Deterministic Sparse Fourier Transform with an ℓ_∞ Guarantee

Yi Li
Nanyang Technological University, Singapore, Singapore
yili@ntu.edu.sg

Vasileios Nakos
Universität des Saarlandes, Saarbrücken, Germany
Max Planck Institut für Informatik, Saarland Informatics Campus, Saarbrücken, Germany
vnakos@mpi-inf.mpg.de

Abstract

In this paper we revisit the deterministic version of the Sparse Fourier Transform problem, which asks to read only a few entries of $x \in \mathbb{C}^n$ and design a recovery algorithm such that the output of the algorithm approximates $\widehat{x}$, the Discrete Fourier Transform (DFT) of x. The randomized case has been well-understood, while the main work in the deterministic case is that of Merhi et al. (J Fourier Anal Appl 2018), which obtains $O(k^2 \log^{-1} k \cdot \log^{5.5} n)$ samples and a similar runtime with the ℓ_2/ℓ_1 guarantee. We focus on the stronger ℓ_∞/ℓ_1 guarantee and the closely related problem of incoherent matrices. We list our contributions as follows.

1. We find a deterministic collection of $O(k^2 \log n)$ samples for the ℓ_∞/ℓ_1 recovery in time $O(nk \log^2 n)$, and a deterministic collection of $O(k^2 \log^2 n)$ samples for the ℓ_∞/ℓ_1 sparse recovery in time $O(k^2 \log^3 n)$.
2. We give new deterministic constructions of incoherent matrices that are row-sampled submatrices of the DFT matrix, via a derandomization of Bernstein's inequality and bounds on exponential sums considered in analytic number theory. Our first construction matches a previous randomized construction of Nelson, Nguyen and Woodruff (RANDOM'12), where there was no constraint on the form of the incoherent matrix.

Our algorithms are nearly sample-optimal, since a lower bound of $\Omega(k^2 + k \log n)$ is known, even for the case where the sensing matrix can be arbitrarily designed. A similar lower bound of $\Omega(k^2 \log n / \log k)$ is known for incoherent matrices.

2012 ACM Subject Classification Theory of computation → Approximation algorithms analysis; Theory of computation → Streaming, sublinear and near linear time algorithms

Keywords and phrases Fourier sparse recovery, derandomization, incoherent matrices

Digital Object Identifier 10.4230/LIPIcs.ICALP.2020.77

Category Track A: Algorithms, Complexity and Games

Related Version A full version of the paper is available at `https://arxiv.org/abs/1903.00995`.

Funding *Vasileios Nakos*: This work is part of the project TIPEA that has received funding from the European Research Council (ERC) under the European Unions Horizon 2020 research and innovation programme (grant agreement No. 850979). Part of the work was completed when the author was a Ph.D. student in Harvard University and supported in part by NSF CAREER award CCF-1350670.

Acknowledgements We would like to thank anonymous reviewers for their valuable feedback.

47th International Colloquium on Automata, Languages, and Programming (ICALP 2020).
Editors: Artur Czumaj, Anuj Dawar, and Emanuela Merelli; Article No. 77; pp. 77:1–77:14
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

1 Introduction

Compressed sensing is a subfield of discrete signal processing, based on the principle that a high-dimensional signal can be approximately reconstructed, by exploiting its sparsity, in fewer samples than those demanded by the Shannon-Nyquist theorem. An important subtopic is the Sparse Fourier Transform, where we desire to detect and approximate the largest coordinates of a high-dimensional signal, given a few samples from its Fourier spectrum. Fewer samples play a crucial role, for example, in medical imaging, where reconstructing an image corresponds exactly to reconstructing a signal from its Fourier representation. Thus, the number of Fourier coefficients needed for (approximate) reconstruction is proportional to the radiation dose a patient receives as well as the time the patient needs to remain in the scanner. Furthermore, exploiting the sparsity of the signal has given researchers the hope of defeating the FFT algorithm of Cooley and Tukey, in the special (but of high practical value) case where the signal is approximately sparse. Thus, since FFT serves as an important computational primitive, and has been recognized as one of the 10 most important algorithms of the 20th century [16], every place where it has found application can possibly be benefited from a faster algorithm. The main intuition and hope is that signals arising in practice often exhibit certain structure, such as concentration of energy in a small number of Fourier coefficients.

Since vectors in practice are never exactly sparse, and it is impossible to reconstruct a generic vector $\widehat{x} \in \mathbb{C}^n$ from $o(n)$ samples, researchers resort to approximation. More formally, a sparse recovery scheme consists of a sample set $S \subseteq \{1, \dots, n\}$ and a recovery algorithm $\mathcal{R}$ such that for any given $x \in \mathbb{C}^n$, the scheme approximates $\widehat{x}$ by $\widehat{x}' = \mathcal{R}(x_S)$, where x_S denotes the vector of x restricted to the coordinates in S. The fineness of approximation is measured with respect to the best k-sparse approximation to $\widehat{x}$. The breakthrough work of Candès, Tao and Donoho [12, 20] first showed that $k \log^{O(1)} n$ samples of $x \in \mathbb{C}^n$ suffices to reconstruct a $O(k)$-sparse vector $\widehat{x}'$ which is "close" to the best k-approximation of $\widehat{x}$. More formally, the reconstruction $\widehat{x}'$ satisfies the so-called ℓ_2/ℓ_1 guarantee, i.e.,

$$\|\widehat{x} - \widehat{x}'\|_2 \leq \frac{1}{\sqrt{k}} \|\widehat{x}_{-k}\|_1,$$

where $\widehat{x}_{-k}$ is the tail vector, obtained from restricting $\widehat{x}$ to its smallest $n-k$ coordinates in magnitude. The strength of their algorithms lies in the uniformity, in the sense that the samples at the same coordinates can be used to approximate every $x \in \mathbb{C}^n$. However, the running time is polynomial in the vector length n, giving thus only sample-efficient, but not necessarily time-efficient algorithms. Furthermore, the samples are not obtained via a deterministic procedure, but are chosen at random. Regarding non-uniform randomized algorithms that run in sublinear time, numerous researchers have worked on the problem and obtained a series of algorithms with different recovery guarantees [29, 46, 42, 24, 3, 27, 30, 31, 43, 38, 53, 34, 33, 39, 40, 41, 50]. See Table 1 for a list of common recovery guarantees. The state of the art is the seminal algorithm of Kapralov [40], which shows that $O(k \log n)$ samples and $O(k \log^{O(1)} n)$ time are simultaneously possible for the ℓ_2/ℓ_2 guarantee (which is strictly stronger[1] than the ℓ_2/ℓ_1). The fastest algorithm is due to [30], needing $O(k \log n \cdot \log(n/k))$ time and samples. We note also the algorithm of Indyk and Kapralov [33] that runs in $O(n \log^2 n)$ time, uses $O(k \log n)$ samples but gives a stronger

[1] Here we mean that given an algorithm giving the ℓ_2/ℓ_2 guarantee, one can create an algorithm, using the ℓ_2/ℓ_2 algorithm as a black box, with sparsity parameter $k' = O(k)$, achieving the ℓ_2/ℓ_1 guarantee with the same order of number of samples.

Table 1 Common guarantees of sparse recovery. Only the ℓ_2/ℓ_2 case requires a parameter $C > 1$. The guarantees are listed in the descending order of strength.

Guarantee	Formula	Deterministic Lower Bound
ℓ_∞/ℓ_2	$\|\widehat{x} - \widehat{x}'\|_\infty \le \|\widehat{x}_{-k}\|_2/\sqrt{k}$	$\Omega(n)$ [17]
ℓ_2/ℓ_2	$\|\widehat{x} - \widehat{x}'\|_2 \le C\|\widehat{x}_{-k}\|_2$	$\Omega(n)$ [17]
ℓ_∞/ℓ_1	$\|\widehat{x} - \widehat{x}'\|_\infty \le \|\widehat{x}_{-k}\|_1/k$	$\Omega(k^2 + k\log n)$ [23, 21]
ℓ_2/ℓ_1	$\|\widehat{x} - \widehat{x}'\|_2 \le \|\widehat{x}_{-k}\|_1/\sqrt{k}$	$\Omega(k\log(n/k))$ [23, 21]

ℓ_∞/ℓ_2 guarantee than the ℓ_2/ℓ_2 guarantee in the previous two papers. We refer the reader to the next section for comparison of the different guarantees appearing in the literature. Recently there has been also considerable work on recovering k-sparse signals from their continuous Fourier Transform, see [9, 55, 14, 6].

Although our understanding on randomized algorithms is almost complete, there are still important gaps in our knowledge regarding deterministic schemes. The following natural open-ended question has theoretical and practical interest and remains in principle highly unexplored, touching a variety of fields including (sublinear-time) algorithms, pseudorandomness and computational complexity, Additive Combinatorics [10] and analytic number theory.

► **Question 1.** *What are the best bounds we can obtain for the different versions of the deterministic Sparse Fourier Transform problem?*

With sublinear runtime, the earliest work of Iwen [36, 37] gives $O(k^2 \log^4 n)$ samples and time, albeit in a significantly easier (although similar) model: where one wants to learn a band-limited function $f : [0, 2\pi) \to \mathbb{C}$ and can evaluate f at any point. In the discrete case which we are interested in, the state of the art is the work of Merhi et al. [47], which obtains $O(k^2 \log^{11/2} n/\log k)$ samples and the same runtime. A recent work of Bittens et al. [8] showed that the quadratic dependence can be dropped if the signals are sufficiently structured, namely, if the Fourier coefficients are generated by an unknown but small degree polynomial. On the related problem of the Walsh-Hamadard Transform, Indyk and Cheraghchi [15] showed that roughly $O(k^{1+\alpha} \log^{O(1)+6/\alpha} n)$ samples and similar run-time are possible, if one resorts to a slightly weaker guarantee. Interestingly, their approach resides in a novel connection between the Walsh-Hadamard matrix and linear lossless condensers. However, this connection does not extend to the Fourier Transform over $\mathbb{Z}_n$, which is our focus and the most interesting case. Interesting ideas appear also in the work of Akavia [1, 2], where it is shown how to approximate the Fourier Transform of an arithmetic progression in poly-logarithmic time in the length of the progression; due to the worse dependence on the quality of approximation, however, that work obtained an algorithm with sample complexity $(k \cdot (\text{signal-to-noise ratio}))^4$.

The papers above showed how to achieve the ℓ_2/ℓ_1 guarantee in a number of samples that is quadratic in the signal sparsity. It is already known that a nearly linear dependence is possible [12]; however, we do not have efficient deterministic algorithms for finding these samples. The work of [12], as well as subsequent works, proceeds by sampling with repetition rows of the DFT matrix, and showing that the RIP condition (see Definition 5) holds, which in turn implies the desired result, but via a super-linear algorithm. The state-of-the-art analysis of such row subsampling is due to Haviv and Regev [32], who showed that $O(k \log^2 k \log n)$ samples suffice. A lower bound of $\Omega(k \log n)$ rows for this subsampling process has been shown in [7]. In this paper, we follow a different avenue and give a new set of schemes for the Sparse Fourier Transform which allow uniform reconstruction. Although our dependence is still quadratic in k, it is necessary, in contrast to the previous works: our results satisfy the strictly stronger ℓ_∞/ℓ_1 guarantee, for which a quadratic lower bound is known [23], and hence one cannot hope for a sub-quadratic dependence. We also note the deterministic algorithm of [41], which needs a cubic dependence on k but solves a somewhat different problem of finding the multidimensional sparse Fourier transform of a signal with at most k non-zeros in the frequency domain, and thus is not robust to noise.

The focus of our work is the ℓ_∞/ℓ_1 guarantee, defined formally as follows.

▶ **Definition 2** (ℓ_∞/ℓ_1 guarantee). *A sparse recovery scheme is said to satisfy the ℓ_∞/ℓ_1 guarantee with parameter k, if given access to vector x, it outputs a vector $\widehat{x}'$ such that*

$$\|\widehat{x} - \widehat{x}'\|_\infty \leq \frac{1}{k}\|\widehat{x}_{-k}\|_1. \tag{1}$$

ℓ_∞/ℓ_1 versus ℓ_2/ℓ_1: A matter of "find all" versus "miss all"

As we have discussed, previous works satisfied the ℓ_2/ℓ_1 guarantee, while our target is the ℓ_∞/ℓ_1 guarantee. Any algorithm for the latter guarantee also satisfies the former one. But, as we shall demonstrate in Section 2.3, the ℓ_∞/ℓ_1 guarantee is much stronger: there exists an infinite family of vectors for which an ℓ_2/ℓ_1 algorithm might detect none of the heavy frequencies, while an ℓ_∞/ℓ_1 algorithm must detect all of them. This happens because the ℓ_∞/ℓ_1 is a **worst-case** guarantee, in the sense that it requires detection of every frequency just above the noise level, in contrast to the ℓ_2/ℓ_1, which should be regarded as an **average-case** guarantee in the sense that it allows missing a subset of the heavy frequencies if they carry the energy proportional to the noise level.

Previous Work on ℓ_∞/ℓ_1 with arbitrary linear measurements

All approaches described above concerned Fourier measurements, but compressed sensing has a long history using arbitrary linear measurements, for example [19, 56, 35, 25, 28, 48, 26, 45, 44, 49]. Regarding ℓ_∞/ℓ_1, the work of [51] indicated a connection between the aforementioned guarantee and incoherent matrices. More specifically, it was shown that given a $(1/k)$-incoherent matrix one can design an algorithm satisfying the ℓ_∞/ℓ_1 guarantee. The existence of a matrix with $O(k^2 \min\{\log n, (\log n/\log k)^2\})$ rows was also proved. Reconstruction needed $\Omega(nk)$ time, something which was partially remedied by Li and Nakos [44] with a scheme of $O(k^2 \log n \cdot \log^* k)$ measurements and $\mathrm{poly}(k, \log n)$ decoding time. Incoherent matrices are interesting objects on their own, and have been studied before, as they can be used to obtain RIP matrices. Deterministic constructions of $O(k^2(\log n/\log k)^2)$ rows were obtained by DeVore [18] using deep results from the theory of Gelfand widths and by Amini and Marvasti [5] via binary BCH code vectors, where the zeros are replaced by -1s. We note that incoherent matrices matching this bound also follow immediately from the famous Nisan-Wigderson combinatorial designs [52], and serve as a cornerstone for constructions of pseudorandom generators and extractors [59]. Incoherent matrices are also connected with ϵ-biased codes, and thus an almost optimal strongly explicit construction can be obtained by the recent breakthrough work of [57]. On the lower bound side, Alon has shown that $\Omega(k^2 \log n/\log k)$ rows are necessary for a $(1/k)$-incoherent matrix [4].

Our Contribution

In this work we offer several new results for the Sparse Fourier Transform problem across different axis, some of which are nearly optimal. We show how to find in polynomial time a deterministic collection of samples from the time domain, such that we can solve the Sparse Fourier Transform problem in linear and sublinear time and achieve nearly optimal sample complexity. For the closely related problem of incoherent matrices from DFT rows, which is of independent interest, we obtain a nearly optimal derandomized construction via Bernstein's inequality. We also demonstrate strongly explicit constructions, by invoking heavy number-theoretical machinery.

We note that the bounds of our constructions have been known for more than a decade if the sensing/incoherent matrix is allowed to be arbitrary. However, the previous arguments did not facilitate the frequent and relevant scenario where we have access to rows only from the Fourier ensemble. Part of our work is to show that some of these results carry over to the significantly more constrained case. We also note that any progress to deterministic ℓ_2/ℓ_1 schemes with subquadratic sample complexity is connected to the very challenging problem of obtaining a deterministic DFT row-subsampled RIP matrices with subquadratic number of rows[2] which possibly out of reach at the moment.

2 Technical Results

2.1 Preliminaries

For a positive integer n, we define $[n] = \{0, 1 \ldots, n-1\}$ and we shall index the coordinates of a n-dimensional vector or the rows/columns of an $n \times n$ matrix from 0 to $n-1$. We define the Discrete Fourier Transform (DFT) matrix $F \in \mathbb{C}^{n\times n}$ to be the unitary matrix such that $F_{ij} = \frac{1}{\sqrt{n}} e^{2\pi\sqrt{-1}\cdot ij/n}$, and the Discrete Fourier Transform of a vector $x \in \mathbb{C}^n$ to be $\widehat{x} = Fx$.

For a set $S \subseteq [n]$ we define x_S to be the vector obtained from x after zeroing out the coordinates not in S. We also define $H(x, k)$ to be the set of the indices of the largest k coordinates (in magnitude) of x, and $x_{-k} = x_{[n]\setminus H(x,k)}$. We say x is k-sparse if $x_{-k} = 0$. We also define $\|x\|_p = \left(\sum_{i=0}^{n-1} |x_i|^p\right)^{1/p}$ for $p \geq 1$ and $\|x\|_0$ to be the number of nonzero coordinates of x.

For a matrix $F \in \mathbb{C}^{n\times n}$ and subsets $S, T \subseteq [n]$, we define $F_{S,T}$ to be the submatrix of F indexed by rows in S and columns in T.

The median of a collection of complex numbers $\{z_i\}$ is defined to be $\operatorname{median}_i z_i = \operatorname{median}_i \operatorname{Re}(z_i) + \sqrt{-1}\operatorname{median}_i \operatorname{Im}(z_i)$, i.e., taking the median of the real and the imaginary component separately.

For two points x and y on the unit circle, we use $|x - y|_\circ$ to denote the circular distance (in radians, i.e. modulo 2π) between x and y.

2.1.1 ℓ_∞/ℓ_1 Gurantee and incoherent matrices

The quality of the approximation is usually measured in different error metrics, and the main recovery guarantee we are interested in is called the ℓ_∞/ℓ_1 guarantee, as defined in Definition 2. Other types of recovery guarantee, such as the ℓ_∞/ℓ_2, the ℓ_2/ℓ_2 and the

[2] Note that [10] breaks the quadratic barrier for RIP matrices but does not use the Fourier ensemble; the rows are picked from the *discrete chirp-Fourier* ensemble, where the linear functions are substituted by quadratic polynomials.

ℓ_2/ℓ_1, are defined similarly, where (1) is replaced with the respective expression in Table 1. Note that these are definitions of the error guarantee per se and do not have algorithmic requirements on the scheme.

Highly relevant with the ℓ_∞/ℓ_1 guarantee is a matrix condition which we call incoherence.

▶ **Definition 3** (Incoherent Matrix). *A matrix $A \in \mathbb{C}^{m\times n}$ is called ϵ-incoherent if $\|A_i\|_2 = 1$ for all i (where A_i denotes the i-th column of A) and $|\langle A_i, A_j\rangle| \leq \epsilon$.*

▶ **Lemma 4** ([51]). *There exist an absolute constant $c > 0$ such that for any (c/k)-incoherent matrix A, there exists a ℓ_∞/ℓ_1-scheme which uses A as the measurement matrix and whose recovery algorithm runs in polynomial time.*

2.1.2 The Restrictred Isometry Property and its connection with incoherence

Another highly relevant condition is called the renowned restricted isometry property, introduced by Candès et al. in [11]. We show how incoherent matrices are connected to it.

▶ **Definition 5** (Restricted Isometry Property). *A matrix $A \in \mathbb{C}^{m\times n}$ is said to satisfy the (k, ϵ) Restricted Isometry Property (RIP), if for all $x \in \mathbb{C}^n$ with $\|x\|_0 \leq k$, it holds that $(1-\epsilon)\|x\|_2 \leq \|Ax\|_2 \leq (1+\epsilon)\|x\|_2$.*

Candès et al. proved in their breakthrough paper [11] that any RIP matrix can be used for sparse recovery with the ℓ_2/ℓ_1 error guarantee. The following formulation comes from [22, Theorem 6.12].

▶ **Lemma 6.** *Given a $(2k, \epsilon)$-RIP matrix A with $\epsilon < 4/\sqrt{41}$, we can design a ℓ_2/ℓ_1-scheme that uses A as the measurement matrix and has a recovery algorithm that runs in polynomial time.*

Although randomly subsampling the DFT matrix gives an RIP matrix with $O(k\log^2 k \log n)$ rows [32], no algorithm for finding these rows in polynomial time is known; actually, even for $o(k^2)\cdot \mathrm{poly}(\log n)$ rows the problem remains wide open[3]. It is a very important and challenging problem whether one can have an explicit construction of RIP matrices from Fourier measurements that break the quadratic barrier on k.

We state the following two folklore results, connecting the two different guarantees, and their associated combinatorial objects. This indicates the importance of incoherent matrices for the field of compressed sensing.

▶ **Proposition 7** (folklore). *An ℓ_∞/ℓ_1 scheme with a measurement matrix of m rows and recovery time T induces an ℓ_2/ℓ_1 scheme of a measurement matrix of $O(m)$ rows and recovery time $O(T + \|\widehat{x}'\|_0)$, where $\widehat{x}'$ is the output of the ℓ_∞/ℓ_1 scheme.*

▶ **Proposition 8** (folklore). *A (c/k)-incoherent matrix is also a (k, c)-RIP matrix.*

[3] In fact, one of the results of our paper gives the state-of-the-art result even for this problem, with $O(k^2 \log n)$ rows, see Theorem 12.

Table 2 Comparison of our results and the previous results. All O- and Ω-notations are suppressed. The result in the first row follows from Lemma 6 and the RIP matrix in [32].Our algorithms adopt the common assumption in the sparse FT literature that the signal-to-noise ratio is bounded by n^c for some absolute constant $c > 0$.

	Samples	Run-time	Guarantee	Explict Construction	Lower Bound
[32]	$k \log^2 k \log n$	$\text{poly}(n)$	ℓ_2/ℓ_1	No	$k \log(n/k)$
[47]	$k^2 \log^{5.5} n / \log k$	$k^2 \log^{5.5} n / \log k$	ℓ_2/ℓ_1	Yes	$k \log(n/k)$
Theorem 9	$k^2 \log n$	$nk \log^2 n$	ℓ_∞/ℓ_1	Yes	$k^2 + k \log n$[51]
Theorem 10	$k^2 \log^2 n$	$k^2 \log^3 n$	ℓ_∞/ℓ_1	Yes	$k^2 + k \log n$[51]

2.2 Our results

2.2.1 Sparse Fourier Transform Algorithms

▶ **Theorem 9** (Deterministic SFT with super-linear time)**.** *Let n be a power of 2. There exist a set $S \subseteq [n]$ with $|S| = O(k^2 \log n)$ and an absolute constant $c > 0$ such that the following holds. For any vector $x \in \mathbb{C}^n$ with $\|\widehat{x}\|_\infty \leq n^c \|\widehat{x}_{-k}\|_1/k$, one can find an $O(k)$-sparse vector $\widehat{x}' \in \mathbb{C}^n$ such that*

$$\|\widehat{x} - \widehat{x}'\|_\infty \leq \frac{1}{k}\|\widehat{x}_{-k}\|_1,$$

in time $O(nk \log^2 n)$ by accessing $\{x_i\}_{i\in S}$ only. Moreover, the set S can be found in $\text{poly}(n)$ *time.*

▶ **Theorem 10** (Deterministic SFT with sublinear time,)**.** *Let n be a power of 2. There exist a set $S \subseteq [n]$ with $|S| = O(k^2 \log^2 n)$ and an absolute constant $c > 0$ such that the following holds. For any vector $x \in \mathbb{C}^n$ with $\|\widehat{x}\|_\infty \leq n^c \|\widehat{x}_{-k}\|_1/k$, one can find an $O(k)$-sparse vector $\widehat{x}' \in \mathbb{C}^n$ such that*

$$\|\widehat{x} - \widehat{x}'\|_\infty \leq \frac{1}{k}\|\widehat{x}_{-k}\|_1,$$

in time $O(k^2 \log^3 n)$ by accessing $\{x_i\}_{i\in S}$ only. Moreover, the set S can be found in $\text{poly}(n)$ *time.*

▶ Remark 11. The condition $\|\widehat{x}\|_\infty \leq n^c \|\widehat{x}_{-k}\|_1/k$ upper bounds the "signal-to-noise ratio", a common measure in engineering that compares the level of a desired signal to the level of the background noise. This is a common assumption in most algorithms in the Sparse Fourier Transform literature, see, e.g. [30, 33, 39, 13, 40], where the ℓ_2-norm variant $\|\widehat{x}\|_\infty \leq n^c \|\widehat{x}_{-k}\|_2/\sqrt{k}$ was assumed.

2.2.2 From DFT to incoherent matrices

This section contains deterministic constructions of incoherent matrices.

An Explicit Construction: Derandomization in $\text{poly}(n)$ time

▶ **Theorem 12** (Incoherent matrices by derandomized subsampling of DFT)**.** *There exists a set $S \subseteq [n]$ with of cardinality $O(k^2 \log n)$ such that the matrix $\sqrt{\frac{n}{m}} F_{S,[n]}$ is $(1/k)$-incoherent. Moreover, S can be found in* $\text{poly}(n)$ *time.*

The above Theorem yields immediately a different algorithm for ℓ_∞/ℓ_1 Sparse Fourier Tranform with $O(k^2 \log n)$ samples, via the reduction in [51].

Strongly explicit constructions: Derandomization in sub-linear time

▶ **Theorem 13** (Incoherent matrices from DFT via low-degree polynomials). *Let $\epsilon > 0$ be a constant small enough, p be a prime and $d \geq 2$ be an integer. There exists a strongly explicit construction of an $O(m^\epsilon(\frac{1}{m} + \frac{p}{m^d})^{2^{1-d}})$-incoherent matrix $M \in \mathbb{C}^{m\times p}$ such that the rows of $\sqrt{m}M$ are rows of the DFT matrix (a row may appear more than once). The hidden constant in the O-notation depends on d and ϵ. Finding the indices of the rows takes $\widetilde{O}(m)$ time.*

To get an idea of the above result one could for example set $d = 3$ and observe that the results translates to the following: for every $k \geq p^{1/8}$ one can get a $(1/k)$-incoherent matrix with $O(k^{4+\epsilon})$ rows. One needs the condition on k (or equivalently the condition on m) to bound the term p/m^d. The larger the degree d, the looser this condition, but also the worse the dependence of m on k. For example, when $d = 4$, we can expand the regime of k to approximately $k \geq p^{1/24}$, but obtain approximately $m = O(k^{8+\epsilon})$.

The following is a different construction, incomparable with Theorem 13 in multiple ways. First, the construction runs in sublinear time in p but it is not strongly explicit. Second, it gives different trade-offs between the sparsity parameter and the number of rows. Last but not least, the construction depends on the factorization of $p - 1$.

▶ **Theorem 14** (Incoherent matrices from DFT via multiplicative subgroupss). *Let p be a prime number. For every divisor d of $p-1$ such that $d > \sqrt{p}$ we can find in time $O(d \log p)$ a matrix $M \in \mathbb{C}^{d\times p}$ with rows being the rows of the DFT matrix such that $\frac{1}{d}M$ is $(\sqrt{p}/d)$-incoherent.*

This result could give (depending on the factorization of $p - 1$) a better polynomial dependence of m on k in the high-sparsity regime. If $p - 1$ has a large divisor about $p^{1-\gamma}$, this would yield a matrix with sparsity parameter $k \approx p^\gamma$ and $m \approx k^{1/\gamma-1}$ rows. For example, when $\gamma = 1/4$, we obtain $k \approx p^{1/4}$ and $m \approx k^3$, which cannot be obtained from Theorem 13. In general, Theorem 14 will yield useful matrices as long as $p - 1$ has divisors in the range $[\sqrt{p}, p - 1]$, ideally as many as possible. An extreme case is Fermat primes, which have $(\log p)/2$ divisors in the aforesaid interval.

The reader might ask the question if the polynomial dependence of k on p is necessary; ideally one would like a logarithmic dependence, since the polynomial dependence is interesting only in the high-sparsity regime. Regarding strongly explicit constructions, we provide some evidence why this might be a very hard problem in the remark below.

▶ Remark 15. The inferiority of our bounds in the low-sparsity regime is justifiable to some extent: it is because of a common obstacle that has persisted more than a century in the theory of exponential sums, due to the lack of techniques to account for sparse character sums (either additive or multiplicative). In general, the fewer summands the sum has, the harder it is to prove a tight cancellation bound. Thus, owing to the use of heavy machinery from analytic number theory and more specifically the theory of exponential sums over finite fields, our bounds for strongly explicit constructions are quite suboptimal.

2.3 Comparing ℓ_2/ℓ_1 with ℓ_∞/ℓ_1

In this subsection we elaborate why ℓ_∞/ℓ_1 is much stronger than ℓ_2/ℓ_1, and not just a guarantee that implies ℓ_2/ℓ_1. Let $\gamma < 1$ be a constant and consider the following scenario. There are three sets A, B, C of size $\gamma k, (1-\gamma)k$, $n - k$ respectively, and for every $i \in A$ we have $|\widehat{x}_i| = \frac{2}{k}\|\widehat{x}_C\|_1 = \frac{2}{k}\|\widehat{x}_{-k}\|_1$, while every coordinate in B and C has the equal magnitude. It follows immediately that

$$\|\widehat{x}_C\|_1 = \frac{n-k}{n-\gamma k}\|\widehat{x}_{B\cup C}\|_1.$$

Now assume that $k \leq \gamma n$, then $(n - \gamma k)/(n - k) \leq 1 + \gamma$. We claim that the zero vector is a valid solution for the ℓ_2/ℓ_1 guarantee, since

$$\begin{aligned}
\|\vec{0} - \widehat{x}\|_2^2 &= \|\widehat{x}_A\|_2^2 + \|\widehat{x}_{B\cup C}\|_2^2 \\
&\leq \gamma k \cdot \frac{4}{k^2}\|\widehat{x}_{-k}\|_1^2 + \frac{1}{(n-\gamma k)}\|\widehat{x}_{B\cup C}\|_1^2 \\
&\leq \frac{4\gamma}{k}\|\widehat{x}_{-k}\|_1^2 + \frac{n-\gamma k}{(n-k)^2}\|\widehat{x}_C\|_1^2 \\
&\leq \left(\frac{4\gamma}{k} + \frac{1+\gamma}{n-k}\right)\|\widehat{x}_{-k}\|_1^2 \\
&\leq \frac{5\gamma}{k}\|\widehat{x}_{-k}\|_1^2,
\end{aligned}$$

where the last inequality follows provided it further holds that $k \leq \gamma n/(2\gamma + 1)$. Hence when $\gamma \leq 1/5$, we see that the zero vector satisfies the ℓ_2/ℓ_1 guarantee.

Since $\vec{0}$ is a possible output, we may not recover any of the coordinates in S, which is the set of "interesting" coordinates. On the other hand, the ℓ_∞/ℓ_1 guarantee does allow the recovery of **every** coordinate in S. This is a difference of recovering all γk versus 0 coordinates. We conclude from the discussion above that in the case of too much noise, the ℓ_2/ℓ_1 guarantee becomes much weaker than the ℓ_∞/ℓ_1, possibly giving meaningless results in some cases.

3 Overview

Sparse Fourier Transform Algorithms

We first show how to achieve the for-all schemes, i.e., schemes that allow universal reconstruction of all vectors, and then derandomize them. Similarly to the previous works [31, 33, 40], our algorithm hashes, with the filter in [40], the spectrum of x to $O(k)$ buckets using pseudorandom permutations, and repeat $O(k \log n)$ times with fresh randomness. The main part of the analysis is to show that for any vector $\widehat{x} \in \mathbb{C}^n$ and any set $S \subseteq [n]$ with $|S| \leq k$, each $i \in S$, in a constant fraction of the repetitions, receives "low noise" from all other elements, under the pseudorandom permutations. This will boil down to a set of $\Theta(n^2)$ inequalities involving the filter and the pseudorandom permutations. We prove these inequalities with full randomness, and then derandomize the pseudorandom permutations using the method of conditional expectations. This will give us Theorem 9. To do so, we choose the pseudorandom permutations one at a time, repetition by repetition, and keep an (intricate) pessimistic estimator , which we update accordingly. Our argument extends the arguments in [51] and [54], and could be of independent interest. To compare with [51] we have the following observation. The construction in [51] consists of $O(k \log n)$ matrices, joined vertically, each having $O(k)$ rows and exactly one 1 per column. This ensures a small incoherence of the concatenated matrix and gives the ℓ_∞/ℓ_1 guarantee. In the Fourier case, the convolution with the filter functions behaves analogously: instead of having exactly one non-zero element, each column in the ℓ-th matrix has a contiguous segment of 1s of size $\approx n/k$ (where the center of that segment is depends on the choice of the ℓ-th pseudorandom permutation) and polynomially decaying entries away from this segment. Moreover, the positions of the segments across the columns are not fully independent and are defined via the pseudorandom permutations. We show that even in this more restricted setting, derandomization is possible in polynomial time. Several details are omitted in the preceding high-level discussion and we suggest the reader look at the corresponding sections for the complete argument.

The sublinear-time algorithm (Theorem 10) is obtained by bootstrapping the derandomized scheme above with an identification procedure in each bucket, as most previous algorithms have done (e.g. [30]). The major difference is that our identification procedure needs to be deterministic. We show an explicit set of samples that allow the implementation of the desired routine. To illustrate our idea, let us focus on the following 1-sparse case: $\widehat{x} \in \mathbb{C}^n$ and $|\widehat{x}_{i^*}| \geq 3\|\widehat{x}_{[n]\setminus i^*}\|_1$ for some i^*, which we want to locate. Let

$$\theta_j = \left(\frac{2\pi}{n} j\right) \bmod 2\pi,$$

and consider the $\log n$ samples $x_0, x_1, x_2, x_4, \ldots, x_{2^{r-1}}, \ldots$.

Observe that (ignoring $1/\sqrt{n}$ factors)

$$x_\beta = \widehat{x}_{i^*} e^{\sqrt{-1}\beta\theta_{i^*}} + \sum_{j \neq i^*} \widehat{x}_j e^{\sqrt{-1}\beta\theta_j},$$

we can find $\beta\theta_{i^*} + \arg \widehat{x}_{i^*}$ up to $\pi/8$, just by estimating the phase of x_β. Thus we can estimate $\beta\theta_{i^*}$ up to $\pi/4$ from the phase of x_β/x_0. If $i^* \neq j$, then there exists a $\beta \in \{1, 2, 2^2, \ldots, 2^{r-1}, \ldots\}$ such that $|\beta\theta_{i^*} - \beta\theta_j|_\circ > \pi/2$, and so $\beta\theta_j$ will be more than $\pi/4$ away from the phase of the measurement. Thus, by iterating over all $j \in [n]$, we keep the index j for which $\beta\theta_j$ is within $\pi/4$ from $\arg(x_\beta/x_0)$, for every β that is a power of 2 in $\mathbb{Z}_n$.

Unfortunately, although this is a deterministic collection of $O(\log n)$ samples, the above argument gives only $O(n \log n)$ time. For sublinear-time decoding we use x_1/x_0 to find a sector S_0 of the unit circle of length $\pi/4$ that contains θ_{i^*}. Then, from x_2/x_0 we find two sectors of length $\pi/8$ each, the union of which contains θ_{i^*}. Because these sectors are antipodal on the unit circle, the sector S_0 intersects exactly one of those, let the intersection be S_1. The intersection is a sector of length at most $\pi/8$. Proceeding iteratively, we halve the size of the sector at each step, till we find θ_{i^*}, and infer i^*. Plugging this idea in the whole k-sparse recovery scheme yields the desired result. Our argument crucially depends on the fact that in the ℓ_1 norm the phase of θ_{i^*} will always dominate the phase of all samples we take.

Incoherent Matrices from the Fourier ensemble

Our first result for incoherent matrices (Theorem 12) is more general and works for any matrix that has orthonormal columns with entries bounded by $O(1/\sqrt{n})$. We subsample the matrix, invoke a Chernoff bound and Bernstein's inequality to show the small incoherence of the subsampled matrix. We follow a derandomization procedure which essentially mimics the proof of Bernstein's inequality, keeping a pessimistic estimator which corresponds to the sum of the generating functions of the probabilities of all events we want to hold, evaluated at specific points. We obtain an explicit construction, i.e. a derandomization in poly(n) time. This argument could be of independent interest for its generality. As there are many technical obstacles to overcome, we suggest the reader take a careful look at the proof to gain a clearer picture of the argument.

Our next results (Theorem 13 and Theorem 14) construct *strongly explicit* incoherent matrices by making use of technology from the fruitful theory of exponential sums in analytic number theory and additive combinatorics. Roughly speaking, to bound a complex exponential sum over a set S, one would expect that specific choices of the set S lead to non-trivial bounds, i.e. $o(|S|)$, since cancellation takes place in the summation. Ideally, one would desire that the exponentials behave like a random walk and give the optimal cancellation of $O(\sqrt{|S|})$. This intuition is clearly not true, but the results by Weyl and others show that certain sets S can exhibit a nicer behaviour. We exploit their results to build incoherent matrices by taking the rows of the DFT matrix indexed by the "nice" sets. This connection also yields an immediate improvement on the lower bound of an exponential sum obtained by Winterhof [60].

4 Open Problems and Future Direction

A direction of research is to design deterministic schemes that break the quadratic barrier for signals with structured Fourier support. For example, subsampling the rows of the DFT matrix to obtain RIP matrices depends highly on the structure of the vectors we would like to preserve. The more additive structure the support of a k-sparse vector x has, the worse is the concentration of a random Fourier coefficient of x. Equivalently, the less additive structure the support of x has, the flatter its Fourier transform is, and hence, the better concentration bounds we obtain. The concentration in the extreme case, when the support of x is "dissociated", is captured by the renowned Rudin's inequality in additive combinatorics (see, e.g. [58, Lemma 4.33]). We thus believe that it is an interesting direction to use machinery from the field of additive combinatorics and the relevant fields in order to obtain new constructions and algorithms, at least for interesting subclasses of structured signals.

References

1 Adi Akavia. Deterministic sparse Fourier approximation via fooling arithmetic progressions. In *COLT*, pages 381–393, 2010.

2 Adi Akavia. Deterministic sparse Fourier approximation via approximating arithmetic progressions. *IEEE Transactions on Information Theory*, 60(3):1733–1741, 2014.

3 Adi Akavia, Shafi Goldwasser, and Shmuel Safra. Proving hard-core predicates using list decoding. In *FOCS*, volume 44, pages 146–159, 2003.

4 Noga Alon. Perturbed identity matrices have high rank: Proof and applications. *Combinatorics, Probability and Computing*, 18(1-2):3–15, 2009.

5 Arash Amini and Farokh Marvasti. Deterministic construction of binary, bipolar, and ternary compressed sensing matrices. *IEEE Transactions on Information Theory*, 57(4):2360–2370, 2011.

6 Haim Avron, Michael Kapralov, Cameron Musco, Christopher Musco, Ameya Velingker, and Amir Zandieh. A universal sampling method for reconstructing signals with simple Fourier transforms. *arXiv preprint*, 2018. `arXiv:1812.08723`.

7 Afonso S Bandeira, Megan E Lewis, and Dustin G Mixon. Discrete uncertainty principles and sparse signal processing. *Journal of Fourier Analysis and Applications*, pages 1–22, 2017.

8 Sina Bittens, Ruochuan Zhang, and Mark A Iwen. A deterministic sparse FFT for functions with structured Fourier sparsity. *Advances in Computational Mathematics, to appear.*, 2017.

9 Petros Boufounos, Volkan Cevher, Anna C Gilbert, Yi Li, and Martin J Strauss. What's the frequency, Kenneth?: Sublinear Fourier sampling off the grid. In *Algorithmica(A preliminary version of this paper appeared in the Proceedings of RANDOM/APPROX 2012, LNCS 7408, pp.61–72)*, pages 1–28. Springer, 2014.

10 Jean Bourgain, Stephen J Dilworth, Kevin Ford, Sergei V Konyagin, and Denka Kutzarova. Breaking the k^2 barrier for explicit RIP matrices. In *Proceedings of the forty-third annual ACM symposium on Theory of computing*, pages 637–644. ACM, 2011.

11 Emmanuel J Candes, Justin K Romberg, and Terence Tao. Stable signal recovery from incomplete and inaccurate measurements. *Communications on pure and applied mathematics*, 59(8):1207–1223, 2006.

12 Emmanuel J Candes and Terence Tao. Near-optimal signal recovery from random projections: Universal encoding strategies? *IEEE transactions on information theory*, 52(12):5406–5425, 2006.

13 Volkan Cevher, Michael Kapralov, Jonathan Scarlett, and Amir Zandieh. An adaptive sublinear-time block sparse Fourier transform. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, pages 702–715. ACM, 2017.

14 Xue Chen, Daniel M Kane, Eric Price, and Zhao Song. Fourier-sparse interpolation without a frequency gap. In *Foundations of Computer Science (FOCS), 2016 IEEE 57th Annual Symposium on*, pages 741–750. IEEE, 2016.

15 Mahdi Cheraghchi and Piotr Indyk. Nearly optimal deterministic algorithm for sparse walsh-hadamard transform. *ACM Transactions on Algorithms (TALG)*, 13(3):34, 2017.

16 Barry A Cipra. The best of the 20th century: Editors name top 10 algorithms. *SIAM news*, 33(4):1–2, 2000.

17 Albert Cohen, Wolfgang Dahmen, and Ronald DeVore. Compressed sensing and best k-term approximation. *Journal of the American mathematical society*, 22(1):211–231, 2009.

18 Ronald A DeVore. Deterministic constructions of compressed sensing matrices. *Journal of complexity*, 23(4):918–925, 2007.

19 Khanh Do Ba, Piotr Indyk, Eric Price, and David P Woodruff. Lower bounds for sparse recovery. In *Proceedings of the twenty-first annual ACM-SIAM symposium on Discrete Algorithms*, pages 1190–1197. SIAM, 2010.

20 David L. Donoho. Compressed sensing. *IEEE Trans. Information Theory*, 52(4):1289–1306, 2006.

21 Simon Foucart, Alain Pajor, Holger Rauhut, and Tino Ullrich. The gelfand widths of ℓ_p-balls for $0 < p \leq 1$. *Journal of Complexity*, 26(6):629–640, 2010.

22 Simon Foucart and Holger Rauhut. *A Mathematical Introduction to Compressive Sensing*. Applied and Numerical Harmonic Analysis. Birkhäuser Basel, 2013.

23 Sumit Ganguly. Lower bounds on frequency estimation of data streams. In *International Computer Science Symposium in Russia*, pages 204–215. Springer, 2008.

24 Anna C Gilbert, Sudipto Guha, Piotr Indyk, S Muthukrishnan, and Martin Strauss. Near-optimal sparse Fourier representations via sampling. In *Proceedings of the thiry-fourth annual ACM symposium on Theory of computing*, pages 152–161. ACM, 2002.

25 Anna C Gilbert, Yi Li, Ely Porat, and Martin J Strauss. Approximate sparse recovery: optimizing time and measurements. *SIAM Journal on Computing 2012 (A preliminary version of this paper appears in STOC 2010)*, 41(2):436–453, 2010.

26 Anna C Gilbert, Yi Li, Ely Porat, and Martin J Strauss. For-all sparse recovery in near-optimal time. *ACM Transactions on Algorithms (TALG)*, 13(3):32, 2017.

27 Anna C Gilbert, S Muthukrishnan, and Martin Strauss. Improved time bounds for near-optimal sparse Fourier representations. In *Optics & Photonics 2005*, pages 59141A–59141A. International Society for Optics and Photonics, 2005.

28 Anna C Gilbert, Hung Q Ngo, Ely Porat, Atri Rudra, and Martin J Strauss. ℓ_2/ℓ_2-foreach sparse recovery with low risk. In *International Colloquium on Automata, Languages, and Programming*, pages 461–472. Springer, 2013.

29 Oded Goldreich and Leonid A Levin. A hard-core predicate for all one-way functions. In *Proceedings of the twenty-first annual ACM symposium on Theory of computing*, pages 25–32. ACM, 1989.

30 Haitham Hassanieh, Piotr Indyk, Dina Katabi, and Eric Price. Nearly optimal sparse Fourier transform. In *Proceedings of the forty-fourth annual ACM symposium on Theory of computing*, pages 563–578. ACM, 2012.

31 Haitham Hassanieh, Piotr Indyk, Dina Katabi, and Eric Price. Simple and practical algorithm for sparse Fourier transform. In *Proceedings of the twenty-third annual ACM-SIAM symposium on Discrete Algorithms*, pages 1183–1194. SIAM, 2012.

32 Ishay Haviv and Oded Regev. The restricted isometry property of subsampled Fourier matrices. In *SODA*, pages 288–297. arXiv preprint, 2016. `arXiv:1507.01768`.

33 Piotr Indyk and Michael Kapralov. Sample-optimal Fourier sampling in any constant dimension. In *Foundations of Computer Science (FOCS), 2014 IEEE 55th Annual Symposium on*, pages 514–523. IEEE, 2014.

34 Piotr Indyk, Michael Kapralov, and Eric Price. (Nearly) Sample-optimal sparse Fourier transform. In *Proceedings of the Twenty-Fifth Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 480–499. SIAM, 2014.

35 Piotr Indyk, Eric Price, and David P Woodruff. On the power of adaptivity in sparse recovery. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 285–294. IEEE, 2011.

36 Mark A Iwen. A deterministic sub-linear time sparse Fourier algorithm via non-adaptive compressed sensing methods. In *Proceedings of the nineteenth annual ACM-SIAM symposium on Discrete algorithms*, pages 20–29. Society for Industrial and Applied Mathematics, 2008.

37 Mark A Iwen. Combinatorial sublinear-time Fourier algorithms. *Foundations of Computational Mathematics*, 10(3):303–338, 2010.

38 Mark A Iwen. Improved approximation guarantees for sublinear-time Fourier algorithms. *Applied And Computational Harmonic Analysis*, 34(1):57–82, 2013.

39 Michael Kapralov. Sparse Fourier transform in any constant dimension with nearly-optimal sample complexity in sublinear time. In *Symposium on Theory of Computing Conference, STOC'16, Cambridge, MA, USA, June 19-21, 2016*, 2016.

40 Michael Kapralov. Sample efficient estimation and recovery in sparse FFT via isolation on average. In *Foundations of Computer Science (FOCS), 2017 IEEE 58th Annual Symposium on*, pages 651–662. Ieee, 2017.

41 Michael Kapralov, Ameya Velingker, and Amir Zandieh. Dimension-independent sparse Fourier transform. In *Proceedings of the Thirtieth Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, pages 2709–2728. SIAM, 2019.

42 Eyal Kushilevitz and Yishay Mansour. Learning decision trees using the Fourier spectrum. *SIAM Journal on Computing*, 22(6):1331–1348, 1993.

43 David Lawlor, Yang Wang, and Andrew Christlieb. Adaptive sub-linear time Fourier algorithms. *Advances in Adaptive Data Analysis*, 5(01):1350003, 2013.

44 Yi Li and Vasileios Nakos. Deterministic heavy hitters with sublinear query time. In *APPROX-RANDOM*, volume 116 of *LIPIcs*, pages 18:1–18:18. Schloss Dagstuhl - Leibniz-Zentrum fuer Informatik, 2018.

45 Yi Li, Vasileios Nakos, and David P. Woodruff. On low-risk heavy hitters and sparse recovery schemes. In *Approximation, Randomization, and Combinatorial Optimization. Algorithms and Techniques, APPROX/RANDOM 2018, August 20-22, 2018 - Princeton, NJ, USA*, pages 19:1–19:13, 2018. `doi:10.4230/LIPIcs.APPROX-RANDOM.2018.19`.

46 Yishay Mansour. Randomized interpolation and approximation of sparse polynomials. In *International Colloquium on Automata, Languages, and Programming*, pages 261–272. Springer, 1992.

47 Sami Merhi, Ruochuan Zhang, Mark A Iwen, and Andrew Christlieb. A new class of fully discrete sparse Fourier transforms: Faster stable implementations with guarantees. *Journal of Fourier Analysis and Applications*, 2018.

48 Vasileios Nakos, Xiaofei Shi, David P. Woodruff, and Hongyang Zhang. Improved algorithms for adaptive compressed sensing. In *45th International Colloquium on Automata, Languages, and Programming, ICALP 2018, July 9-13, 2018, Prague, Czech Republic*, pages 90:1–90:14, 2018. `doi:10.4230/LIPIcs.ICALP.2018.90`.

49 Vasileios Nakos and Zhao Song. Stronger l_2/l_2 compressed sensing; without iterating. In *Proceedings of the 51st Annual ACM SIGACT Symposium on Theory of Computing, STOC 2019, Phoenix, AZ, USA, June 23-26, 2019*, pages 289–297, 2019. `doi:10.1145/3313276.3316355`.

50 Vasileios Nakos, Zhao Song, and Zhengyu Wang. (nearly) sample-optimal sparse fourier transform in any dimension; ripless and filterless. In *2019 IEEE 60th Annual Symposium on Foundations of Computer Science (FOCS)*, pages 1568–1577. IEEE, 2019.

51 Jelani Nelson, Huy L Nguyên, and David P Woodruff. On deterministic sketching and streaming for sparse recovery and norm estimation. *Linear Algebra and its Applications*, 441:152–167, 2014.

52 Noam Nisan and Avi Wigderson. Hardness vs randomness. *Journal of computer and System Sciences*, 49(2):149–167, 1994.

53 Sameer Pawar and Kannan Ramchandran. A robust R-FFAST framework for computing a k-sparse n-length DFT in O(k log n) sample complexity using sparse-graph codes. In *Information Theory (ISIT), 2014 IEEE International Symposium on*, pages 1852–1856. IEEE, 2014.

54 Ely Porat and Amir Rothschild. Explicit non-adaptive combinatorial group testing schemes. In Luca Aceto, Ivan Damgård, Leslie Ann Goldberg, Magnús M. Halldórsson, Anna Ingólfsdóttir, and Igor Walukiewicz, editors, *Automata, Languages and Programming*, pages 748–759, Berlin, Heidelberg, 2008. Springer Berlin Heidelberg.

55 Eric Price and Zhao Song. A robust sparse Fourier transform in the continuous setting. In *Foundations of Computer Science (FOCS), 2015 IEEE 56th Annual Symposium on*, pages 583–600. IEEE, 2015.

56 Eric Price and David P Woodruff. (1+ eps)-approximate sparse recovery. In *Foundations of Computer Science (FOCS), 2011 IEEE 52nd Annual Symposium on*, pages 295–304. IEEE, 2011.

57 Amnon Ta-Shma. Explicit, almost optimal, epsilon-balanced codes. In *Proceedings of the 49th Annual ACM SIGACT Symposium on Theory of Computing*, STOC 2017, pages 238–251, New York, NY, USA, 2017. ACM.

58 Terence Tao and Van H Vu. *Additive combinatorics*, volume 105. Cambridge University Press, 2006.

59 Luca Trevisan. Extractors and pseudorandom generators. *Journal of the ACM*, 48(4):860–879, 2001.

60 Arne Winterhof. Incomplete additive character sums and applications. In *Finite fields and applications*, pages 462–474. Springer, 2001.