A 2-D extension of the sampling algorithm for sparse fourier representations

Kazutaka Nishimura, Yoshinori Takei, Toshinori Yoshikawa,
Dept. Electrical Engineering
Nagaoka University of Techology
1603-1 Kamitomikamachi
Nagaoka-shi, Niigata 940-2188

Xi Zhang
Dept. Information and Communication Engineering
The University of Electro-Communications
1-5-1 Chofugaoka
Chofu-shi, Tokyo 182-8585

Abstract—Recently, a very efficient sampling algorithm for finding a $B$-term fourier representation of given 1-D discrete signal is presented [Gilbert, Guha, Indyk, Muthukrishnan, Strauss; STOC'02]. In this paper, we present a modified version of the algorithm, which can be applied to 2-D signals. As in the original algorithm, dependence of the running time on the signal length is polylogarithmic.

I. INTRODUCTION

Let

$$x = (x[0], x[1], \ldots, x[n], \ldots, x[N - 1])$$

be a complex-valued discrete time signal of length $N$. The discrete fourier transform (DFT), or the spectrum, of $x$,

$$\tilde{x} = (\tilde{x}[0], \tilde{x}[1], \ldots, \tilde{x}[k], \ldots, \tilde{x}[N - 1])$$

where

$$\tilde{x}[k] := \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] e^{-j2\pi nk/N},$$

is a fundamental tool for analyzing or representing features of the signal $x$. The inverse DFT (IDFT):

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \tilde{x}[k] e^{j2\pi nk/N},$$

recovers the signal $x$ completely from the spectrum $\tilde{x}$, and these transforms preserve the $l^2$ norm:

$$||x||_2 = ||\tilde{x}||_2,$$

where

$$||x||_2 := \left( \sum_{n=0}^{N-1} |x[n]|^2 \right)^{1/2}.$$  \hspace{1cm} (6)

Let $B < N$ be a positive integer and suppose that the frequency list $K(B) = (k_1, \ldots, k_B)$ is taken so that $\tilde{x}[k_1], \ldots, \tilde{x}[k_B]$ are the $B$-maxima among the spectrum $\tilde{x}$, in their absolute values. Then the signal $r_{opt}(B)$ defined by

$$r_{opt}(B)[n] = \frac{1}{\sqrt{N}} \sum_{k \in K(B)} \tilde{x}[k] e^{j2\pi nk/N}$$

is the best possible $l^2$ approximation of the signal $x$ by a $B$-term sum of pure tones $\psi_k[n] := e^{j2\pi nk/N}/\sqrt{N}$. We shall refer $r_{opt}(B)$ as the optimal $B$-term (fourier) representation of $x$.

Theorem 1 ([1]). There exists an algorithm, on input $B < N$, $\epsilon > 0$, and $x$, with cost $(B \log(N) \log(M)/\epsilon)^{O(1)}$, with high probability, output a $B$-term representation $r$ for $x$ which satisfies $||x - r||_2 \leq (1 + \epsilon)||x - r_{opt}(B)||_2$, where $r_{opt}(B)$ is the optimal $B$-term representation for $x$. The algorithm accesses only $(B \log(N) \log(M)/\epsilon)^{O(1)}$ samples $\{x[n] : n \in T\}$ of $x$, where the random sample set $T$ is chosen independently of $x$. (The success probability may be a constant arbitrarily close to 1 and the cost for it is implied in “$O(1)$”.)

Thus, when $B$ is sufficiently smaller than $N$, the algorithm, which we shall refer as the GGIMS algorithm in the sequel, enables very efficient computation of $B$-term representations.

Now, suppose that $x$ is a signal defined over the 2-dimensional domain $\{0, \ldots, N_1 - 1\} \times \{0, \ldots, N_2 - 1\}$:

$$x = (x[n_1, n_2] : 0 \leq n_1 < N_1, 0 \leq n_2 < N_2)$$
and let
\[ \bar{x} = (\bar{x}[k_1, k_2] : 0 \leq k_1 < N_1, 0 \leq k_2 < N_2) \]
be its 2-D DFT, where
\[ \bar{x}[k_1, k_2] := \sum_{n_1, n_2} x[n_1, n_2] \psi_{k_1, k_2}[-n_1, -n_2], \]
and
\[ \psi_{k_1, k_2}[n_1, n_2] := e^{j2\pi \left( \frac{k_1 n_1}{N_1} + \frac{k_2 n_2}{N_2} \right)} \sqrt{\frac{N_1 N_2}{N^2}}. \]

The optimal B-term representation for \( x \), i.e. the lists \( K(B) = ((k_1, 1, k_2, 1), \ldots, (k_1, B, k_2, B)) \) of frequencies and the corresponding coefficients \( \bar{x}[k_1, k_2, 1], \ldots, \bar{x}[k_1, B, k_2, B] \) where the latter are B-maxima of \( \bar{x} \) in absolute value, will be useful in this case also. The goal of this paper is to extend the GGIMS algorithm to an algorithm that efficiently compute near-optimal B-term representations for signals defined over a 2-dimensional domain.

The rest of the paper is organized as follows: Section II. is a brief description of the GGIMS algorithm [1] for 1-dimensional signals, where we shall discuss what parts of the algorithm are dimension-dependent. In Section III., we modify parts of the GGIMS algorithm so that it can be applied for 2-dimensional signals. Section IV. concludes the paper.

In the sequel, the time or frequency domain \( \{0, 1, \ldots, N - 1\} \) is identified with \( \mathbb{Z}/N\mathbb{Z} \), the ring of integers modulo \( N \). For simplicity, we assume that \( N, N_1, N_2 \) are odd prime numbers. \( E[X] \) means the expectation of a random variable \( X \).

II. THE GGIMS ALGORITHM

The GGIMS algorithm composed of 3 parts: Identification, Estimation and Iteration. Given the input signal \( x \), the algorithm start with the current list \( S \) set to null, and the corresponding at-most-B-term representation \( r \) set to 0.

A. Identification

Let \( \eta := \epsilon/(2B) \). The purpose of this part is to output a list \( \Lambda \) of frequencies, whose length is \( 2m + 1 = O(1/\eta) \), which contains (with high probability) all frequencies \( k \) satisfying \( |\bar{x}[k]|^2 \geq \eta ||x||^2_2 \). In this part there exist two procedures: Isolation and Group Testing. The latter procedure call a subroutine that (roughly) estimate the norm of given signal by sampling.

1) Isolation: This procedure constructs \( 2m + 1 \) signals \( f_0, \ldots, f_{2m} \) that satisfy (with high probability)

1. For each \( k' \) such that \( |\bar{x}[k']|^2 \geq \eta ||x||^2_2 \), there exists \( i \) (0 \leq i \leq 2m + 1) such that \( \hat{f}_i[k'] \geq 0.98 ||f_i||^2_2 \).

2. Each \( f_i \) can be sampled by sampling non-adaptively from \( x \) in \( O(m) \) places.

Here we note that the word “constructs” does not mean actual output of the entire signal points \( (f_i[0], f_i[1], \ldots, f_i[N - 1]) \).

The procedure only sets up parameters which enable later procedures to sample arbitrary signal points of \( \hat{f}_i \). To show the construction we define the operator \( R_{\theta, \sigma} \) for \( \theta, \sigma \in \mathbb{Z}/N\mathbb{Z} \) through
\[ (R_{\theta, \sigma} f)[n] = e^{j2\pi \theta n/N} f[n]. \]
Then it holds that
\[ (\hat{R}_{\theta, \sigma} f)[\sigma k + \theta] = \hat{f}[k]. \]

Also we define the Fejér kernel of length \( 2m + 1 \) as
\[ H_m[n] := \left\{ \begin{array}{ll} \sqrt{\frac{N}{2m+1}} & n \in [-m, m] \\ 0 & \text{otherwise.} \end{array} \right. \]

It holds that
\[ \hat{H}_m[k] = \left\{ \begin{array}{ll} \sin(\pi(2m+1)/N)/\sin(\pi k/N) & (k \neq 0) \\ 1 & (k = 0) \end{array} \right. \]
and in particular \( \hat{H}_m[k] \geq 2/\pi \) for \( k \in [-N/(2(2m + 1)), N/(2(2m + 1))] \). We regard the last interval as “pass region” of \( H_m \).

Now, the signals \( f_0, \ldots, f_{2m} \) are defined as follows: first, pick \( \theta \) and \( \sigma \) at random in \( \mathbb{Z}/N\mathbb{Z} \), with \( \sigma \) invertible. Then put, for each \( 0 \leq i \leq 2m \),
\[ f_i := (e^{j2\pi in/(2m+1)}H_m) * (R_{\theta, \sigma}x), \]
where * denotes the convolution. In the frequency domain, the spectrum \( \hat{x} \) is first permuted by \( k \mapsto \sigma k + \theta \), then filtered by “\( N/(2m+1) \)-shifted version of” \( \hat{H}_m \) to form \( \hat{f}_i \).

Using the fact that \( k \mapsto \sigma k + \theta \) is a pair-wise independent permutation (i.e., for all \( k_1 \neq k_2 \) and \( k_3 \neq k_4 \), \( k_1 \leftrightarrow k_3 \) and \( k_2 \leftrightarrow k_4 \) with probability \( 1/(N(N - 1)) \)), it can be shown that the following holds with high probability: for each \( k' \) such that \( |\bar{x}[k']|^2 \geq \eta ||x||^2_2 \), there exists \( i \) (0 \leq i \leq 2m + 1) such that \( \hat{f}_i[k'] \geq 0.98 ||f_i||^2_2 \). We note that sampling one point from \( f_i \) can be done by sampling \( 2m + 1 \) points from \( x \), by construction.

2) Group Testing: For each \( f \in \{f_0, \ldots, f_{2m}\} \), the procedure determines the frequency \( k' \) satisfying \( \hat{f}[k'] \geq 0.98 ||f||^2_2 \), if such \( k' \) exists. Then gather such \( k' \) from \( f_i \)'s to form the frequency list \( \Lambda \). By the construction of \( f_i \)'s, \( \Lambda \) catches all such frequencies that \( |\bar{x}[k']|^2 \geq \eta ||x||^2_2 \).

The procedure utilizes 16 filters
\[ \hat{G}_\ell[k] = \frac{1}{2} + \cos(\frac{2\pi \ell}{N} - \frac{2\pi \ell}{16}) \quad (\ell = 0, \ldots, 15), \]
each is of length 3. If we define the pass region of \( \hat{G}_\ell \) by
\[ \text{pass}_\ell := \{k : 2\pi \ell/N - 2\pi \ell/16 \leq 2\pi/32\}, \]
then \( \hat{G}_\ell[k] \geq 0.99 \) for \( k \in \text{pass}_\ell \) and these pass regions tile the entire frequency domain. The procedure also uses a subroutine, which will be explained in the next subsection, satisfying the
following: The subroutine makes \( O(\log \log M) \) samples, runs in time polynomial in \( \log(M) \), on input \( f \) (of length \( N \)) with biggest fourier coefficient \( \hat{f}[k] \), it returns, with high probability, a random output \( X \) such that:

1. \( X \leq ||f||_2^2 \), \( \forall f \).
2. If \( |\hat{f}[k]|^2 \geq 0.95||f||_2^2 \) then \( X \geq ||f||_2^2 \).

Now assume that we are seeking unknown \( k' \) which have at least 98% energy of \( f \). Without loss of generality, we assume that \( k' \in \text{ pass}_0 \). For each \( \ell = 0, 1, \ldots, 15 \), we estimate the norm \( ||f \ast G_{\ell}|| \) using the subroutine. For \( \ell = 0 \), it can be shown that \( ||f \ast G_{\ell}||^2 \geq 0.95||f \ast G_0||^2 \) by combining the facts \( k' \in \text{ pass}_0 \), \( G_0[k'] \geq 0.99 \) and \( \hat{f}[k'] \geq 0.98||f||_2^2 \). Then the second property of the subroutine, the returned value \( X \) satisfies \( X \geq 0.5||f \ast G_{\ell}||_2^2 \). It also can be shown that \( 0.5||f \ast G_{\ell}||_2^2 \geq 0.48||f||_2^2 \). On the other hand, for \( \ell = 4 \), it can be shown that the returned value \( X \) of the subroutine satisfies \( X \leq ||f \ast \tilde{G}_{\ell}||_2^2 \leq 0.38||f||_2^2 \); from the facts that \( k' \) is far from \( \text{ pass}_3 \) and that the contributions to \( ||f \ast \tilde{G}_{\ell}||_2^2 \) from frequencies other than \( k' \) is small (at most 0.02\( ||f||_2^2 \)). For \( \ell = 5, 6, \ldots, 12 \), we obtain \( X \leq 0.38||f||_2^2 \) in the same way. So, we can eliminate the possibilities that \( k' \in \text{ pass}_4 \) for \( \ell = 4, 5, \ldots, 12 \). In this manner, we can always eliminate 9/16 possibilities. The remaining region forms a cyclic interval of length at most \( 7N/16 \). By applying the operator \( R_{0/2} \), we can dilate the remaining frequency region by two. Then applying 16 filters, we can halve the possibilities again. Repeating that \( O(\log(N)) \) times, we learn \( k' \).

3) Estimating Norm: The subroutine for estimating norm is shown in Fig. 1, where \( \beta > 0 \) is a small constant (independent of \( N \) and \( M \)), and the number \( r \) of samples is \( O(\log \log M) \), and in fact we can reuse the same samples in the loop. \( K_c f \) stands means “\( f \) clipped at the ceiling \( c \)”, i.e.,

\[
(K_c f)[n] := \begin{cases} f[n] & (|f[n]| \leq c/\sqrt{N}) \\ 0 & \text{otherwise} \end{cases}
\]

(19)

Intuition behind the algorithm is as follows: Let \( Y = \frac{1}{N} \sum_{i=1}^{r} N|f[n_i]|^2 \). Then \( E[Y] = ||f||_2^2 \) and it would be a good estimate if the variance of \( Y \) were small. The variance comes from spikes, so we use the clipped version \( X \) instead. It is satisfied that \( 0 \leq E[X] \leq E[Y] = ||f||_2^2 \) and \( X \) has small variance.

If 95% of \( f \) is concentrated in a pure frequency, then the energy of spikes are small, so \( E[X] \approx E[Y] = ||f||_2^2 \), and since variance of \( X \) is small, \( X \) is a good estimation. The algorithm sets a ceiling \( c \), above which all values of the function are clipped as if they were spikes. The algorithm gradually lowers the clipping ceiling \( c \) until the energy estimate is consistent with the ceiling value.

### B. Estimation

This part estimates \( \tilde{x}[k] \), for each \( k \in \Lambda \), by a sampling algorithm.

The sampling algorithm used here takes inputs \( (a, k, \mu) \), where \( a \) is a signal of length \( N \), \( k \) a frequency, \( \mu \) an accuracy parameter. Then the algorithm runs with cost \( O(\log(M)/\mu) \), making \( O(1/\mu) \) samplings from \( a \), and returns the estimation \( \hat{a}[k] \), which satisfies with high probability that \( |\hat{a}[k] - \tilde{a}[k]|^2 \leq \epsilon ||a||_2^2 \). The estimation \( \tilde{a}[k] \) is computed as follows: let \( \ell = O(1/\mu) \) and pick \( \ell \) random positions \( n_1, \ldots, n_\ell \). Put \( a^{\ell} = N/\sum_{h=1}^{\ell} \Delta_{n_h} \), where \( \Delta_n \) is the delta function. Then compute \( X = a^{\ell}[k] \) and return it as the estimate. It can be shown that \( E[X] = \tilde{a}[k] \) and \( E[|X - \tilde{a}[k]|^2] \leq O(\epsilon ||a||_2^2) \).

For the estimation of \( \tilde{x}[k] \), an accuracy of \( \mu = 1/poly(B \log(N) \log(M/\delta)/\epsilon) \) (where \( \delta = ||x - r_{opt(B)}||_2 \)) is required.

### C. Iteration

Now we have the estimated fourier coefficients \( \tilde{x}[k] \) for \( k \in \Lambda \). Let \( \tilde{x}[k'] \) be the maximum (in absolute value) among them. Add \( k' \) to the current coefficients list \( S \), and update the current representation \( r \) by \( r \leftarrow r + \tilde{x}[k'] \psi_{k'} \). Then set \( x \leftarrow x - \tilde{x}[k'] \psi_{k'} \) and \( A \). However, we should note that the update of \( x \) just before goto is in fact impossible, since we can access \( x \) only via sampling. Actually, after 2nd iteration, we sample from \( x - r \) whenever we need sample from \( x \), using the maintained representation \( r \).

The algorithm may or may not halt just after that the representation \( r \) has grown to \( B \)-terms. Here we do not mention the accurate halting condition, however, it is guaranteed to halt within \( poly(B \log(N) \log(M/\delta)/\epsilon) \) iterations.

### III. MODIFICATIONS FOR 2-D SIGNALS

The strategy of GGIMS algorithm is not dimension-dependent. However, two key steps in the Identification part needs a modification for 2-D signals. Here we describe modifications for Isolation step and Group Testing step. In the sequel, signals are defined over the 2-dimensional domain \( Z/N_1Z \times Z/N_2Z \).

#### A. Identification (2-D)

1) Isolation (2-D): Let \( \eta := \epsilon/(2B) \). We will construct \((2n_1 + 1) \times (2n_2 + 1)\) signals \( f_{i_1,i_2}[n_1, n_2][i_1 =
0, 1, ..., 2m1, i2 = 0, 1, ..., 2m2), where (2m1 + 1) \times (2m2 + 1) = O(1/\eta).

We define the operator \( R_{\theta_1, \sigma_1, \theta_2, \sigma_2} \) through

\[
R_{\theta_1, \sigma_1, \theta_2, \sigma_2} x[n_1, n_2] = e^{2\pi i (\theta_1 n_1 + \theta_2 n_2 / N_2)} x[n_1, \sigma_2 n_2] (20)
\]

and 2-D Fejér kernel by

\[
H_{m_1, m_2}[n_1, n_2] = \begin{cases} \sqrt{N_1 N_2}/(2m_1 + 1)(2m_2 + 1) & (n_1, n_2) \in [-m_1, m_1] \times [-m_2, m_2] \\ 0 & \text{(otherwise)} \end{cases} (21)
\]

See Fig. 2 for the spectrum.

Then the signals \( f_{i_1, i_2}[n_1, n_2] \) for \( i_1 = 0, 1, ..., 2m_1, i_2 = 0, 1, ..., 2m_2 \) are constructed as follows: pick \( \sigma_1, \theta_1 \in \mathbb{Z}/N_1\mathbb{Z} \) at random, with \( \sigma_1 \) invertible. Independently, pick \( \sigma_2, \theta_2 \in \mathbb{Z}/N_2\mathbb{Z} \) at random, with \( \sigma_2 \) invertible. Then set

\[
f_{i_1, i_2}[n_1, n_2] := \left( e^{2\pi i (i_1 n_1/(2m_1 + 1) + i_2 n_2/(2m_2 + 1))} H_{m_1, m_2}[n_1, n_2] \right) \ast R_{\theta_1, \sigma_1, \theta_2, \sigma_2} x[n_1, n_2]. (22)
\]

It can be shown that they have the following (expected) properties:

1. For each \( k_1’, k_2’ \) such that \( |\hat{x}|[k_1’, k_2’]|^2 \geq \eta |x|^2 \), there exists \( (i_1, i_2) \) \((0 \leq i_1 \leq 2m_1 + 1, 0 \leq i_2 \leq 2m_2 + 1)\) such that \( \hat{f}_{i_1, i_2}[k_1’, k_2’] \leq 0.98 |f_{i_1, i_2}|^2 \).

2. Each \( f_{i_1, i_2} \) can be sampled by sampling non-adaptively from \( x \) in \( O(m_1 m_2) \) places.

2) **Group Testing (2-D):** In 2-D case, we rely on the following 256 filters (see Fig. 3):

\[
\hat{G}_{\ell_1, \ell_2}[k_1, k_2] = \frac{1}{2} \left\{ 1 + \cos \left( \frac{2\pi k_1}{N_1} - \frac{2\pi \ell_1}{16} \right) \right\} \left\{ 1 + \cos \left( \frac{2\pi k_2}{N_2} - \frac{2\pi \ell_2}{16} \right) \right\} (0 \leq \ell_1 < 16, 0 \leq \ell_2 < 16) (23)
\]

In the time domain, \( G_{\ell_1, \ell_2} \) has a support of \( 3 \times 3 \) square.

With these filters, we can eliminate \( 1 - (7/16)^2 = 207/256 \) (instead of \( 1 - 7/16 \) in 1-D case) of possibilities in one step of Group Testing.

**B. Complexity of 2-D algorithm**

By the above modifications of the algorithm, we obtain a natural 2-D extension of Theorem 1:

**Theorem 2.** There exists an algorithm, on input \( B < N_1 N_2, \epsilon > 0, \) and a signal \( x \) over the 2-dimensional domain \{0, ..., N_1 - 1\} \times \{0, ..., N_2 - 1\}, with cost \((B \log(N_1) \log(N_2) \log(M)/\epsilon)^O(1)\), with high probability, output a \( B \)-term representation \( \mathbf{r} \) for \( x \) which satisfies \( ||x - \mathbf{r}||_2 \leq (1 + \epsilon)||x - \mathbf{r}_{\text{opt}}(B)||_2 \), where \( \mathbf{r}_{\text{opt}}(B) \) is the optimal \( B \)-term representation for \( x \). The algorithm accesses only \((B \log(N_1) \log(N_2) \log(M)/\epsilon)^O(1)\) samples \{\( x[n_1, n_2] \) : \((n_1, n_2) \in T\)\} of \( x \), where the random sample set \( T \) is chosen independently of \( x \). (The success probability may be a constant arbitrarily close to 1 and the cost for it is implied in “O(1)”.)

**IV. CONCLUSION**

In this paper, we have presented an efficient algorithm that finds a near-optimal \( B \)-term fourier representation of 2-D signals, by modifying a part of GGI-M algorithm [1].

The running time of the algorithm depends on signal amount only polylogarithmically, so the algorithm will be useful for digesting vast 2-D signals.

The algorithm has many hidden constants that would affect practical use. So more fine analysis of the algorithm is still needed.

**REFERENCES**