \dots, x_{n+1}) = \sum_{1 \le i < j \le n+1} \mathrm{Eqv}(x_i, x_j),$$

where

$$\mathrm{Eqv} : \mathbb{Z}_n^2 \mapsto \mathbb{C}, \quad \mathrm{Eqv}(x, y) = \begin{cases} 1 & \text{if } x = y, \\ 0 & \text{otherwise.} \end{cases}$$

Proposition 6.2. *For each $n \in \mathbb{N}$, the unsatisfiability of $\overline{\mathrm{PHP}_n^{n+1}}$ is equivalent to the positivity of p_n. Moreover, p_n admits an FSOS of sparsity at most $O(n^2)$.*

Proof. Let $\chi_k(x) = \exp\left(\frac{2\pi i k x}{n}\right)$ and let NOR be the function defined on $\mathbb{Z}_n$:

$$\mathrm{NOR}(x) = \begin{cases} 1 & \text{if } x = 0, \\ 0 & \text{otherwise.} \end{cases}$$

By the inner product $\langle \chi_k, \mathrm{NOR} \rangle = \sum_{x \in \mathbb{Z}_n} \mathrm{NOR}(x)\chi_k(x) = \chi_k(0) = 1$, we can conclude that $\mathrm{NOR} = \frac{1}{n}\sum_{k=1}^n \chi_k$. It is easy to verify that $\mathrm{Eqv}(x, y) = \mathrm{NOR}(x - y) = \frac{1}{n}\sum_{k=1}^n \chi_k(x)\chi_{n-k}(y)$, thus

$$p_n = \frac{n+1}{2} + \frac{1}{n}\sum_{k=1}^{n-1} \sum_{1 \le i < j \le n+1} \chi_k(x_i)\chi_{n-k}(x_j).$$

By the fact that

$$\left| \sum_{i=1}^{n+1} \chi_k(x_i) \right|^2 = n + 1 + \sum_{1 \le i \ne j \le n+1} \chi_k(x_i)\chi_{n-k}(x_j),$$

we have

$$p_n = \frac{n+1}{2n} + \sum_{k=1}^{n-1} \frac{1}{2n} \left| \sum_{i=1}^{n+1} \chi_k(x_i) \right|^2.$$

This implies $p_n > \frac{1}{2} > 0$ for all $n \in \mathbb{N}$, with an FSOS certificate of sparsity $O(n^2)$. □

6.2. Sum of Hermitian squares (SOHS) of polynomials on $\mathbb{T}^n$

As a counterpart of SOS for non-negative polynomials over $\mathbb{C}$, SOHS has also been extensively studied in polynomial optimization and mathematical physics [34–36]. As an interesting application of sparse FSOS, we show below that one can construct an SOHS of $f \ge 0$ on $\mathbb{T}^n$ from an FSOS of $f \circ \tau$:

$$\underbrace{\Gamma_N \times \cdots \times \Gamma_N}_{n \text{ copies}} \overset{\tau}{\hookrightarrow} \mathbb{T}^n := \underbrace{\mathbb{S}^1 \times \cdots \times \mathbb{S}^1}_{n \text{ copies}} \overset{f}{\longrightarrow} \mathbb{R}_+, \tag{21}$$

Here N is a positive integer, $\Gamma_N = \{\exp(2\pi i k/N)\}_{k=0}^{N-1} \simeq \mathbb{Z}_N$, and τ is the natural inclusion map. It is obvious that an SOHS of f provides an FSOS of $f \circ \tau$. On the other hand, if N is chosen sufficiently large, then one can construct an SOHS of f from an FSOS of $f \circ \tau$ by simply replacing χ_k by z^k and χ_{N-k} by $\overline{z}^k$ for $k < N/2$.

Example 6.3. In order to compute an SOHS of $f(z) = 1 - (z + \overline{z})/2$ on $\mathbb{S}^1$, we choose $N = 6$, and compute an FSOS of $f \circ \tau$ on Γ_6:

$$f \circ \tau = 1 - \frac{1}{2}\chi_1 - \frac{1}{2}\chi_5 = \frac{1}{2}\left|1 - \chi_1\right|^2.$$

Replacing χ_1 by z, we obtain an SOHS of f on $\mathbb{S}^1$:

$$f(z) = \frac{1}{2}(1 - z)(1 - \overline{z}).$$

Let I be the ideal generated by $\overline{x_i}x_i - 1$, $i = 1, 2, \dots, n$. Let ρ be a natural homomorphism.

$$\rho : \mathbb{C}[x_1, \overline{x}_1, \dots, x_n, \overline{x}_n] \mapsto \mathbb{C}[x_1, \overline{x}_1, \dots, x_n, \overline{x}_n]/I.$$

Given a polynomial $f \in \mathbb{C}[x_1, \overline{x}_1, x_2, \overline{x}_2, \ldots, x_n, \overline{x}_n]$, if we have an SOHS representation $\rho(f) = \sum_{i=1}^m h_i \overline{h}_i$ for some $h_i \in \mathbb{C}[x_1, \overline{x}_1, \ldots, x_n, \overline{x}_n]/I, 1 \le i \le m$, then we have $f = \sum_{i=1}^m h_i \overline{h}_i + g$ for some $g \in I$ and $h_i \in \mathbb{C}[x_1, \overline{x}_1, \ldots, x_n, \overline{x}_n]$. Since $g(x) = 0$ for all $x \in \mathbb{T}^n$, we obtain

$$f(x) = \sum_{i=1}^m |h_i(x)|^2, \quad x \in \mathbb{T}^n.$$

Therefore, the nonnegativity of f on $\mathbb{T}^n$ can be certified by an SOHS of $\rho(f)$. For convenience, we denote both f and $\rho(f)$ simply by f in the rest of this section, when there is no risk of confusion.

Given a polynomial

$$f = \sum_{\alpha,\beta \in \mathbb{N}^n} a_{\alpha,\beta} \prod_{i=1}^n x_i^{\alpha_i} \overline{x}_i^{\beta_i} \in \mathbb{C}[x_1, \overline{x}_1, \cdots, x_n, \overline{x}_n]/I,$$

we define $\deg_{x_i}(f) := \max_{i=1}^n \{\alpha_i + \beta_i\}$ for each $i = 1, \ldots, n$. Moreover, the *restriction* $f|_G$ of f to $G = \mathbb{Z}_{k_1} \times \mathbb{Z}_{k_2} \times \cdots \times \mathbb{Z}_{k_n}$ is defined by

$$f|_G(x_1, x_2, \cdots, x_n) = \sum_{\alpha,\beta \in \mathbb{N}^n} a_{\alpha,\beta} \prod_{i=1}^n \chi_{\alpha_i - \beta_i}(x_i) = \sum_{\alpha,\beta \in \mathbb{N}^n} a_{\alpha,\beta} \prod_{i=1}^n \exp\left(\frac{2i\pi(\alpha_i - \beta_i)x_i}{k_i}\right). \tag{22}$$

For each positive integer k, we denote by $[k]$ the set $\{0, 1, 2, \cdots, k-1\}$. We recall that a function g on $G = \mathbb{Z}_{k_1} \times \mathbb{Z}_{k_2} \times \cdots \times \mathbb{Z}_{k_n}$ can be written as

$$g(x_1, x_2, \cdots, x_n) = \sum_{\alpha \in [k_1] \times [k_2] \times \cdots \times [k_n]} a_\alpha \prod_{i=1}^n \chi_{\alpha_i}(x_i) = \sum_{\alpha \in [k_1] \times [k_2] \times \cdots \times [k_n]} a_\alpha \prod_{i=1}^n \exp\left(\frac{2i\pi\alpha_i x_i}{k_i}\right).$$

We define the *lift* of g to $\mathbb{T}^n$ by

$$L(g) = \sum_{\alpha \in [k_1] \times [k_2] \times \cdots \times [k_n]} a_\alpha \prod_{i=1}^n x_i^{\ell_{k_i}(\alpha_i)} \in \mathbb{C}[x_1, \overline{x}_1, \cdots, x_n, \overline{x}_n]/I. \tag{23}$$

Here $x_i^{-j} = \overline{x}_i^j$ in $\mathbb{C}[x_1, \overline{x}_1, \cdots, x_n, \overline{x}_n]/I$ and for each positive integer k, m_k is the map

$$\ell_k(j) = \begin{cases} j, & \text{if } 0 \le j < \frac{k}{2}, \\ j - k, & \text{if } \frac{k}{2} \le j < k. \end{cases} \tag{24}$$

It is clear that the lift of g is a linear map satisfying $L(\overline{g}) = \overline{L(g)}$. The following theorem provides conditions to ensure that the lift of an FSOS is an SOHS.

Theorem 6.4. *Let $f \in \mathbb{C}[x_1, \overline{x}_1, \cdots, x_n, \overline{x}_n]$ be a polynomial defined on $\mathbb{T}^n$ and let S be a subset of $\widehat{G}$. We denote by $f|_G$ the restriction of f on the group $G = \mathbb{Z}_{k_1} \times \mathbb{Z}_{k_2} \times \cdots \times \mathbb{Z}_{k_n}$. Assume the following conditions are satisfied:*

(a) For all $i = 1, 2, \ldots, n$, $k_i > 4\deg_{x_i}(f)$;
(b) $f|_G = \sum_{i=1}^m |h_i|^2$ is an FSOS on the group G, with $\bigcup_{i\in I} \operatorname{supp}(h_i) \subseteq S$;
(c) For all $\chi_\alpha \in S$, and all $i = 1, 2, \ldots, n$, we have $0 \le \alpha_i < k_i/4$ or $3k_i/4 < \alpha_i \le k_i$ holds.

Then we can lift polynomials $h_1, \ldots, h_m$ to give an SOHS of f on $\mathbb{T}^n$.

Proof. According to (22) and (23), $f|_G$ is obtained from f by substituting $x_i^{\alpha_i}$ (resp. $\overline{x}_i^{\beta_i}$) by χ_{α_i} (resp. $\chi_{-\beta_i}$), while $L(g)$ is obtained from g by replacing $\chi_{\alpha_i}(x_i)$ by $x_i^{\ell_{k_i}(\alpha_i)}$. A direct calculation together with (24) implies that $f = L(f|_G)$ if (a) holds.

For functions $h, g : G \mapsto \mathbb{C}$ supported on S, we have

$$L(hg) = L\left(\sum_{\chi_\alpha, \chi_\beta \in S} \widehat{h}(\alpha)\widehat{g}(\beta)\chi_\alpha\chi_\beta\right) = \sum_{\chi_\alpha, \chi_\beta \in S} \widehat{h}(\alpha)\widehat{g}(\beta) \prod_{i=1}^n x_i^{\ell_{k_i}(\alpha_i) + \ell_{k_i}(\beta_i)}, \tag{25}$$

where $\ell_{k_i} : \mathbb{Z}_{k_i} \to \mathbb{Z}$ is the map defined in (24). If (c) holds, then we have

$$\ell_i(\alpha_i) + \ell_i(\beta_i) = \ell_i(\alpha_i + \beta_i), \tag{26}$$

$$L(hg) = L(h)L(g). \tag{27}$$

In fact, since α_i, β_i are elements in $\mathbb{Z}_{k_i}$, (c) leads to three cases below.

- If $0 \le \alpha_i, \beta_i < k_i/4$, then $0 \le \alpha_i + \beta_i < k_i/2$.
- If $0 \le \alpha_i < k_i/4$ and $3k_i/4 < \beta_i \le k_i$, then $3k_i/4 < \alpha_i + \beta_i \le k_i$ or $0 \le \alpha_i + \beta_i \le k_i/4$.
- If $3k_i/4 < \alpha_i, \beta_i \le k_i$, then $k_i/2 < \alpha_i + \beta_i \le k_i$.

Thus, (26) follows immediately from (24). Furthermore, it is clear that (27) is a direct consequence of (25) and (26). Therefore, we obtain

$$f = L(f|_G) = L\left(\sum_{i=1}^m h_i \overline{h}_i\right) = \sum_{i=1}^m L\left(h_i\right) L\left(\overline{h}_i\right) = \sum_{i=1}^m L\left(h_i\right) \overline{L\left(h_i\right)}. \quad \square$$

Remark 6.5. As an important application of SOHS of polynomials on $\mathbb{T}^n$, one can apply it to verify that a polynomial $f(x_1, \ldots, x_n)$ is nonnegative on a hypercube: we can first map the hypercube to $[-2,2]^n$, then we properly choose an abelian group G to construct an SOHS of $f(z_1 + \overline{z}_1, \ldots, z_n + \overline{z}_n)$ on $\mathbb{T}^n$. Comparing with known methods (based on computing SOS over constraints), this new method may provide us a simpler certificate of nonnegativity of polynomials on intervals.

We show below that a sparse FSOS on finite abelian groups can be used to certify the nonnegativity of the Motzkin polynomial in $[-2,2] \times [-2,2]$.

Example 6.6 *(Nonnegativity of the Motzkin polynomial in a square).* Let $M(x, y) = x^4y^2 + x^2y^4 - 3x^2y^2 + 1$ be the Motzkin polynomial. It is well-known that M has no SOS over reals [37,38]. However, we can compute an SOS of $M(x, y)$ subject to constraints $4 - x^2 \ge 0$ and $4 - y^2 \ge 0$. The SOS of M on $[-2,2] \times [-2,2]$ computed by TSSOS [39][1] has sparsity 9:

$$M(x, y) = \sum_{j=1}^3 v_j Q_j v_j^{\mathsf{T}} + (4 - x^2) v_1 Q_4 v_1^{\mathsf{T}} + (4 - y^2) v_1 Q_5 v_1^{\mathsf{T}},$$

where $v_1 = [1, x^2, y^2], v_2 = [x, x^3, xy^2], v_3 = [y, x^2y, y^3]$ and

$$Q_1 = \begin{bmatrix} 0.114 & -0.057 & -0.057 \\ -0.057 & 0.0425 & 0.014 \\ -0.057 & 0.014 & 0.043 \end{bmatrix}, Q_2 = \begin{bmatrix} 1.111 & -0.277 & -0.834 \\ -0.277 & 0.074 & 0.203 \\ -0.834 & 0.203 & 0.630 \end{bmatrix},$$
$$Q_3 = \begin{bmatrix} 1.111 & -0.834 & -0.277 \\ -0.834 & 0.630 & 0.203 \\ -0.277 & 0.203 & 0.074 \end{bmatrix}, Q_4 = \begin{bmatrix} 0.111 & -0.087 & -0.024 \\ -0.087 & 0.074 & 0.013 \\ -0.024 & 0.013 & 0.011 \end{bmatrix},$$
$$Q_5 = \begin{bmatrix} 0.111 & -0.024 & -0.087 \\ -0.024 & 0.011 & 0.013 \\ -0.087 & 0.013 & 0.074 \end{bmatrix}.$$

For a better demonstration, elements in $Q_1, \ldots, Q_5$ are rounded to three decimal places. Our new method works as follows: First, we substitute of x, y by $x = z_1 + \overline{z}_1, y = z_2 + \overline{z}_2$ and consider

$$f(z_1, z_2) := M(z_1 + \overline{z}_1, z_2 + \overline{z}_2), \quad z_1, z_2 \in \mathbb{S}^1.$$

Then we take $N = 8$ and compute a sparse FSOS of the function $f \circ \tau$ on $\Gamma_8 \times \Gamma_8$ by Algorithm 1, which can be lifted further to give an SOHS of f on $\mathbb{T}^2$:

$$f(z_1, z_2) = \left| z_1^2 z_2^2 + \overline{z}_1^2 + \overline{z}_2^2 + z_1^2 \overline{z}_2^2 + 2z_1^2 + z_2^2 + 2 \right|^2.$$

In particular, we obtain a rank one SOHS of $M(z_1 + \overline{z}_1, z_2 + \overline{z}_2)$ for $z_1, z_2 \in \mathbb{S}^1$ of sparsity 7. This provides a proof for the nonnegativity of $M(x, y)$ on $[-2,2] \times [-2,2]$ since $M(x, y)$ is equal to

$$|\exp(2(\theta + \psi)i) + \exp(-2\theta i) + \exp(-2\psi i) + \exp(2(\theta - \psi)i) + 2\exp(2\theta i) + \exp(2\psi i) + 2|^2,$$

where $\theta = \arccos(x/2), \psi = \arccos(y/2)$ for $x, y \in [-2,2]$.

It is proved in [40] that there are nonnegative quadratic functions on $\mathbb{Z}_2^n$ with no FSOS of degree less than $n/2$. We present below a similar example for $\mathbb{Z}_n$.

Example 6.7 *(A function on $\mathbb{Z}_n$ with no low degree FSOS).* For any integer $n \ge 5$, we consider the function

$$g : \mathbb{Z}_n \to \mathbb{C}, \ x \mapsto 2\cos\left(\frac{\pi}{n}\right) - \exp\left(\frac{\pi i}{n}\right) \chi_1(x) - \exp\left(-\frac{\pi i}{n}\right) \chi_{-1}(x).$$

[1] We thank Jie Wang for his help.

It is straightforward to verify that g is nonnegative on $\mathbb{Z}_n$. We claim that g has no FSOS supported on

$$\left\{\chi_{\lfloor -\frac{n}{4}\rfloor}, \chi_{\lfloor -\frac{n}{4}\rfloor+1}, \ldots, \chi_{\lfloor \frac{n}{4}\rfloor-1}, \chi_{\lfloor \frac{n}{4}\rfloor}\right\}, \tag{28}$$

where $\chi_k(x) = \exp(2i\pi kx/n)$ and $\chi_{-k} := \chi_{n-k}$. Indeed, according to Theorem 6.4, an FSOS of g satisfying (28) can be lifted to an SOHS of $L(g) = 2\cos(\pi/n) - \exp(\pi i/n)x - \exp(-\pi i/n)\overline{x}$. However, this contradicts to the fact that

$$L(g)\left(\exp\left(-\frac{\pi i}{n}\right)\right) = 2\cos\left(\frac{\pi}{n}\right) - 2 < 0$$

when $n \geq 5$.

This also implies that for each integer $n \geq 5$, the polynomial optimization problem

$$\begin{array}{ll} \min\limits_{x\in\mathbb{C}} & 2\cos\left(\frac{\pi}{n}\right) - \exp\left(\frac{\pi i}{n}\right)x - \exp\left(-\frac{\pi i}{n}\right)\overline{x} \\ \text{s.t.} & x^n = 1 \end{array}$$

has no SOS certificate of degree less than $n/4$.

Data availability

No data was used for the research described in the article.

Acknowledgments

Ke Ye and Lihong Zhi are supported by the National Key R&D Program of China (2023YFA1009401) and the National Natural Science Foundation of China (12071467). We are very grateful to referees for their careful reading and helpful comments and suggestions.

References

[1] H. Fawzi, J. Saunderson, P.A. Parrilo, Sparse sums of squares on finite abelian groups and improved semidefinite lifts, Math. Program. 160 (1–2) (2016) 149–191.
[2] W. Fulton, J. Harris, Representation Theory: a First Course, vol. 129, Springer Science & Business Media, 2013.
[3] S. Sakaue, A. Takeda, S. Kim, N. Ito, Exact semidefinite programming relaxations with truncated moment matrix for binary polynomial optimization problems, SIAM J. Optim. 27 (1) (2017) 565–582.
[4] M.L. Bonet, J. Levy, F. Manya, Resolution for max-sat, Artif. Intell. 171 (8–9) (2007) 606–618.
[5] W. Rudin, Fourier Analysis on Groups, Interscience Tracts in Pure and Applied Mathematics, vol. 12, Interscience Publishers (a division of John Wiley & Sons, Inc.), New York-London, 1962, ix+285 pp.
[6] R.H. Hammack, W. Imrich, S. Klavžar, W. Imrich, S. Klavžar, Handbook of Product Graphs, vol. 2, CRC Press, Boca Raton, 2011.
[7] N. Kakimura, A direct proof for the matrix decomposition of chordal-structured positive semidefinite matrices, Linear Algebra Appl. 433 (4) (2010) 819–823.
[8] L. Vandenberghe, M.S. Andersen, Chordal graphs and semidefinite optimization, Found. Trends Optim. 1 (4) (2015) 241–433.
[9] D.S. Dummit, R.M. Foote, Abstract Algebra, vol. 3, Wiley, Hoboken, 2004.
[10] E.J. Candes, J.K. Romberg, T. Tao, Stable signal recovery from incomplete and inaccurate measurements, Commun. Pure Appl. Math., J. Iss. Courant Inst. Math. Sci. 59 (8) (2006) 1207–1223.
[11] D. Donoho, M. Elad, V. Temlyakov, Stable recovery of sparse overcomplete representations in the presence of noise, IEEE Trans. Inf. Theory 52 (1) (2006) 6–18, https://doi.org/10.1109/TIT.2005.860430.
[12] C. Eckart, G. Young, The approximation of one matrix by another of lower rank, Psychometrika 1 (3) (1936) 211–218.
[13] D. Elliott, Truncation errors in two Chebyshev series approximations, Math. Comput. 19 (90) (1965) 234–248.
[14] G. Blatman, B. Sudret, Adaptive sparse polynomial chaos expansion based on least angle regression, J. Comput. Phys. 230 (6) (2011) 2345–2367.
[15] K.E. Atkinson, An Introduction to Numerical Analysis, John Wiley & Sons, 2008.
[16] J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, vol. 12, Springer Science & Business Media, 2013.
[17] M. Grant, S. Boyd, CVX: Matlab software for disciplined convex programming, version 2.1, http://cvxr.com/cvx, Mar. 2014.
[18] M. Grant, S. Boyd, Graph implementations for nonsmooth convex programs, in: V. Blondel, S. Boyd, H. Kimura (Eds.), Recent Advances in Learning and Control, Lecture Notes in Control and Information Sciences, Springer-Verlag Limited, 2008, pp. 95–110, http://stanford.edu/~boyd/graph_dcp.html.
[19] K.-C. Toh, M. Todd, SDPT3 – a Matlab software package for semidefinite programming, Optim. Methods Softw. 11 (1999).
[20] R.Y. Zhang, J. Lavaei, Sparse semidefinite programs with guaranteed near-linear time complexity via dualized clique tree conversion, Math. Program. 188 (1) (2021) 351–393.
[21] F. Bach, A. Rudi, Exponential convergence of sum-of-squares hierarchies for trigonometric polynomials, 2022.
[22] L. Slot, M. Laurent, Sum-of-squares hierarchies for binary polynomial optimization, Math. Program. 197 (2) (2023) 621–660.
[23] J.-Y. Cai, X. Chen, Complexity of counting csp with complex weights, in: Proceedings of the Forty-Fourth Annual ACM Symposium on Theory of Computing, 2012, pp. 909–920.
[24] A. Caprara, H. Kellerer, U. Pferschy, The multiple subset sum problem, SIAM J. Optim. 11 (2) (2000) 308–319.
[25] C. Chekuri, S. Khanna, A polynomial time approximation scheme for the multiple knapsack problem, SIAM J. Comput. 35 (3) (2005) 713–728.
[26] C. Lecoutre, Constraint Networks: Targeting Simplicity for Techniques and Algorithms, John Wiley & Sons, 2013.
[27] S. Pettie, V. Ramachandran, An optimal minimum spanning tree algorithm, J. ACM (JACM) 49 (1) (2002) 16–34.
[28] E. Tsang, Foundations of Constraint Satisfaction: the Classic Text, BoD–Books on Demand, 2014.
[29] J. Yang, K. Ye, L. Zhi, Short certificates for max-sat via Fourier sum of squares, arXiv preprint, arXiv:2207.08076, 2022.
[30] H. van Maaren, L. van Norden, M. Heule, Sums of squares based approximation algorithms for max-sat, Discrete Appl. Math. 156 (10) (2008) 1754–1779.
[31] J. Yang, K. Ye, L. Zhi, Fourier sum of squares certificates, arXiv:2207.08076, 2023.
[32] J. Yang, K. Ye, L. Zhi, Lower bounds of functions on finite abelian groups, in: International Computing and Combinatorics Conference, Springer, 2023, pp. 157–170.
[33] M. Davis, R. Sigal, E.J. Weyuker, Computability, Complexity, and Languages: Fundamentals of Theoretical Computer Science, Elsevier, 1994.

[34] C. Josz, D.K. Molzahn, Lasserre hierarchy for large scale polynomial optimization in real and complex variables, SIAM J. Optim. 28 (2) (2018) 1017–1048.
[35] I. Klep, M. Schweighofer, Sums of Hermitian squares and the bmv conjecture, J. Stat. Phys. 133 (4) (2008) 739–760.
[36] J. Wang, V. Magron, Exploiting sparsity in complex polynomial optimization, J. Optim. Theory Appl. 192 (1) (2022) 335–359.
[37] E.L. Kaltofen, B. Li, Z. Yang, L. Zhi, Exact certification in global polynomial optimization via sums-of-squares of rational functions with rational coefficients, J. Symb. Comput. 47 (1) (2012) 1–15.
[38] T.S. Motzkin, The arithmetic-geometric inequality, in: Inequalities, Proc. Sympos., Wright-Patterson Air Force Base, Ohio, 1965, 1965, pp. 205–224.
[39] J. Wang, V. Magron, J.-B. Lasserre, Tssos: a moment-sos hierarchy that exploits term sparsity, SIAM J. Optim. 31 (1) (2021) 30–58.
[40] G. Blekherman, J. Gouveia, J. Pfeiffer, Sums of squares on the hypercube, Math. Z. 284 (1) (2016) 41–54.