Deterministic Sparse Fourier Approximation via Fooling Arithmetic Progressions

Adi Akavia[*]
The Weizmann Institute of Science, Rehovot, Israel.
adi.akavia@weizmann.ac.il

Abstract

A significant Fourier transform (SFT) algorithm, given a threshold τ and oracle access to a function f, outputs (the frequencies and approximate values of) all the τ-significant Fourier coefficients of f, i.e., the Fourier coefficients whose magnitude exceeds $\tau\|f\|_2^2$. In this paper we present the first *deterministic* SFT algorithm for functions f over $\mathbb{Z}_N$ which is: (1) Local, i.e., its running time is polynomial in $\log N$, $1/\tau$ and $L_1(\widehat{f})$ (the L_1 norm of f's Fourier transform). (2) Robust to random noise. This strictly extends the class of compressible/Fourier sparse functions over $\mathbb{Z}_N$ efficiently handled by prior deterministic algorithms. As a corollary we obtain deterministic and robust algorithms for sparse Fourier approximation, compressed sensing and sketching.

As a central tool, we prove that there are:

1. Explicit sets A of size $poly((\ln N)^d, 1/\varepsilon)$ with ε-discrepancy in all rank d Bohr sets in $\mathbb{Z}_N$. This extends the Razborov-Szemeredi-Wigderson result on ε-discrepancy in arithmetic progressions to Bohr sets, which are their higher rank analogue.
2. Explicit sets A_P of size $poly(\ln N, 1/\varepsilon)$ that ε-approximate the uniform distribution over a given arithmetic progression P in $\mathbb{Z}_N$, in the sense that $|\mathbb{E}_{x\in A}\,\chi(x) - \mathbb{E}_{x\in P}\,\chi(x)| < \varepsilon$ for all linear tests χ in $\mathbb{Z}_N$. This extends results on small biased sets, which are sets approximating the uniform distribution over the entire domain, to sets approximating uniform distributions over (arbitrary size) arithmetic progressions.

These results may be of independent interest.

1 Introduction

Computing the Fourier transform is a basic building block used in numerous applications. Its complexity is well understood: Quasi-linear running time $O(N \log N)$ for N the input size is achieved by the Fast Fourier Transform (FFT) algorithm [CT65], and believed to be optimal. For data intensive applications, however, achieving a *sub-linear* running time is desired. In general, this is infeasible, because the input and output are already of size N. Nevertheless, in settings where the input is given via *oracle access*, and it suffices to output only the few *"significant"* Fourier coefficients, sub-linear algorithms do exist.

We say that a Fourier coefficient is τ-significant if its magnitude is at least a τ-fraction (say, 1%) of the sum of squared Fourier coefficients. A significant Fourier transform (SFT) algorithm is an algorithm that, given a significance threshold τ and oracle access to a function f, outputs all the τ-significant Fourier coefficients of f (i.e., their frequencies and approximate values). The complexity of such algorithms is measured primarily in terms of $1/\tau$ and the size N of (the truth table of) f.

Randomized SFT algorithms achieving complexity polynomial in $\log N$ and $1/\tau$ for functions over any finite abelian group were developed in a sequence of works [GL89, KM93, Man95, GGI$^+$02, AGS03, GMS05, Aka09].

Deterministic SFT algorithms were given for restricted functions:

- Functions over the *boolean hypercube* $\{0,1\}^n$ in Kushilevitz-Mansour's (KM) algorithm [KM93].

[*] This research was supported in part by NSF grant CCF-0514167, by NSF grant CCF-0832797, and by Israel Science Foundation 700/08.

- *Compressible* or *Fourier sparse* functions over $\mathbb{Z}_N$ in Iwen's algorithms [Iwe07, Iwe08, IS08]. (A function is compressible if its Fourier coefficients decay as fast as the series $c(1/i)^p$ for absolute constants $c > 0$ and $p > 1$. A function is Fourier sparse if it has at most $poly(\log N)$ non-zero Fourier coefficients.)

The KM algorithm [KM93] is given as an extra input an upper bound t on the sum of (absolute values of) Fourier coefficients of the input function f, $L_1(\widehat{f}) \overset{def}{=} \sum_\alpha \left|\widehat{f}(\alpha)\right|$. Its running time is polynomial in $\log N$, $1/\tau$ and t. We say that a deterministic SFT algorithm *achieves the KM benchmark* if its complexity is polynomial in $\log N$, $1/\tau$ and t.

1.1 Main Result: Deterministic & Robust SFT Algorithm

In this paper we present a deterministic SFT algorithm achieving the KM benchmark for all functions f over $\mathbb{Z}_N$. Furthermore, our SFT algorithm is *robust* to random noise. That is, the algorithm succeeds even if the oracle to f is noisy in the sense that on queries x the oracle returns the value $f'(x) = f(x) + \eta(x)$ for $\eta \colon \mathbb{Z}_N \to \mathbb{C}$ an ε-random noise, i.e., values $\eta(x)$ are drawn independently at random from distributions D_x of expected absolute value at most $\mathbb{E}_{\eta(x)\sim D_x}[|\eta(x)|] \leq \varepsilon$.

Theorem 1 *There is a deterministic algorithm such that:*

- *Given N, τ, t, and oracle access to a function $f \colon \mathbb{Z}_N \to \mathbb{C}$ s.t. $L_1(\widehat{f}) \leq t$, the algorithm outputs all the τ-significant Fourier coefficients of f.*
- *Given N, τ, t, and oracle access is to a function $f' \colon \mathbb{Z}_N \to \mathbb{C}$, where $f' = f + \eta$ s.t. $L_1(\widehat{f}) \leq t$ and η is a $\tau/3$-random noise, the algorithm outputs all the τ-significant Fourier coefficients of f (with probability at least $1 - 1/N^{\Theta(1)}$ over the random noise η).*

The running time and query complexity are polynomial in $\log N$*,* $1/\tau$ *and* t*.*

Remarks. (i) The KM benchmark is matched by taking $t = L_1(\widehat{f})$. (ii) We stress that the complexity of our algorithm depends on the bound t on $L_1(\widehat{f})$, and not on a bound on $L_1(\widehat{f'})$. This is crucial, because even if $L_1(\widehat{f}) \leq t$ is small, typically $L_1(\widehat{f'}) \approx \sqrt{N}$ is very large.

Our algorithm is better than the prior deterministic SFT algorithms for functions over $\mathbb{Z}_N$ in:

1. Achieving the KM benchmark. In particular, our algorithm efficiently (i.e., in time polynomial in $\log N$) handles a much wider class of functions than handled by prior works: all functions f s.t. $L_1(\widehat{f}) \leq poly(\log N)$, instead of only the compressible/Fourier sparse functions.[1]

   Handling this wider class of functions is motivated by functions arising in applications, e.g., threshold functions $f_\theta(x) = 1$ iff $x \leq \theta$ and 0 otherwise.

2. Achieving robustness to random noise. In contrast, in other deterministic algorithms, noisy functions $f' = f + \eta$ are out of the scope of functions handled efficiently, because typically f' is not compressible/Fourier sparse (even if f were).[2]

   Robustness to noise is motivated for example by measurement noise in signal processing applications.

1.2 New Tools: Fooling Bohr Sets and Arithmetic Progressions

As a central ingredient for our deterministic SFT algorithm, we prove that there exists explicit constructions of: (1) Sets with small discrepancy on all rank d Bohr sets; and (2) Sets that ε-approximate the uniform distribution on a given arithmetic progression (definitions follow). These results may be of independent interest.

[1] In the context of SFT algorithms, compressible functions f are a strict subclass of the functions with poly-logarithmic $L_1(\widehat{f})$. This is because without loss of generality we may assume that f is normalized to have (approximately) unit energy (as significance is determined by *ratios* of Fourier coefficient magnitude to total energy), and for normalized f, compressibility implies that $L_1(\widehat{f}) = O(1)$.

[2] A few remarks. (i) Having some restriction on the class of functions efficiently handled by deterministic algorithms is unavoidable (because two input functions f, g may be identical on the small set of entries read by the deterministic algorithm, while differing widely on their Fourier transform, implying the algorithm fails on at least one out of f, g). The class of functions handled by our algorithm is wide enough to include typical noise. (ii) Our analysis extends to show that KM's deterministic algorithm for functions over $\{0, 1\}^n$ is also robust to random noise.

1.2.1 Definitions

For sets A, S in a group G, we denote by U_S the uniform distribution over S, and say that:

- A has ε-discrepancy on S in G if the intersection $|A \cap S|$ is roughly as expected if A were random:
$$D_{A,G}(S) \overset{def}{=} \left| \frac{|A \cap S|}{|A|} - \frac{|S|}{|G|} \right| < \varepsilon.$$
For a family $\mathcal{S}$ of sets, A has ε-discrepancy on $\mathcal{S}$, if $D_{A,G}(\mathcal{S}) \overset{def}{=} \max_{S \in \mathcal{S}} D_{A,G}(S) < \varepsilon$.

- A ε-approximates U_S in G if for all linear tests $\chi \colon G \to \mathbb{C}$ in G, the expected outcome $\chi(x)$ over uniform x in A is ε-close to its expected outcome over uniform x in S:
$$\left| \mathop{\mathbb{E}}_{x \in A} \chi(x) - \mathop{\mathbb{E}}_{x \in S} \chi(x) \right| < \varepsilon.$$
We focus on $G = \mathbb{Z}_N$, where linear tests are the functions $\chi_\alpha(x) \overset{def}{=} e^{2\pi i \alpha x / N}$ indexed by $\alpha \in \mathbb{Z}_N$.

- A is explicit if there is a deterministic algorithm that, given G (by its generators and their orders) and ε, outputs A in time polynomial in $|A|$. We usually focus on explicit sets A of size polynomial in $\log |G|$ and $1/\varepsilon$.

We are particularly interested in sets S that are either arithmetic progressions or Bohr sets (which are the higher rank analogue of arithmetic progressions used in many additive combinatorics works [TV06]):

- Arithmetic progressions in $\mathbb{Z}_N$ are sets $P_{\alpha,I} \overset{def}{=} \{x \cdot \alpha \bmod N \mid x \in I\}$ for $\alpha \in \mathbb{Z}_N$ a multiplier and $I = [a..b]$ an interval (i.e., the set of integers in $[a, b]$) with endpoints $0 \le a \le b < N$.

- Rank d Bohr sets in $\mathbb{Z}_N$ are sets $B_{\{\alpha_i, I_i\}_{i=1}^d} \overset{def}{=} \{x \in \mathbb{Z}_N \mid \alpha_i \cdot x \bmod N \in I_i \; \forall i = 1, \ldots, d\}$ for $\alpha_i \in \mathbb{Z}_N$ multipliers and $I_i = [a_i..b_i]$ intervals with endpoints $0 \le a_i \le b_i < N$. Denote by $\mathcal{B}_{N,d}$ the set of all rank d Bohr sets in $\mathbb{Z}_N$.

1.2.2 Our Results

We show that there exists (1) explicit sets with small discrepancy on all rank d Bohr sets, and (2) explicit sets approximating the uniform distribution on a given arithmetic progression:

Theorem 2 *1. For any N, ε, d, there is an explicit set $A \subseteq \mathbb{Z}_N$ of size polynomial in $1/\varepsilon$ and $(\ln N)^d$ with* ε-discrepancy *on $\mathcal{B}_{N,d}$.*

2. For any N, ε, and an arithmetic progression P in $\mathbb{Z}_N$, there is an explicit set $A \subseteq \mathbb{Z}_N$ of size polynomial in $1/\varepsilon$ and $\ln |P|$ that ε-approximates *the distribution U_P.*

Remarks and comparison to prior works.

1. Our proof is by reduction to explicit constructions of small biased sets. The exact size of our sets A depends on the small biased set we use. For example, using the ε'-biased set of size $O((\log N)^2/\varepsilon')$ of [Kat89] results in sets A of sizes $O((\log N)^{2+d}/\varepsilon)$ and $O((\log N)^4/\varepsilon^3)$ in Theorem 2 Parts 1 and 2 respectively.[3]

2. For $d = 1$, our proof of Theorem 2 Part 1 gives a new –and much simpler– proof for the Razborov-Szemeredi-Wigderson [RSW93] result on ε-discrepancy on arithmetic progressions.

   Our result (even when restricted to $d = 1$) is better than the latter in: (i) Achieving ε-discrepancy on *all* arithmetic progressions $P_{\alpha,I}$, in contrast to only $P_{\alpha,I}$ s.t. α is co-prime to N. (ii) Achieving a better set size whenever $\varepsilon < 1/(\log N)^{1/8}$ (as for $d = 1$ our set size is $\Theta((\log N)^3/\varepsilon)$ compared with $\Theta((\log N)^2/\varepsilon^9)$ in [RSW93]).

3. Theorem 2 Part 2 can be viewed as a generalization of small biased sets in $\mathbb{Z}_N$: Small biased sets are sets approximating the uniform distribution over the entire domain, that is, the arithmetic progression $P_{\alpha,I}$ for $\alpha = 1$ and $I = [0..N-1]$, whereas our result addresses arbitrary arithmetic progressions P.

[3] The size quoted here for the ε'-biased set of [Kat89] is taken from the accounts of [AIK$^+$90] on [Kat89]. We remark that using the (much simpler) ε'-biased set of size $O((\log N)^2/\varepsilon'^3)$ of [AIK$^+$90] results in sets A of sizes $O((\log N)^{2+d}/\varepsilon^3)$ and $O((\log N)^4/\varepsilon^9)$ in Theorem 2 Parts 1 and 2 respectively.

1.3 Paper Organization

In the rest of this paper we present: An overview of our proof (Sect. 2); Preliminaries (Sect. 3); Our explicit constructions (Sect. 4); Our deterministic SFT algorithm (Sect. 5); Concluding remarks (Sect. 6).

2 Proof Overview

Our starting point is the randomized SFT algorithm of [Aka09]. Randomness there is employed solely for constructing a set $S = S_{N,\tau,t} \subseteq \mathbb{Z}_N$ of queries to the oracle to the input function f, such that (with high probability) S is good according to Definition 3 below. Their analysis shows that: (i) If S is good, then for *all* functions f over $\mathbb{Z}_N$ s.t. $L_1(\widehat{f}) \leq t$, their algorithm finds the τ-significant Fourier coefficients of f (even in the presence of noise) while querying f only on the entries in S. (ii) With high probability, their S is good.

Definition 3 (Good queries [Aka09]) *A set $S = S_{N,\tau,t} \subseteq \mathbb{Z}_N$ is (N,τ,t)-*good *(*good*, in short) if $S = \bigcup_{\ell=0}^{\lfloor(\log N)\rfloor}(A - B_\ell)$ s.t. for $\varepsilon = \Theta(\tau/(t^2 \log N))$ sufficiently small:*

- *A is an ε-biased set in $\mathbb{Z}_N$*
- *For each ℓ, B_ℓ ε-approximates the distribution $U_{[0..2^\ell-1]}$ in $\mathbb{Z}_N$*
- *The sizes $|A|$ and $|B_1|, \ldots, \left|B_{\lfloor(\log N)\rfloor}\right|$ are polynomial in $\log N$, $1/\tau$ and t*

Remark. *Exact setting of ε may vary depending on the desired tradeoff: We take $\varepsilon = \tau/(3t^2 \ln N)$ when there is no noise, and $\varepsilon = \tau/(49t^2 \ln N)$ to tolerate up to $\tau/3$-random noise.*

In this work we obtain a deterministic (and robust) SFT algorithm by replacing the randomized construction of sets S in [Aka09] with an explicit (i.e., efficient and deterministic) construction.

The heart of our construction is a novel analysis, where, for any arithmetic progression P (say, $[0..2^\ell]$), we reduce the problem of finding explicit sets approximating the distribution U_P in $\mathbb{Z}_N$ to the problem of finding explicit sets with small bias in $\mathbb{Z}_M$ for $M = |P|$.[4] We then obtain a good set S by utilizing known constructions of explicit small-biased sets in $\mathbb{Z}_M$ [Kat89, AIK+90, RSW93].

Our reduction is composed of the three parts detailed in the theorem below.

Theorem 4 (Our reduction) *For any positive integers $d, M < N$, a positive real ε, and $A \subseteq I = [0..M-1]$,*

1. *If A is $\varepsilon/(4\ln M)^d$-biased in $\mathbb{Z}_M$, then A has ε-discrepancy on all rank d Bohr sets in $\mathbb{Z}_M$.*
2. *If A has $\varepsilon^3/(128\pi^2)$-discrepancy on all rank 2 Bohr sets in $\mathbb{Z}_M$, then A ε-approximates U_I in $\mathbb{Z}_N$.*
3. *For every $\alpha, s \in \mathbb{Z}_N$, if A ε-approximates U_I in $\mathbb{Z}_N$, then $\alpha(A+s)$ ε-approximates $U_{P_{\alpha,I+s}}$ in $\mathbb{Z}_N$.*

Remark. The converse of Theorem 4 Part 1 is known (and simple to prove): If A has ε-discrepancy on all arithmetic progressions in $\mathbb{Z}_N$, then A is $2\pi\varepsilon$-biased in $\mathbb{Z}_N$ (see [RSW93], Proposition 4.1).

Proving part 3 of our reduction is straightforward, whereas Parts 1-2 require more insight (details follow).

Theorem 4, Part 1. We relate having small bias in $\mathbb{Z}_M$ to having small discrepancy on all rank d Bohr sets in $\mathbb{Z}_M$ as follows. First we upper bound the discrepancy of A on any set R (say, a rank d Bohr set) in $\mathbb{Z}_M$ by $D_{A,\mathbb{Z}_N}(R) \leq \varepsilon' \cdot L_1(\widehat{R})$ for ε' the bias of A in $\mathbb{Z}_M$ and for $L_1(\widehat{R}) \overset{def}{=} \sum_{\alpha\in\mathbb{Z}_M}\left|\widehat{R}(\alpha)\right|$ the L_1-norm of the Fourier transform of (the characteristic function of) R. Next we apply Fourier analysis and some elementary number theory to show that for any rank d Bohr set R in $\mathbb{Z}_M$, $L_1(\widehat{R}) \leq (4\ln M)^d$. We conclude that if A is $\varepsilon' = \varepsilon/(4\ln M)^d$-biased in $\mathbb{Z}_M$, then A has ε-discrepancy on all rank d Bohr sets in $\mathbb{Z}_M$.

Theorem 4, Part 2. We relate approximating U_I in $\mathbb{Z}_N$ to having small discrepancy on rank 2 Bohr sets in $\mathbb{Z}_M$ as follows.

First, we identify each element $\alpha \in \mathbb{Z}_N$ with the pair (q_α, r_α) of its quotient and remainder in the division-with-remainder by (the typically, non-integer value) N/M. We then rewrite each linear test $\chi_\alpha(x) = e(\alpha x/N)$ in $\mathbb{Z}_N$ as: $\chi_\alpha(x) = e\left(\left(\left(\frac{q_\alpha x}{M}\right)_1 + \left(\frac{r_\alpha x}{N}\right)_1\right)_1\right)$ (where for any real number r, $(r)_1$ denotes its non-integer part, i.e., its remainder modulo 1; and $e(r) = e^{2\pi i r}$).

Second, we embed the sets $\left\{\left(\frac{q_\alpha x}{M}\right)_1\right\}_{x\in I}$ and $\left\{\left(\frac{r_\alpha x}{N}\right)_1\right\}_{x\in I}$ into $\mathbb{Z}_M$, and use this embedding to show that if A has ε'-discrepancy on all rank 2 Bohr sets in $\mathbb{Z}_M$, then the joint distribution of pairs $\left(\left(\frac{q_\alpha x}{M}\right)_1, \left(\frac{r_\alpha x}{N}\right)_1\right)$

[4] We remark that the connection between approximating U_P in $\mathbb{Z}_N$ and having small bias in $\mathbb{Z}_{|P|}$ may seem surprising. For example, achieving the former requires satisfying N linear tests modulo N, whereas achieving the latter requires satisfying only $|P|$ linear tests and these tests are modulo $|P|$ (where $|P| < N$ is arbitrary).

over uniform x in A is "close" to their distribution over uniform x in I; where closeness is in the sense that for every two length $\rho = \varepsilon/8\pi$ intervals $J_1, J_2 \subseteq [0,1]$,

$$\left|\Pr_{x\in A}\left[\left(\frac{q_\alpha x}{M}\right)_1 \in J_1 \;\&\; \left(\frac{r_\alpha x}{N}\right)_1 \in J_2\right] - \Pr_{x\in I}\left[\left(\frac{q_\alpha x}{M}\right)_1 \in J_1 \;\&\; \left(\frac{r_\alpha x}{N}\right)_1 \in J_2\right]\right| < \varepsilon'$$

Finally, we prove that ε'-closeness of these two joint distributions implies $((\varepsilon'/\rho^2) + 4\pi\rho)$-closeness of the expected value of $\chi_\alpha(x) = e\left(\left(\left(\frac{q_\alpha x}{M}\right)_1 + \left(\frac{r_\alpha x}{N}\right)_1\right)_1\right)$ over uniform x in A and over uniform x in I. Assigning $\varepsilon' = \varepsilon^3/(128\pi^2)$, we conclude that A ε-approximates U_I in $\mathbb{Z}_N$.

3 Preliminaries

In this section we summarize some preliminary terminology, notations and facts.

Notations. Let $\mathbb{Z}$ and $\mathbb{C}$ denote the integer and complex numbers respectively. Let $\mathbb{Z}_N$ and $\mathbb{Z}_N^*$ denote the additive and the multiplicative groups of integers modulo N. We identify the elements of $\mathbb{Z}_N$ with integers in $0,\dots,N-1$, and denote $\mathsf{abs}(\alpha) = \min\{\alpha, N-\alpha\}$ for all $\alpha \in \mathbb{Z}_N$. We denote by $[a..b]$ the set of integers in the closed interval $[a,b]$. For any element $a \in \mathbb{Z}_N$ and sets $S, S' \subseteq \mathbb{Z}_N$, denote $aS = \{as\}_{s\in S}$ and $S - S' = \{s - s'\}_{s\in S, s'\in S'}$. For any real number r, denote $e(r) = e^{2\pi i r}$, and denote by $(r)_1$ the remainder of r in division by 1.

3.1 Significant Fourier Transform Coefficients

We give definitions and properties for normed spaces and Fourier transform.

Inner product, norms, convolution. The *inner product* of complex valued functions f, g over a domain G is $\langle f, g\rangle \stackrel{def}{=} \frac{1}{|G|}\sum_{x\in G} f(x)\overline{g(x)}$. Denote the ℓ_2*-norm* of f by $\|f\|_2 \stackrel{def}{=} \sqrt{\langle f, f\rangle}$, and its L_1*-norm* by $L_1(f) \stackrel{def}{=} \sum_{x\in G}|f(x)|$. The *convolution* of f and g is the function $f * g\colon G \to \mathbb{C}$ defined by $f * g(x) \stackrel{def}{=} \frac{1}{|G|}\sum_{y\in G} f(y)\overline{g(x-y)}$.

Characters and Fourier transform. The *characters* of a finite abelian groups G are all the homomorphisms $\chi\colon G \to \mathbb{C}$ from G into the complex unit sphere. Denote by $\widehat{G}$ the set of characters of G. The *Fourier transform* of a complex valued function f over G is the function $\widehat{f}\colon \widehat{G} \to \mathbb{C}$ defined by $\widehat{f}(\chi) \stackrel{def}{=} \langle f, \chi\rangle$. A character χ is *trivial* if $\chi(x) = 1$ for all x.

In particular, the characters of $\mathbb{Z}_N$ are the functions $\chi_\alpha\colon \mathbb{Z}_N \to \mathbb{C}$, $\alpha \in \mathbb{Z}_N$, defined by $\chi_\alpha(x) \stackrel{def}{=} e^{2\pi i \alpha x/N}$. Abusing notation, we view $\widehat{f}$ as a function over G, defined by $\widehat{f}(\alpha) \stackrel{def}{=} \langle f, \chi_\alpha\rangle$.

Significant Fourier coefficients. For any $\alpha \in \mathbb{Z}_N$ and $\tau \in [0,1]$, we say that α is a τ*-significant* Fourier coefficient iff $\left|\widehat{f}(\alpha)\right|^2 \geq \tau\|f\|_2^2$. Denote by $\mathsf{Heavy}_\tau(f)$ the set of all τ-significant Fourier coefficients of f.

Useful Fourier transform properties. For every positive integer N, functions $f, g\colon \mathbb{Z}_N \to \mathbb{C}$, and elements $s \in \mathbb{Z}_N$ and $t \in \mathbb{Z}_N^*$, the following holds (where subtraction, multiplication and inverse operations are modulo N): *Parseval Identity*: $\frac{1}{N}\sum_{x\in\mathbb{Z}_N}|f(x)|^2 = \sum_{\alpha\in\mathbb{Z}_N}\left|\widehat{f}(\alpha)\right|^2$. *Convolution Theorem*: $\widehat{(f * g)}(\alpha) = \widehat{f}(\alpha)\cdot\widehat{g}(\alpha)$ and similarly $\frac{1}{N}\widehat{f\cdot g}(\alpha) = (\widehat{f} * \widehat{g})(\alpha)$. *Phase Shift*: If $g = f\cdot\chi_{-s}$, then $\widehat{g}(\alpha) = \widehat{f}(\alpha - s)$ for all α. *Scaling*: If $g(x) = f(tx)\ \forall x$, then $\widehat{g}(\alpha) = \widehat{f}(\alpha\cdot t^{-1})$ for all α. Finally, for all integers α, $\frac{1}{N}\sum_{x\in\mathbb{Z}_N} e(\alpha x/N) = 1$ iff $\alpha = 0 \bmod N$ and it is 0 otherwise.

3.2 Small Biased Sets

Let G be a finite abelian group and $A \subseteq G$. Denote $bias_A(\chi) \stackrel{def}{=} \frac{1}{|A|}\sum_{a\in A}\chi(a)$. A is ε*-biased* in G if for all non-trivial characters χ of G, $|bias_A(\chi)| < \varepsilon$.

Fact 1 (ε-biased sets in $\mathbb{Z}_N$ [AIK$^+$90, RSW93, Kat89]) *For any integer $N > 0$ and real $\varepsilon > 0$, there exists an explicit set A of size polynomial in $\log N$ and $1/\varepsilon$ which is ε-biased in $\mathbb{Z}_N$.*

For the sake of completeness we specify the details of one of the small biased sets constructed in [AIK$^+$90] (chosen due to its simplicity): For each N, ε, let $AIKPS(N,\varepsilon)$ denote the ε-biased set in $\mathbb{Z}_N$ of [AIK$^+$90] defined by:

$$AIKPS(N,\varepsilon) \stackrel{def}{=} \left\{ sp^{-1} \bmod N \mid s \in S, p \in P, p \not| s\right\} \tag{1}$$

where, for $\varepsilon' = -(\log\varepsilon)/(\log\log N)$, $P = \left\{p \mid p \text{ is prime}, (\log N)^{1+\varepsilon'}/2 < p \leq (\log N)^{1+\varepsilon'}\right\}$ and $S = [1..(\log N)^{1+2\varepsilon'}]$.

4 Our Results on Explicit Constructions

We present our results on explicit sets with small discrepancy on Bohr sets (Sect. 4.1), and explicit sets approximating distributions uniform over arithmetic progressions (Sect. 4.2). See definitions in Sect. 1.2.1.

4.1 Small Discrepancy in Bohr Set

We show that there are explicit small sets A with small discrepancy on all rank d Bohr sets in $\mathbb{Z}_N$.

Theorem 2 Part 1 (Small discrepancy on Bohr sets). For any integers $N, d > 0$ and real $\varepsilon > 0$, there is an explicit set $A \subseteq \mathbb{Z}_N$ of size polynomial in $1/\varepsilon$ and $(\ln N)^d$ s.t. $D_{A,\mathbb{Z}_N}(\mathcal{B}_{N,d}) < \varepsilon$.

Proof: Fix N, d, ε. Denote $\varepsilon' = \varepsilon/(4\ln N)^d$. Let A be any explicit ε'-biased set A in $\mathbb{Z}_N$ of size polynomial in $\log N$ and $1/\varepsilon'$ (such sets exists by Fact 1, Sect. 3.2). By Lemma 5, for any set $B \subseteq \mathbb{Z}_N$,

$$D_{A,\mathbb{Z}_N}(B) < \varepsilon' L_1(\widehat{B})$$

for $L_1(\widehat{B}) \overset{def}{=} \sum_{\alpha\in\mathbb{Z}_N} \left|\widehat{B}(\alpha)\right|$ the L_1-norm of the Fourier transform of the characteristic function $B(x) = 1$ iff $x \in B$ and 0 otherwise. By Lemma 6, for any rank d Bohr set B,

$$L_1(\widehat{B}) \le (4\ln N)^d.$$

We conclude that $D_{A,\mathbb{Z}_N}(B) < \varepsilon' \cdot (4\ln N)^d = \varepsilon$. ■

For any subsets A, B of a finite abelian group G, we bound the discrepancy of A on B in G by the product of the bias of A in G and the L_1-norm of the Fourier transform of B.

Lemma 5 *For any finite abelian group G, sets $A, B \subseteq G$ and a real number $\varepsilon > 0$, if A is ε-biased in G, then $D_{A,G}(B) < \varepsilon L_1(\widehat{B})$. Furthermore, this bound is tight.*

Proof: We show that $D_{A,G}(B) < \varepsilon L_1(\widehat{B})$. Recall that $D_{A,G}(B) = \left|\frac{|A\cap B|}{|A|} - \frac{|B|}{|G|}\right|$. Replacing B with its characteristic function, we get that $D_{A,G}(B) = \left|\frac{1}{|A|}\sum_{a\in A} B(a) - \frac{1}{|G|}\sum_{a\in G} B(a)\right|$. Replacing the characteristic function of B with its Fourier representation $B(a) = \sum_\chi \widehat{B}(\chi)\chi(a)$ and rearranging the summation we get that $D_{A,G}(B) = \left|\sum_\chi \widehat{B}(\chi)\left(\frac{1}{|A|}\sum_{a\in A}\chi(a) - \frac{1}{|G|}\sum_{a\in G}\chi(a)\right)\right|$ which is in turn equal to

$$D_{A,G}(B) = \left|\sum_{\text{non-trivial }\chi} \widehat{B}(\chi) \cdot bias_A(\chi)\right| \tag{2}$$

(since for the trivial χ, the difference in the parenthesis is zero, and for all non-trivial χ, $\frac{1}{|G|}\sum_{a\in G}\chi(a) = 0$). Using the triangle inequality and bounding the bias of A by its upper bound ε, we conclude that

$$D_{A,G}(B) \le \max_{\text{non trivial }\chi} |bias_A(\chi)| \sum_\chi \left|\widehat{B}(\chi)\right| < \varepsilon L_1(\widehat{B}).$$

We prove the bound is tight. Let $G = \{0,1\}^n$. Given any ε-biased set $A \subseteq \{0,1\}^n$ and $\alpha \in \{0,1\}^n$ s.t. $bias_A(\chi_\alpha) = \varepsilon$, define $B = \{\, x \mid x \cdot \alpha = 0\}$ ($x \cdot \alpha$ is the dot product). Then $D_{A,G}(B) = \varepsilon L_1(\widehat{B})$, because $D_{A,G}(B) = \left|\sum_{\alpha\neq 0} \widehat{B}(\chi_\alpha) bias_A(\chi_\alpha)\right|$ (by Eq. 2) and $\widehat{B}(\chi_\beta) \neq 0$ iff $\beta = \alpha$ (by Fourier analysis of B). ■

We bound the L_1-norm of the Fourier transform of rank d Bohr sets B.

Lemma 6 *For any positive integers N, d, and B the characteristic function of a rank d Bohr set in $\mathbb{Z}_N$, $L_1(\widehat{B}) \le (4\ln N)^d$.*

Proof: Fix $B_{\{\alpha_i,I_i\}_{i=1}^d}$ a rank d Bohr set in $\mathbb{Z}_N$. Observe that $B_{\{\alpha_i,I_i\}_{i=1}^d} = \bigcap_{i=1}^d B_{\{\alpha_i,I_i\}}$ is the intersection of the d rank 1 Bohr sets $B_{\{\alpha_i,I_i\}}$. Denote by B and $B_1,\dots,B_d$ the characteristic functions of $B_{\{\alpha_i,I_i\}_{i=1}^d}$ and $B_{\{\alpha_1,I_1\}},\dots,B_{\{\alpha_d,I_d\}}$ respectively. Then $B = \prod_{i=1}^d B_i$. By Proposition 2, the above implies that $L_1(\widehat{B}) \le \prod_{i=1}^d L_1(\widehat{B_i})$. By Lemma 7, $L_1(\widehat{B_i}) \le 4\ln N$ for any rank 1 Bohr set B_i. We conclude that $L_1(\widehat{B}) \le (4\ln N)^d$. ■

We bound the L_1-norm of the Fourier transform of a product of functions.

Proposition 2 *Let $f, g\colon G \to \mathbb{C}$, then $L_1(\widehat{f \cdot g}) \leq L_1(\widehat{f}) \cdot L_1(\widehat{g})$.*

Proof: By the convolution theorem, $\widehat{f \cdot g}(\chi) = N(\widehat{f} * \widehat{g})(\chi)$. Thus, $L_1(\widehat{f \cdot g}) = \sum_\chi \left|N(\widehat{f} * \widehat{g})(\chi)\right|$. By definition of the convolution operator, the latter is equal to the sum $\sum_\chi \left|\sum_\psi \widehat{f}(\psi) \cdot \widehat{g}(\chi \cdot \psi^{-1})\right|$ over χ, ψ characters of the group G, which is upper bounded by $\sum_\psi \left|\widehat{f}(\psi)\right| \sum_\chi \left|\widehat{g}(\chi \cdot \psi^{-1})\right| = L_1(\widehat{f}) \cdot L_1(\widehat{g})$. ■

We bound the L_1-norm of the Fourier transform of rank 1 Bohr sets.

Lemma 7 *For any positive integer N, and B the characteristic function of any rank 1 Bohr set in $\mathbb{Z}_N$, $L_1(\widehat{B}) \leq 4 \ln N$.*

Remark. The bound is tight up to constants, that is, there are Bohr sets B s.t. $L_1(\widehat{B}) = \Omega(\ln N)$.

Proof: Fix a rank 1 Bohr set in $\mathbb{Z}_N$, $B = B_{\alpha,I}$, for $\alpha \in \mathbb{Z}_N$ and $I = [s..t] \subseteq \mathbb{Z}_N$. Denote by $g \overset{def}{=} gcd(N, \alpha)$ the greatest common divisor of N and α, and let $\beta \overset{def}{=} \alpha/g$. We prove that $L_1(\widehat{B}) \leq 4 \ln N$.

For $g = 1$ the proof is simple: Since α is co-prime to N, then $\mathbb{Z}_N = \left\{\alpha^{-1}x \bmod N\right\}_{x \in \mathbb{Z}_N}$, implying that B is equal to the arithmetic progression $\left\{(\alpha^{-1}x \bmod N) \in \mathbb{Z}_N \,\middle|\, \alpha(\alpha^{-1}x) \in I\right\} = \alpha^{-1}I$. By the scaling property of the Fourier transform, for α^{-1} co-prime to N, $L_1(\widehat{\alpha^{-1}I}) = L_1(\widehat{I})$. Thus, by Proposition 3, $L_1(\widehat{B}) \leq 4 \ln N$.

For $g > 1$ the proof is more involved, details are given in Sect. 4.1.1. ■

We bound the L_1-norm of the Fourier transform of an interval I.

Proposition 3 *Let I be the characteristic function of an interval in $\mathbb{Z}_N$. Then $L_1(\widehat{I}) < 4 \ln N$.*

Proof: By [AGS03], the Fourier coefficients of any length at most $N/2$ interval I' are upper bounded by $\left|\widehat{I'}(\alpha)\right| \leq 1/\mathsf{abs}(\alpha)$ for $\mathsf{abs}(\alpha) = \min\{\alpha, N - \alpha\}$, and $\left|\widehat{I}(0)\right| \leq 1$. For longer intervals I, we write I as the sum of two intervals each of length at most $N/2$. By linearity of the Fourier transform, the Fourier coefficients of I are the sum of the Fourier coefficients of these intervals, and are therefore bounded by $\widehat{I}(\alpha) \leq 2/\mathsf{abs}(\alpha)$. We conclude that $L_1(\widehat{I}) \leq 1 + 2\sum_{\alpha=1}^{N/2}(2/\mathsf{abs}(\alpha)) < 4 \ln N$ (where for the last inequality we use the bound $\sum_{i=1}^{n}(1/i) \leq 1 + \ln n$ on harmonic numbers). ■

4.1.1 Proof of Lemma 7

We prove Lemma 7 for the case of rank 1 Bohr sets in $\mathbb{Z}_N$ with a multiplier α that is not co-prime to N.

Proof of Lemma 7. Fix a rank 1 Bohr set $B \overset{def}{=} \{x \in \mathbb{Z}_N \mid (\alpha x \bmod N) \in I\}$ in $\mathbb{Z}_N$ for $\alpha \in \mathbb{Z}_N$ and $I = [s..t] \subseteq \mathbb{Z}_N$. Denote by $g \overset{def}{=} gcd(N, \alpha)$ the greatest common divisor of N and α, and let $\beta \overset{def}{=} \alpha/g$. We prove that $L_1(\widehat{B}) \leq 4 \ln N$, focusing on the case that $g > 1$.

By Claim 8, for $J \overset{def}{=} [\lceil(\frac{s}{g})\rceil..\lfloor(\frac{t}{g})\rfloor]$ we can rewrite B as

$$B \;=\; \left\{\beta^{-1}(i\frac{N}{g} + x_0) \in \mathbb{Z}_N \,\middle|\, x_0 \in J, i \in [0..g-1]\right\} \tag{3}$$

for β^{-1} the inverse of β modulo N. Denote

$$\widehat{J}_{N/g}(\alpha) \overset{def}{=} \frac{1}{N/g} \sum_{x \in J} e(\alpha x/(N/g)).$$

By Claim 9, for every index $\gamma \in \mathbb{Z}_N$, the γ-Fourier coefficient of B is

$$\widehat{B}(\gamma) \;=\; \begin{cases} 0 & g \not| \gamma \\ \widehat{J}_{N/g}(\gamma\beta^{-1}/g) & g|\gamma \end{cases}$$

Thus,

$$L_1(\widehat{B}) = \sum_{\gamma' \in \mathbb{Z}_{N/g}} \widehat{J_{N/g}}((\gamma' g)\beta^{-1}/g) = \sum_{\gamma' \in \mathbb{Z}_{N/g}} \left|\widehat{J_{N/g}}(\gamma'\beta^{-1})\right| = \sum_{\gamma' \in \mathbb{Z}_{N/g}} \left|\widehat{J_{N/g}}(\gamma')\right|$$

where the last equality holds because β^{-1} is co-prime to N/g (because it is co-prime to N, and N/g is a divisor of N). By the definition of $\widehat{J_{N/g}}$, $\sum_{\gamma' \in \mathbb{Z}_{N/g}} \left|\widehat{J_{N/g}}(\gamma')\right|$ is equal to $L_1(\widehat{J})$ where the Fourier transform of J is computed with respect to the characters of $\mathbb{Z}_{N/g}$. By Proposition 3, the latter is upper bounded by $4\ln(N/g) \leq 4\ln N$. We conclude that $L_1(\widehat{B}) \leq 4\ln N$. ■

Claim 8 $B = \left\{ \beta^{-1}(i\frac{N}{g} + x_0) \in \mathbb{Z}_N \,\middle|\, x_0 \in J, i \in [0..g-1] \right\}$.

Proof: Since $\left\{\beta^{-1}x\right\}_{x \in \mathbb{Z}_N} = \mathbb{Z}_N$ (as by the maximality of g, $gcd(N, \beta) = 1$), then B is equal to the set $\left\{ \beta^{-1}x \in \mathbb{Z}_N \mid (\alpha(\beta^{-1}x) \bmod N) \in I \right\}$. Since $\alpha(\beta^{-1}x) = (g\beta)(\beta^{-1}x) = gx$ we get that

$$B = \left\{ \beta^{-1}x \in \mathbb{Z}_N \mid (gx \bmod N) \in I \right\}.$$

Thus, by Proposition 4, for $J = [\lceil(\frac{s}{g})\rceil..\lfloor(\frac{t}{g})\rfloor]$,

$$B = \left\{ \beta^{-1}x \in \mathbb{Z}_N \mid (x \bmod \frac{N}{g}) \in J \right\}.$$

Expressing each $x \in \mathbb{Z}_N$ according to its division with remainder by N/g, i.e., $x = i\frac{N}{g} + x_0$ for $i = \lfloor(x/(N/g))\rfloor \in [0..g-1]$ and $x_0 = x - i\frac{N}{g} \in [0..\frac{N}{g} - 1]$, we get that

$$B = \left\{ \beta^{-1}(i\frac{N}{g} + x_0) \in \mathbb{Z}_N \,\middle|\, x_0 \in J, i \in [0..g-1] \right\}.$$

■

Claim 9 *For every index $\gamma \in \mathbb{Z}_N$, if $g \nmid \gamma$, then $\widehat{B}(\gamma) = 0$, and $\widehat{B}(\gamma) = \widehat{J}_{N/g}(\gamma\beta^{-1}/g)$ otherwise.*

Proof: Fix $\gamma \in \mathbb{Z}_N$. By definition of the Fourier transform in $\mathbb{Z}_N$ and the set B,

$$\widehat{B}(\gamma) \quad = \quad \frac{1}{N}\sum_{x \in B} e(\gamma x/N) = \frac{1}{N}\sum_{x_0 \in J}\sum_{i \in [0..g-1]} e(\gamma\beta^{-1}(i\frac{N}{g} + x_0)/N)$$

where the last equality holds by Eq. 3. Rearranging this sum we get that

$$\begin{aligned} \widehat{B}(\gamma) &= \frac{1}{N}\sum_{x_0 \in J} e(\gamma\beta^{-1}x_0/N) \cdot \sum_{i \in [0..g-1]} e(\gamma\beta^{-1}i/g) \\ &= \begin{cases} 0 & g \nmid \gamma\beta^{-1} \\ \frac{g}{N}\sum_{x_0 \in J} e(\gamma\beta^{-1}x_0/N) & otherwise \end{cases} \end{aligned}$$

where the last equality holds since $\frac{1}{g}\sum_{i \in [0..g-1]} e(\gamma\beta^{-1}i/g) = 0$ if $g \nmid \gamma\beta^{-1}$, and it is 1 otherwise (see Sect. 3.1).

Note that $g|\gamma\beta^{-1}$ iff $g|\gamma$: This trivially holds for $g = 1$, and holds for $g > 1$, since $g \nmid \beta$ (because $g|N$ and $gcd(N, \beta) = 1$). We conclude that $\widehat{B}(\gamma) = 0$ if $g \nmid \gamma$.

We focus next on the case that $g|\gamma$. Denote $\gamma' = \gamma/g$. Substituting γ with $\gamma' g$ in the above and rearranging, we get that, $\widehat{B}(\gamma) = \frac{g}{N}\sum_{x_0 \in J} e(\gamma'\beta^{-1}x_0/(N/g))$. This in turn is equal to $\widehat{J}_{N/g}(\gamma'\beta^{-1})$ (by definition of $\widehat{J}_{N/g}$). We conclude that $\widehat{B}(\gamma) = \widehat{J}_{N/g}(\gamma'\beta^{-1}) = \widehat{J}_{N/g}(\gamma\beta^{-1}/g)$ if $g|\gamma$. ■

Proposition 4 *Fix any positive integers N, g s.t. $g|N$, and any $x \in \mathbb{Z}_N$ and $I = [s..t] \subseteq \mathbb{Z}_N$. Then $(gx \bmod N) \in I$ iff $(x \bmod \frac{N}{g}) \in [\lceil(\frac{s}{g})\rceil..\lfloor(\frac{t}{g})\rfloor]$.*

Proof: Observe that since $g|N$, then $(gx \bmod N) = (x \bmod \frac{N}{g}) \cdot g$. (Because $x = x_1\frac{N}{g} + x_2$ for $x_1 < g$, $x_2 < \frac{N}{g}$, implying that $(x \bmod \frac{N}{g}) = x_2$ and $(gx \bmod N) = (x_1 N + gx_2 \bmod N) = gx_2$.) Thus, $(gx \bmod N) \geq s$ iff $(x \bmod \frac{N}{g}) \geq s/g$, which in turn happens iff $(x \bmod \frac{N}{g}) \geq \lceil(s/g)\rceil$ (because $x \bmod \frac{N}{g}$ is an integer). Similarly, $(gx \bmod N) \leq t$ iff $(x \bmod \frac{N}{g}) \leq \lfloor(t/g)\rfloor$. We conclude that $(gx \bmod N) \in I$ iff $(x \bmod \frac{N}{g}) \in [\lceil(\frac{s}{g})\rceil..\lfloor(\frac{t}{g})\rfloor]$. ■

4.2 Approximating Distributions Uniform over Arithmetic Progressions

We show that there are explicit small sets approximating the uniform distribution over a given arithmetic progression.

Theorem 2 Part 2 (Approximating arithmetic progressions). For any integer $N > 0$, real $\varepsilon \in (0,1)$, and a length $M \leq N$ arithmetic progression B in $\mathbb{Z}_N$, there is an explicit set $A \subseteq \mathbb{Z}_N$ of size polynomial in $1/\varepsilon$ and $(\ln M)^2$ that ε-approximates U_B in $\mathbb{Z}_N$.

Proof: Fix N, M and ε. We focus here on the case $B = [0..M-1]$ is an *interval* starting at zero; extensions to arbitrary arithmetic progressions appear in section 4.2.1.

Let $\varepsilon' = \varepsilon\rho^2/2$ for $\rho = \varepsilon/8\pi$. We show below that if $D_{A,\mathbb{Z}_M}(\mathcal{B}_{M,2}) < \varepsilon'$ (i.e., A has ε'-discrepancy on all rank 2 Bohr sets in $\mathbb{Z}_M$), then A ε-approximates U_B. This completes our proof, as by Theorem 2 Part 1 there exists sets A of size polynomial in $1/\varepsilon' = O(1/\varepsilon^3)$ and $(\ln M)^2$ s.t. $D_{A,\mathbb{Z}_M}(\mathcal{B}_{M,2}) < \varepsilon'$.

To relate approximating the distribution U_B over $\mathbb{Z}_N$ to having ε'-discrepancy on all rank 2 Bohr sets in $\mathbb{Z}_M$, we do the following: For each $\alpha \in \mathbb{Z}_N$, denote by (q_α, r_α) its quotient and remainder in division-with-remainder by (the typically non-integer value) N/M. That is, q_α is the largest integer s.t. $q_\alpha \frac{N}{M} \leq \alpha$ and $r_\alpha = \alpha - q_\alpha \frac{N}{M}$ is the remainder. Write α as:

$$\alpha = N\left(\frac{q_\alpha}{M} + \frac{r_\alpha}{N}\right)$$

Rewrite each linear test $\chi_\alpha(x) = e(\alpha x/N)$ in $\mathbb{Z}_N$ as:

$$\chi_\alpha(x) = e\left(\left(\left(\frac{q_\alpha x}{M}\right)_1 + \left(\frac{r_\alpha x}{N}\right)_1\right)_1\right)$$

by replacing αx with $N\left(\frac{q_\alpha x}{M} + \frac{r_\alpha x}{N}\right)_1 = N\left(\left(\frac{q_\alpha x}{M}\right)_1 + \left(\frac{r_\alpha x}{N}\right)_1\right)_1$. By Lemma 10, if $D_{A,\mathbb{Z}_M}(\mathcal{B}_{M,2}) < \varepsilon'$, then the joint distribution of pairs $\left(\left(\frac{q_\alpha x}{M}\right)_1, \left(\frac{r_\alpha x}{N}\right)_1\right)$ with x drawn uniformly at random from A is "close" to their distribution with x drawn uniformly at random from B. Closeness is in the sense that for every two length ρ intervals $J_1, J_2 \subseteq [0,1]$,

$$\left|\Pr_{x\in A}\left[\left(\frac{q_\alpha x}{M}\right)_1 \in J_1 \;\&\; \left(\frac{r_\alpha x}{N}\right)_1 \in J_2\right] - \Pr_{x\in B}\left[\left(\frac{q_\alpha x}{M}\right)_1 \in J_1 \;\&\; \left(\frac{r_\alpha x}{N}\right)_1 \in J_2\right]\right| < \varepsilon'$$

By Lemma 11, if the above equation holds, then A $(\frac{\varepsilon'}{\rho^2} + 4\pi\rho)$-approximates U_B. Noting that $(\frac{\varepsilon'}{\rho^2} + 4\pi\rho) = \varepsilon$, we conclude that $D_{A,\mathbb{Z}_M}(\mathcal{B}_{M,2}) < \varepsilon'$ implies that A ε-approximates U_B. ■

Lemma 10 *For any positive integers $M \leq N$, positive real ε and a subset $A \subseteq \mathbb{Z}_M$, if $D_{A,\mathbb{Z}_M}(\mathcal{B}_{M,2}) < \varepsilon'$, then for all integers $\alpha \in [0, M)$, reals $\beta \in [0, N/M)$, and intervals $J_1, J_2 \subseteq [0,1]$,*

$$\left|\Pr_{x\in A}\left[\left(\frac{\alpha x}{M}\right)_1 \in J_1 \textit{ and } \left(\frac{\beta x}{N}\right)_1 \in J_2\right] - \Pr_{x\in B}\left[\left(\frac{\alpha x}{M}\right)_1 \in J_1 \textit{ and } \left(\frac{\beta x}{N}\right)_1 \in J_2\right]\right| < \varepsilon' \qquad (4)$$

Remark. In particular, Eq. 4 holds for any $\alpha, \beta, J_1, J_2, \varepsilon'$ s.t. $(\alpha, \beta) = (q_z, r_z)$ are the quotient and remainder in the division-with-remainder by N/M of some $z \in \mathbb{Z}_N$, $J_1, J_2 \subseteq [0,1]$ and length ρ intervals (because q_z is an integer in $[0, M)$ and r_z is a real in $[0, N/M)$).

Proof: Fix parameters α, β, J_1, J_2 and ε' for Eq. 4. Fix $A \subseteq \mathbb{Z}_M$ s.t. $D_{A,\mathbb{Z}_M}(\mathcal{B}_{M,2}) < \varepsilon'$. In the following we first show that Eq. 4 above holds iff Eq. 5 below holds. We then argue that Eq. 5 holds as it bounds the discrepancy of A on a rank 2 Bohr set in $\mathbb{Z}_M$. We conclude that Eq. 4 holds.

We map the sets $\left(\frac{\alpha B}{M}\right)_1$ and $\left(\frac{\beta B}{N}\right)_1$ into $\mathbb{Z}_M$ by the one-to-one mappings:

$$(i) \qquad \left(\frac{\alpha x}{M}\right)_1 \mapsto \alpha x \bmod M$$

$$(ii) \qquad \left(\frac{\beta x}{N}\right)_1 \mapsto x$$

(The latter map is a one-to-one because $\beta < N/M$ and hence for $x \in B = [0..M-1]$, $\left(\frac{\beta x}{N}\right)_1$ does not wrap around 0.) Denote by $\bar{J}_1$ and $\bar{J}_2$ the intervals in $\mathbb{Z}_M$ that are the images of the sets $J_1' = J_1 \cap \left(\frac{\alpha B}{M}\right)_1$ and $J_2' = J_2 \cap \left(\frac{\beta B}{N}\right)_1$ under mappings (i) and (ii), respectively. Since mappings (i),(ii) are one-to-one, then Eq. 4 holds iff the following holds:

$$\left|\Pr_{x\in A}\left[(\alpha x \bmod M) \in \bar{J}_1 \text{ and } x \in \bar{J}_2\right] - \Pr_{x\in B}\left[(\alpha x \bmod M) \in \bar{J}_1 \text{ and } x \in \bar{J}_2\right]\right| < \varepsilon' \qquad (5)$$

Observing that the left hand side of Eq. 5 is the discrepancy of A on the rank 2 Bohr set

$$R = \left\{ x \in \mathbb{Z}_M \mid (\alpha x \bmod M) \in \bar{J}_1 \text{ and } x \in \bar{J}_2 \right\},$$

we conclude that Eq. 5 holds and hence Eq. 4 holds. ■

Lemma 11 *If Eq. 4 holds for every integer $\alpha \in [0, M)$, real $\beta \in [0, N/M)$, and length ρ intervals $J_1, J_2 \subseteq [0, 1]$, then A $(\frac{\varepsilon'}{\rho^2} + 4\pi\rho)$-approximates U_B in $\mathbb{Z}_N$.*

Proof Sketch. Intuitively, if the distribution of pairs $\left(\left(\frac{q_\alpha x}{M}\right)_1, \left(\frac{r_\alpha x}{N}\right)_1\right)$ with x drawn from U_A is "close" to the distribution of such pairs with x drawn from U_B, then the distribution of sums $\chi_\alpha(x) = \left(\left(\frac{q_\alpha x}{M}\right)_1 + \left(\frac{r_\alpha x}{N}\right)_1\right)_1$ over $x \in A$ is "close" to the distribution of such sums over $x \in B$. Namely, the expected value of $\chi_\alpha(x)$ with x drawn from U_A is "close" to its expected value with x drawn from U_B.

Capturing this intuition requires some technical work, where we rewrite $\mathbb{E}_{x \in S}\, \chi_\alpha(x)$, $S = A, B$, as a sum over expressions depending on pairs of intervals J_1, J_2, and use Eq. 4 (and other manipulations) to bound these expressions. Details are omitted from this extended abstract. ■

4.2.1 From Approximating Intervals to Approximating Arithmetic Progressions

We show that given any length M arithmetic progression $P = P_{\alpha,[s..s+M-1]}$ and any set A that ε-approximates $U_{[0..M-1]}$ in $\mathbb{Z}_N$, the set $A' = \alpha(A + s)$ is a set of size $|A'| \le |A|$ that ε-approximates U_P.

Lemma 12 *For any positive integers $M \le N$ and any length M arithmetic progression $P = P_{\alpha,[s..M-1+s]}$ in $\mathbb{Z}_N$, if $A \subseteq \mathbb{Z}_N$ ε-approximates $U_{[0..M-1]}$, then $A' \overset{def}{=} \alpha(A + s)$ ε-approximates U_P.*

Proof: For any $\beta \in \mathbb{Z}_N$ and subsets $S, S' \subseteq \mathbb{Z}_N$, denote $\text{diff}_\beta(S, S') \overset{def}{=} \mathbb{E}_{x \in S}\, e(\beta x/N) - \mathbb{E}_{x \in S'}\, e(\beta x/N)$. We prove that $|\text{diff}_\beta(A', P)| < \varepsilon$ for all $\beta \in \mathbb{Z}_N$. Fix β. By definition of A' and P, $|\text{diff}_\beta(A', P)| = |\,\mathbb{E}_{x \in A}\, e(\beta\alpha(x+s)/N) - \mathbb{E}_{x \in [0..M-1]}\, e(\beta\alpha(x+s)/N)|$. The latter is equal to $|e(\beta\alpha s/N)|\,|\,\mathbb{E}_{x \in A}\, e(\beta\alpha x/N) - \mathbb{E}_{x \in [0..M-1]}\, e(\beta\alpha x/N)| = 1 \cdot |\text{diff}_{\alpha\beta}(A, [0..M-1])| < \varepsilon$ where the last inequality holds, because A ε-approximates $U_{[0..M-1]}$ means that $|\text{diff}_{\beta'}(A, [0..M-1])| < \varepsilon$ for all $\beta' \in \mathbb{Z}_N$. ■

5 Deterministically Finding Significant Fourier Coefficients

We present our deterministic and robust SFT algorithm, and prove Theorem 1.

5.1 The SFT Algorithm

At a high level, our SFT algorithm is a binary search algorithm that repeatedly: (1) Partitions the set of potentially significant Fourier coefficients into two subsets. (2) Tests each subset to decide if it (potentially) contains a significant Fourier coefficient. (3) Continues recursively on any subset with a positive test result.

Elaborating on the above, at each step of this search, the set of potentially significant Fourier coefficients is maintained as a collection $\mathcal{J}$ of intervals, starting with $\mathcal{J}$ containing the four intervals $[i\frac{N}{4}..(i+1)\frac{N}{4}]$, $i = 0, \ldots, 3$. To maintain efficiency, the intervals $[a..b]$ are represented by their endpoints pairs $\{a, b\}$. The partition step partitions every interval $J = [a..b] \in \mathcal{J}$ into its lower and upper halves: $J_1 = [a..c]$ and $J_2 = [c + 1..b]$ for $c = \lfloor((a + b)/2)\rfloor$ its center. For $i = 1, 2$, the test estimates the Fourier weight of f on J_i, denoted $\widehat{f}(J_i)^2 \overset{def}{=} \sum_{\alpha \in J_i} \left|\widehat{f}(\alpha)\right|^2$, returning YES if this weight exceeds the significance threshold τ. The set $\mathcal{J}$ is updated by removing J, and inserting each J_i ($i = 1, 2$) iff it passes the test. After $O(\log N)$ steps this search terminates with a collection $\mathcal{J}$ of length one intervals containing the frequencies of the (potentially) significant Fourier coefficients. The algorithm wraps up by executing a sieving step, where for each frequency α of a potentially significant Fourier coefficient, a $O(\tau)$-approximation for $\widehat{f}(\alpha)$ is computed: $val_\alpha = \frac{1}{|A|}\sum_{x \in A - y} f(x)\overline{\chi_\alpha(x)}$ (for an arbitrary $y \in \cup_{\ell=1}^{\lfloor(\log N)\rfloor} B_\ell$); and the algorithm outputs the pairs (α, val_α) for all α found to be significant, i.e., $val_\alpha \ge \tau/2$.

The heart of the algorithm is the test deciding which intervals potentially contain a significant Fourier coefficient. This test, given an interval $J = [a..b]$, answers YES if the Fourier weight on J, exceeds the significance threshold τ, and answers NO if the Fourier weight of a slightly larger interval $J' \supseteq J$ is less than $\tau/2$. This is achieved by estimating the ℓ_2 norm (i.e., sum of squared Fourier coefficients) of a filtered version of the input function f, when using a filter h that passes Fourier coefficients in J and decays fast outside of J. The filters we use are functions $h = h_{\ell,c}$ which are (normalized) periodic square functions with support size 2^ℓ when phase shifted by $-c$, for $c = \lfloor((a+b)/2)\rfloor$ the center of J and $\ell = \lfloor(\log\frac{N}{4(b-a)})\rfloor$ a function of J's length:

$$h_{\ell,c}(y) \overset{def}{=} \begin{cases} \frac{N}{2^\ell}\cdot\chi_{-c}(y) & y\in[0..2^\ell-1] \\ 0 & otherwise \end{cases} \tag{6}$$

The filter $h_{\ell,c}$ passes all frequencies that lie within the length $N/2^{\ell+2}$ interval J centered around c, and decays fast outside of J. The filtered version of f is $f*h$, and we estimate its ℓ_2 norm $\|f*h\|_2^2$ by the estimator:

$$\mathsf{est}_{\ell,c}(f) \overset{def}{=} \frac{1}{|A|}\sum_{x\in A}\left(\frac{1}{|B_\ell|}\sum_{y\in B_\ell}\chi_{-c}(y)\overline{f(x-y)}\right)^2 \tag{7}$$

for A a small biased set in $\mathbb{Z}_N$, and B_ℓ approximating the uniform distribution over $[0..2^\ell-1]$ in $\mathbb{Z}_N$.

A pseudo-code of the algorithm follows.

Algorithm 5 SFT.
Input: $N\in\mathbb{N}$, $\tau\in(0,1]$, oracle access to $f\colon\mathbb{Z}_N\to\mathbb{C}$.

```
1. Initialize: $L=\phi$; $\mathcal{J}\leftarrow\left\{\left\{i\frac{\lambda N}{2},(i+1)\frac{\lambda N}{2}\right\}\right\}_{i=0}^{\frac{2}{\lambda}-1}$ for $\lambda=1/2$; $\gamma=\tau/(49t^2\ln N)$
2. Construct queries:
     • $A := AIKPS(N,\gamma)$
     • For $\ell=1,...,\log N$, $B_\ell := AIKPS(2^\ell,\gamma_\ell)$ for $\gamma_\ell := \gamma^3/(128\pi^2\ell^2)$
   Remark: $AIKPS(N,\varepsilon)$ is as specified in Sect. 3.2, Eq. 1.
3. Main Loop: While $\exists\{a,b\}\in\mathcal{J}$ s.t. $b\neq a$ do:
   (a) Remove $\{a,b\}$ from $\mathcal{J}$, denote $c'=\lfloor(\frac{a+b}{2})\rfloor$
   (b) For each pair $\{a',b'\}$ out of the pairs $\{a,c'\}$ and $\{c'+1,b\}$ do:
        i. Let $\ell=\lfloor(\log\frac{\lambda N}{2(b'-a')})\rfloor$ and $c=\lfloor(\frac{a'+b'}{2})\rfloor$
       ii. Compute $\mathsf{est}_{\ell,c}\leftarrow\frac{1}{|A|}\sum_{x\in A}\left(\frac{1}{|B_\ell|}\sum_{y\in B_\ell}\chi_{-c}(y)\overline{f(x-y)}\right)^2$
      iii. If $\mathsf{est}_{\ell,c}\geq\tau/2$, insert $\{a',b'\}$ to $\mathcal{J}$
4. Sieving: For each $\{a,b\}\in\mathcal{J}$, and each $\alpha\in[a..b]$
   (a) Compute $val(\alpha)\leftarrow\left|\frac{1}{|A|}\sum_{x\in A}\chi_\alpha(x)\overline{f(x)}\right|^2$
   (b) If $val(\alpha)\geq\tau/2$, insert $\alpha$ to $L$
5. Output $\{(\alpha,val_\alpha)\}_{\alpha\in L}$
```

Remarks. (1) To keep this paper self contained, in Step 2 of the pseudo-code we use the small biased sets of [AIK$^+$90] which were specified in Sect. 3.2, Eq. 1. Nevertheless, any other construction of small biased sets may be used (with set sizes $|A|, |B_\ell|$ varying accordingly). (2) To simplify notations, we assume without loss of generality that $0\in\cup_\ell B_\ell$ (otherwise add it), and $\|f\|_2 = 1$ (otherwise normalize f by dividing each read value by an energy estimator $\frac{1}{|A|}\sum_{x\in A}\overline{f(x)}^2$).

5.2 Proof of Theorem 1

The proof of Theorem 1 is simple given our new result on explicit sets approximating given arithmetic progressions in $\mathbb{Z}_N$ (Theorems 2 and 4) together with the work of [Aka09]:

Proof:[Proof of Theorem 1.] Our algorithm builds on the algorithm of [Aka09] while replacing their randomized construction of sets $S=\bigcup_\ell(A-B_\ell)$ with a deterministic construction. Our deterministic construction produces a set S which is (N,τ,t)-good (see Corollary 13 below). When S is (N,τ,t)-good, the analysis of [Aka09] shows that the algorithm succeeds (see Theorem 14 below). Namely, the algorithm outputs $L\supseteq \mathsf{Heavy}_\tau(f)$ (with probability at least $1-1/N^{\Theta(1)}$ over the random noise η) in running time is $O(\frac{1}{\tau^2}\log N\cdot|S|)$. Finally, this running time polynomial in $\log N$, $1/\tau$ and t by the definition of good sets. ■

As a corollary of our results on explicit constructions (Theorems 2,4), the queries constructed in our SFT algorithm are (N,τ,t)-good.

Corollary 13 *The set $S = \bigcup_\ell (A - B_\ell)$ for A, B_ℓ as in Algorithm 5 is an (N, τ, t)-good set.*

Proof: The set S is good, because: (i) A is a γ-biased set in $\mathbb{Z}_N$. (ii) B_ℓ γ-approximates $U_{[0..2^\ell-1]}$ (by Theorem 4 and the fact that B_ℓ is a $\gamma_\ell = \gamma^3/(128\pi^2\ell^2)$-biased set in $\mathbb{Z}_{2^\ell}$). Finally, $|A|, |B_\ell|$ are polynomial in $\log N$ and $1/\gamma = O(t^2 \ln N/\tau)$. ■

Theorem 14 ([Aka09]) *If the queries $S = \bigcup_\ell (A - B_\ell)$ for A, B_ℓ as in Algorithm 5 are (N, τ, t)-good, then the following holds. Given N, τ, t, and $\{(x, f'(x))\}_{x \in S}$ for $f' = f + \eta$ a complex valued function over $\mathbb{Z}_N$ s.t. $L_1(\widehat{f}) \leq t$ and η is a $\tau/3$-random noise, Algorithm 5 outputs $\{(\alpha, val_\alpha)\}_{\alpha \in L}$ s.t. (with probability at least $1 - 1/N^{\Theta(1)}$ over the random noise η) $L \supseteq \mathsf{Heavy}_\tau(f)$ and val_α are $O(\tau)$-approximations of $\widehat{f}(\alpha)$. The running time of the algorithm is $O(\frac{1}{\tau^2}\log N \cdot |S|)$.*

6 Conclusions

We presented the first deterministic SFT algorithm for functions over $\mathbb{Z}_N$, which is: (1) Local, i.e., its running time is polynomial in $\log N$, $1/\tau$ and $L_1(\widehat{f})$; and (2) Robust to random noise. Our main technical novelty lies in proving that there exists explicit constructions of small sets with small discrepancy in all rank d Bohr sets, as well as small sets approximating the uniform distribution over a given arithmetic progression.

Extensions. Our deterministic SFT algorithm extends to handle functions over all finite abelian groups G (given by their generators $g_1, \ldots, g_k$ and their orders $N_1, \ldots, N_k$). This extension is motivated for example by functions over domains $G = \mathbb{Z}_{N_1} \times \ldots \times \mathbb{Z}_{N_k}$ arising in image/video processing applications ($k = 2, 3$) and machine learning applications (large k). As a central ingredient for this extension we present explicit small sets approximating the uniform distribution over rectangles $R_{t,\ell} = [0..N_1] \times \ldots \times [0..N_{t-1}] \times [0..2^\ell - 1] \times \{0\}^{k-t}$ for given $t \in [k]$ and $\ell \in [\log N_t]$. Details are omitted from this extended abstract.

Applications to sparse Fourier approximation, compressed sensing and sketching. Using our SFT algorithm we obtain deterministic and robust algorithms for sparse Fourier approximation, compressed sensing and sketching. Our algorithms, given N, m, ε, t and oracle access to a signal $x \in \mathbb{C}^N$ s.t. $L_1(\widehat{x}) \leq t$ (resp., a sketch Ax for $A = A_{N,m,\varepsilon,t} \in \mathbb{C}^{poly(\log N, m/\varepsilon, t) \times N}$ an explicit measurement matrix), output a near optimal m-sparse approximation R of x, i.e., $\|x - R\|_2^2 \leq \|x - R_{opt}\|_2^2 + \varepsilon$ for R_{opt} the best m-terms approximation of x in the Fourier (resp., standard) basis. Our algorithms are robust to $m/3\varepsilon$-random noise, and their running time is polynomial in $\log N$, m/ε and t. Given our SFT algorithm, the derivation of these algorithms is standard; details are omitted from this extended abstract.

Open questions. Our results on explicit constructions yields sets sizes that are efficient but not optimal: Probabilistic method arguments show that there are randomized constructions of sets of size $O(d \log N/\varepsilon^2)$ with ε-discrepancy in all rank d Bohr sets, as well as sets of size $O((\log N)/\varepsilon^2)$ ε-approximating the uniform distribution over a given arithmetic progression. Whether these bounds can be matched by explicit constructions is an open problem.

Acknowledgments

The author is most grateful to Noga Alon, Shafi Goldwasser, Venkat Guruswami, Piotr Indyk, Mark Iwen, Martin Strauss and Avi Wigderson for helpful comments and discussions.

References

[AGS03] A. Akavia, S. Goldwasser, and S. Safra. Proving Hard-Core Predicates using List Decoding. In *Proc. of 44th IEEE Annual Symposium on Foundations of Computer Science (FOCS'03)*, pages 146–157. IEEE Computer Society, 2003.

[AIK$^+$90] M. Ajtai, H. Iwaniec, J. Komlos, J. Pintz, and E. Szemeredi. Constructions of a this set with small fourier coefficients. *Bull. London Math. Soc.*, 22:583–590, 1990.

[Aka09] A. Akavia. Solving Hidden Number Problem with One Bit Oracle and Advice. In *proceedings of the 29th Annual International Cryptology Conference (Crypto'09)*, 2009.

[AM95] N. Alon and Y. Mansour. ε-discrepancy sets and their application for interpolation of sparse polynomials. *IPL: Information Processing Letters*, 54, 1995.

[CM05] G. Cormode and S. Muthukrishnan. Towards an algorithmic theory of compressed sensing. In *DIMACS TR: 2005-25*, 2005.

[CM06] G. Cormode and S. Muthukrishnan. Combinatorial algorithms for compressed sensing. In Paola Flocchini and Leszek Gasieniec, editors, *SIROCCO*, volume 4056 of *Lecture Notes in Computer Science*, pages 280–294. Springer, 2006.

[CRT06] E. J. Candes, J. K. Romberg, and T. Tao. Robust uncertainty principles: exact signal reconstruction from highly incomplete frequency information. *IEEE Transactions on Information Theory*, 52(2):489–509, 2006.

[CT65] J.W. Cooley and J.W. Tukey. An algorithm for machine calculation of complex fourier series. *Mathematics of Computation*, 19:297–301, Apr 1965.

[DeV07] R. A. DeVore. Deterministic constructions of compressed sensing matrices. *J. Complex.*, 23(4-6):918–925, 2007.

[Don05] D. Donoho. Compressed sensing. *IEEE Trans. on Information Theory*, 42(4):1289–1306, April, 2005.

[GGI$^+$02] A. C. Gilbert, S. Guha, P. Indyk, S. Muthukrishnan, and M. Strauss. Near-optimal sparse fourier representations via sampling. In *Proc. of 34 ACM Annual Symposium on Theory of Computing (STOC'02)*, pages 152–161. ACM Press, 2002.

[GL89] O. Goldreich and L. Levin. A hard-core predicate for all one-way functions. In *Proc. 27th ACM Annual Symposium on Theory of Computing (STOC'89)*, pages 25–32, 1989.

[GLR08] V. Guruswami, J. R. Lee, and A. Razborov. Almost euclidean subspaces of ℓ_1 via expander codes. In *SODA '08: Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 353–362, Philadelphia, PA, USA, 2008. Society for Industrial and Applied Mathematics.

[GM06] Sumit Ganguly and Anirban Majumder. Cr-precise: A deterministic summary structure for update data streams. *CoRR*, abs/cs/0609032, 2006.

[GMS05] A. C. Gilbert, S. Muthukrishnan, and M. Strauss. Improved time bounds for near-optimal sparse fourier representation via sampling. In *in Proc. SPIE Wavelets XI*, 2005.

[GSTV06] A. C. Gilbert, M. J. Strauss, J. A. Tropp, and R. Vershynin. Algorithmic linear dimension reduction in the l_1 norm for sparse vectors. *CoRR*, abs/cs/0608079, 2006.

[Ind07] P. Indyk. Sketching, streaming and sub-linear space algorithms. graduate course notes, available at http://stellar.mit.edu/s/course/6/fa07/6,895/, 2007.

[Ind08] P. Indyk. Explicit constructions for compressed sensing of sparse signals. In *SODA'08: Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 30–33, Philadelphia, PA, USA, 2008. Society for Industrial and Applied Mathematics.

[IS08] M. A. Iwen and C. V. Spencer. Improved bounds for a deterministic sublinear-time sparse fourier algorithm. In *Conference on Information Sciences and Systems (CISS), Princeton, NJ*, 2008.

[Iwe07] M. A. Iwen. A deterministic sub-linear time sparse fourier algorithm via non-adaptive compressed sensing methods. *CoRR*, abs/0708.1211, 2007.

[Iwe08] M. A. Iwen. A deterministic sub-linear time sparse fourier algorithm via non-adaptive compressed sensing methods. In *SODA '08: Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 20–29, Philadelphia, PA, USA, 2008. Society for Industrial and Applied Mathematics.

[Kat89] M. Katz. An estimate for characters sum. *J. AMS*, 2(2):197–200, 1989.

[KM93] E. Kushilevitz and Y. Mansour. Learning decision trees using the Fourier spectrum. *SICOMP*, 22(6):1331–1348, 1993.

[Man95] Y. Mansour. Randomized interpolation and approximation of sparse polynomials. *SIAM J. on Computing*, 24(2):357–368, 1995.

[Mut03] S. Muthukrishnan. Data streams: Algorithm and applications (invited talk at soda'03). available at http://athos.rutgers.edu/ muthu/stream-1-1.ps, 2003.

[RSW93] A. Razborov, E. Szemeredi, and A. Wigderson. Constructing small sets that are uniform in arithmetic progressions. *Combinatorics, Probability & Computing*, 2:513–518, 1993.

[TV06] Terence Tao and Van H. Vu. *Additive Combinatorics*. Series: Cambridge Studies in Advanced Mathematics (No. 105), September 2006.