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

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

and

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; STOC02]. 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

$$\mathbf{x} = (x[0], x[1], \ldots, x[n], \ldots, x[N-1]) \tag{1}$$

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

$$\widehat{\mathbf{x}} = (\widehat{x}[0], \widehat{x}[1], \ldots, \widehat{x}[k], \ldots, \widehat{x}[N-1]) \tag{2}$$

where

$$\widehat{x}[k] := \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} x[n] e^{-j2\pi nk/N}, \tag{3}$$

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

$$x[n] = \frac{1}{\sqrt{N}} \sum_{k=0}^{N-1} \widehat{x}[k] e^{j2\pi nk/N}, \tag{4}$$

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

$$||\mathbf{x}||_2 = ||\widehat{\mathbf{x}}||_2, \tag{5}$$

where

$$||\mathbf{x}||_2 := \sqrt{\sum_{n=0}^{N-1} |x[n]|^2}. \tag{6}$$

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

$$r_{\mathrm{opt}(B)}[n] = \frac{1}{\sqrt{N}} \sum_{k \in K(B)} \widehat{x}[k] e^{j2\pi nk/N} \tag{7}$$

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

In those applications that the optimal B-term representation permit a good approximation to the original signal with $B << N$, the B-term representation, or equivalently, the lists $K(B)$ of frequencies and $\widehat{x}[k_1], \ldots, \widehat{x}[k_B]$ of coefficients, would be a nice "digest" of the original signal, in the views of conciseness and accuracy. The computation is obviously feasible, by a combination of the fast fourier transform (FFT) algorithm and some sorting algorithm. The running time has a lower bound $\Omega(N \log N)$ (here we regard the required precision as a constant), which sounds not so big. However, when we want to make B-term digest for each of $\mathbf{x}$'s, in a huge database, each having a really long length N, the cost $N \log N$ by the number of $\mathbf{x}$'s will be so expensive. The formula (3) shows that every fourier coefficient $\widehat{x}[k]$ depends on entire signal $\mathbf{x}$, and it would be natural that one think it were impossible to reduce the computational complexity even with the sparseness condition $B << N$ (say $B = 10$ independent of N.)

Gilbert, Guha, Indyk, Muthukrishnan, and Strauss [1] have shown a somewhat surprising result that a *near*-optimal B-term representation of any signal can be obtained *with high probability*, within a time whose dependence on N is poly-logarithmic. Let ι be a small number such that, if $||\mathbf{x} - \mathbf{y}||_2^2 \le \iota$ then we consider that two signals $\mathbf{x}$ and $\mathbf{y}$ are identical. By a normalization, we assume that $\iota = 1$ and M is a crude upper bound on $||x||_2$ after the normalization.

Theorem 1 ([1]). *There exists an algorithm, on input $B < N$, $\epsilon > 0$, and $\mathbf{x}$, with cost $(B \log(N) \log(M)/\epsilon)^{O(1)}$, with high probability, output a B-term representation $\mathbf{r}$ for $\mathbf{x}$ which satisfies $||\mathbf{x} - \mathbf{r}||_2 \le (1+\epsilon)||\mathbf{x} - \mathbf{r}_{\mathrm{opt}(B)}||_2$, where $\mathbf{r}_{\mathrm{opt}(B)}$ is the optimal B-term representation for $\mathbf{x}$. The algorithm accesses only $(B \log(N) \log(M)/\epsilon)^{O(1)}$ samples $\{x[n] : n \in T\}$ of $\mathbf{x}$, where the random sample set T is chosen independently of $\mathbf{x}$. (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 $\mathbf{x}$ is a signal defined over the 2-dimensional domain $\{0, \ldots, N_1 - 1\} \times \{0, \ldots, N_2 - 1\}$:

$$\mathbf{x} = (x[n_1, n_2] : 0 \le n_1 < N_1, 0 \le n_2 < N_2) \tag{8}$$

and let

$$\widehat{\mathbf{x}} = (\widehat{x}[k_1, k_2] : 0 \leq k_1 < N_1, 0 \leq k_2 < N_2) \tag{9}$$

be its 2-D DFT, where

$$\widehat{x}[k_1, k_2] := \sum_{n_1, n_2} x[n_1, n_2]\psi_{k_1,k_2}[-n_1, -n_2], \tag{10}$$

and

$$\psi_{k_1,k_2}[n_1, n_2] := \frac{e^{j2\pi\left(\frac{k_1 n_1}{N_1} + \frac{k_2 n_2}{N_2}\right)}}{\sqrt{N_1 N_2}}. \tag{11}$$

The optimal B-term representation for $\mathbf{x}$, 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 $\widehat{x}[k_{1,1}, k_{2,1}], \ldots, \widehat{x}[k_{1,B}, k_{2,B}]$ where the latter are B-maxima of $\widehat{x}$ 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 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 $\mathbf{Z}/N\mathbf{Z}$, the ring of integers modulo N. For simplicity, we assume that N, N_1, N_2 are odd prime numbers. $\mathsf{E}[X]$ 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 $\mathbf{x}$, the algorithm start with the current list S set to null, and the corresponding at-most-B-term representation $\mathbf{r}$ is set to 0.

A. Identification

Let $\eta := \epsilon/(2B)$. The purpose of this part is to output a list Λ of frequencies, whose length is $2m + 1 = O(1/\eta)$, which contains (with high probability) all frequencies k satisfying $|\widehat{x}[k]|^2 \geq \eta||\mathbf{x}||_2^2$. 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 $\mathbf{f}_0, \ldots \mathbf{f}_{2m}$ that satisfy (with high probability)

1. For each k' such that $|\widehat{x}[k']|^2 \geq \eta||\mathbf{x}||_2^2$, there exists $i(0 \leq i \leq 2m + 1)$ such that $\widehat{f}_i[k'] \geq 0.98||\mathbf{f}_i||_2^2$.
2. Each $\mathbf{f}_i$ can be sampled by sampling non-adaptively from $\mathbf{x}$ 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 $\vec{f}_i$. To show the construction we define the operator $R_{\theta,\sigma}$ for $\sigma, \theta \in \mathbf{Z}/N\mathbf{Z}$ through

$$(R_{\theta,\sigma}f)[n] = e^{j2\pi\theta n/N} f[\sigma n]. \tag{12}$$

Then it holds that

$$(\widehat{R_{\theta,\sigma}f})[\sigma k + \theta] = \widehat{f}[k]. \tag{13}$$

Also we define the Fejér kernel of length $2m + 1$ as

$$H_m[n] := \begin{cases} \frac{\sqrt{N}}{2m+1} & n \in [-m, m] \\ 0 & \text{otherwise.} \end{cases} \tag{14}$$

It holds that

$$\widehat{H}_m[k] = \begin{cases} \frac{\sin(\pi(2m+1)/N)}{(2m+1)\sin(\pi k/N)} & (k \neq 0) \\ 1 & (k = 0) \end{cases} \tag{15}$$

and in particular $\widehat{H}_m[k] \geq 2/\pi$ for $k \in [-N/(2(2m + 1)), N/(2(2m + 1))]$. We regard the last interval as "pass region" of $\mathbf{H}_m$

Now, the signals $\mathbf{f}_0, \ldots \mathbf{f}_{2m}$ are defined as follows: first, pick θ and σ at random in $\mathbf{Z}/N\mathbf{Z}$, with σ invertible. Then put, for each $0 \leq i \leq 2m$,

$$\mathbf{f}_i := (e^{j2\pi in/(2m+1)}\mathbf{H}_m) * (R_{\theta,\sigma}\mathbf{x}), \tag{16}$$

where $*$ denotes the convolution. In the frequency domain, the spectrum $\widehat{x}$ is first permuted by $k \mapsto \sigma k + \theta$, then filtered by "$iN/(2m + 1)$-shifted version of" $\widehat{\mathbf{H}}_m$ to form $\widehat{\mathbf{f}}_i$.

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 $k_3 \neq k_4$, $k_1 \mapsto k_3$ and $k_2 \mapsto k_4$ with probability $1/(N(N-1))$), it can be shown that the following holds with high probability: for each k' such that $|\widehat{x}[k']|^2 \geq \eta||\mathbf{x}||_2^2$, there exists $i(0 \leq i \leq 2m + 1)$ such that $\widehat{f}_i[k'] \geq 0.98||\mathbf{f}_i||_2^2$. We note that sampling one point from $\mathbf{f}_i$ can be done by sampling $2m+1$ points from $\mathbf{x}$, by construction.

2) Group Testing: For each $\mathbf{f} \in \{\mathbf{f}_0, \ldots \mathbf{f}_{2m}\}$, the procedure determines the frequency k' satisfying $\widehat{f}[k'] \geq 0.98||\mathbf{f}||_2^2$, if such k' exists. Then gather such k' from $\mathbf{f}_i$'s to form the frequency list Λ. By the construction of $\mathbf{f}_i$'s, Λ catches all such frequencies that $|\widehat{x}[k']|^2 \geq \eta||\mathbf{x}||_2^2$.

The procedure utilizes 16 filters

$$\hat{G}_\ell[k] = \frac{1}{2}(1 + \cos(\frac{2\pi k}{N} - \frac{2\pi\ell}{16})) \quad (\ell = 0, \ldots, 15), \tag{17}$$

each is of length 3. If we define the pass region of $\hat{\mathbf{G}}_\ell$ by

$$\mathrm{pass}_\ell := \{k : |2\pi k/N - 2\pi\ell/16| \leq 2\pi/32\}, \tag{18}$$

then $\hat{G}_\ell[k] \geq 0.99$ for $k \in \mathrm{pass}_\ell$ 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

$c \leftarrow 2M/\beta$
Do
{
  Pick r random independent samples $n_i \in \mathbf{Z}/N\mathbf{Z}$
  $X \leftarrow \frac{1}{r}\sum_{i=1}^{r} N|(K_c f)[n_i]|^2$
  $c \leftarrow c/(1+\beta)$
} while $X < \beta^2 c^2$ and $c \geq 1$
 Output $X/(1+\beta)$

Fig. 1. Estimating Norm

following: The subroutine makes $O(\log\log M)$ samples, runs in time polynomial in $\log(M)$, on input $\mathbf{f}$ (of length N) with biggest fourier coefficient $\widehat{f}[k]$, it returns, with high probability, a random output X such that:

1. $X \leq ||\mathbf{f}||_2^2, \forall \mathbf{f}$.
2. If $|\widehat{f}[k]|^2 \geq 0.95||\mathbf{f}||_2^2$ then $X \geq 0.5||\mathbf{f}||_2^2$.

Now assume that we are seeking unknown k' which have at least 98% energy of $\mathbf{f}$. Without loss of generality, we assume that $k' \in \mathrm{pass}_0$. For each $\ell = 0, 1, \ldots, 15$, We estimate the norm $||\mathbf{f} * G_\ell||$ using the subroutine. For $\ell = 0$, it can be shown that $|\widehat{\mathbf{f} * \mathbf{G}}_0|^2 \geq 0.95||\mathbf{f} * \mathbf{G}_0||_2^2$ by combining the facts $k' \in \mathrm{pass}_0$, $\hat{G}_0[k'] \geq 0.99$ and $\widehat{f}[k'] \geq 0.98||\mathbf{f}||_2^2$. Then the second property of the subroutine, the returned value X satisfies $X \geq 0.5||\mathbf{f} * \hat{G}_0||_2^2$. It also can be shown that $0.5||\mathbf{f} * \hat{G}_0||_2^2 \geq 0.48||\mathbf{f}||_2^2$. On the other hand, for $\ell = 4$, it can be shown that the returned value X of the subroutine satisfies $X \leq ||\mathbf{f} * \hat{G}_4||_2^2 \leq 0.38||\mathbf{f}||_2^2$, from the facts that k' is far from pass_4 and that the contributions to $||\mathbf{f} * \hat{G}_4||_2^2$ from frequencies other than k' is small (at most $0.02||\mathbf{f}||_2^2$). For $\ell = 5, 6, \ldots, 12$, we obtain $X \leq 0.38||\mathbf{f}||_2^2$ in the same way. So, we can eliminate the possibilities that $k' \in \mathrm{pass}_\ell$ for $\ell = 4, \ldots, 12$. In this manner, we can always eliminate $9/16$ possibilities. The remaining region forms a cyclic interval of lentgth 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,

$$(K_c f)[n] := \begin{cases} f[n] & (|f[n]| \leq c/\sqrt{N}) \\ 0 & \text{otherwise.} \end{cases} \tag{19}$$

Intuition behind the algorithm is as follows: Let $Y = \frac{1}{r}\sum_{i=1}^{r} N|f[n_i]|^2$. Then $\mathsf{E}[Y] = ||\mathbf{f}||_2^2$ 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 $0 \leq \mathsf{E}[X] \leq \mathsf{E}[Y] = ||\mathbf{f}||_2^2$ and X has small variance. If 95% of $\mathbf{f}$ is concentrated in a pure frequency, then the energy of spikes are small, so $\mathsf{E}[X] \approx \mathsf{E}[Y] = ||\mathbf{f}||_2^2$, 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 celing c until the energy estimate is consistent with the ceiling value.

B. Estimation

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

The sampling algorithm used here takes inputs $(\mathbf{a}, k, \mu)$, where $\mathbf{a}$ is a signal of length N, k a frequency, μ an accuracy parameter. Then the algorithm runs with cost $O(\log(M)/\mu)$, making $O(1/\mu)$ samplings from $\mathbf{a}$, and returns the estimation $\widetilde{\widehat{a}}[k]$, which satisfies with high probability that $|\widetilde{\widehat{a}}[k] - \widehat{a}[k]|^2 \leq \epsilon||\mathbf{a}||_2^2$. The estimation $\widetilde{\widehat{a}}[k]$ is computed as follows: let $t = O(1/\mu)$ and pick t random positions $n_1, \ldots, n_t$. Put $\mathbf{a}'' = N/t\sum_{h=1}^{t} \Delta_{n_h}$, where Δ_n is the delta function. Then compute $X = a''[k]$ and return it as the estimate. It can be shown that $\mathsf{E}[X] = \widehat{a}[k]$ and $\mathsf{E}[|X - \widehat{a}[k]|^2] \leq O(\epsilon||\mathbf{a}||_2^2)$.

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

C. Iteration

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

The algorithm may or may not halt just after that the representation $\mathbf{r}$ 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 $\mathbf{Z}/N_1\mathbf{Z} \times \mathbf{Z}/N_2\mathbf{Z}$.

A. Identification (2-D)

1) Isolation (2-D): Let $\eta := \epsilon/(2B)$.We will construct $(2m_1 + 1) \times (2m_2 + 1)$ signals $f_{i_1,i_2}[n_1, n_2] (i_1 =$

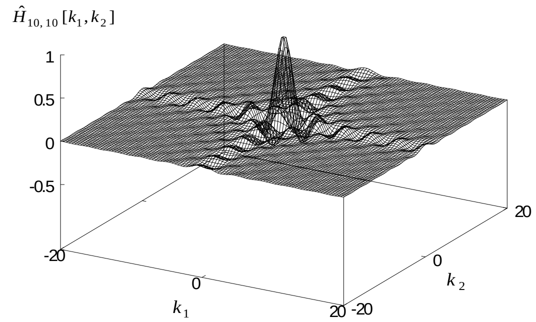

Fig. 2. $\hat{H}_{10,10}[k_1, k_2](N_1 = N_2 = 40)$

$0, 1, \ldots, 2m_1, i_2 = 0, 1, \ldots, 2m_2)$, where $(2m_1+1)\times(2m_2+1) = O(1/\eta)$.

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

$$R_{\theta_1,\sigma_1,\theta_2,\sigma_2}x[n_1, n_2] = e^{j2\pi(\theta_1 n_1/N_1+\theta_2 n_2/N_2)} x[\sigma_1 n_1, \sigma_2 n_2] \quad (20)$$

and 2-D Fejér kernel by

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

See Fig.2 for the spectrum.

Then the signals $f_{i_1,i_2}[n_1, n_2](i_1 = 0, 1, \ldots, 2m_1, i_2 = 0, 1, \ldots, 2m_2)$ are constructed as follows: pick $\sigma_1, \theta_1 \in \mathbf{Z}/N_1\mathbf{Z}$ at random, with σ_1 invertible. Independently, pick $\sigma_2, \theta_2 \in \mathbf{Z}/N_2\mathbf{Z}$ at random, with σ_2 invertible. Then set

$$f_{i_1,i_2}[n_1, n_2] := (e^{j2\pi(i_1 n_1/(2m_1+1)+i_2 n_2/(2m_2+1))} H_{m_1,m_2}[n_1, n_2]) * R_{\theta_1,\sigma_1,\theta_2,\sigma_2}x[n_1, n_2]. \quad (22)$$

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

1. For each k'_1, k'_2 such that $|\widehat{x}[k'_1, k'_2]|^2 \geq \eta||\mathbf{x}||_2^2$, there exists $(i_1, i_2)(0 \leq i_1 \leq 2m_1 + 1, 0 \leq i_2 \leq 2m_2 + 1)$ such that $\widehat{f}_{i_1,i_2}[k'_1, k'_2] \leq 0.98||\mathbf{f}_{i_1,i_2}||_2^2$.
2. Each $\mathbf{f}_{i_1,i_2}$ can be sampled by sampling non-adaptively from $\mathbf{x}$ 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):

$$\hat{G}_{\ell_1,\ell_2}[k_1, k_2] = \frac{1}{4}\left\{1 + \cos\left(\frac{2\pi k_1}{N_1} - \frac{2\pi\ell_1}{16}\right)\right\} \cdot \left\{1 + \cos\left(\frac{2\pi k_2}{N_2} - \frac{2\pi\ell_2}{16}\right)\right\} \quad (0 \leq \ell_1 < 16, 0 \leq \ell_2 < 16) \quad (23)$$

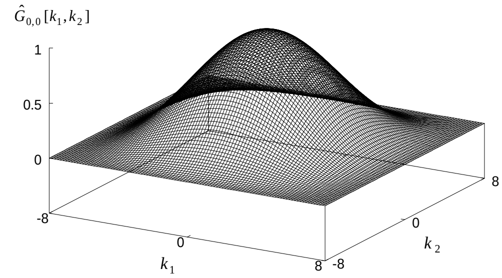

Fig. 3. The filter $\hat{G}_{0,0}[k_1, k_2]$ $(N_1 = N_2 = 16)$

In the time domain, $\mathbf{G}_{\ell_1,\ell_2}$ has a support of 3×3 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 $B < N_1N_2$, $\epsilon > 0$, and a signal $\mathbf{x}$ over the 2-dimensional domain $\{0, \ldots, N_1 - 1\} \times \{0, \ldots, N_2 - 1\}$, 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 $\mathbf{x}$ which satisfies $||\mathbf{x} - \mathbf{r}||_2 \leq (1+\epsilon)||\mathbf{x} - \mathbf{r}_{\mathrm{opt}(B)}||_2$, where $\mathbf{r}_{\mathrm{opt}(B)}$ is the optimal B-term representation for $\mathbf{x}$. 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 $\mathbf{x}$, where the random sample set T is chosen independently of $\mathbf{x}$. (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 GGIMS 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

[1] A. C. Gilbert, S. Guha, P. Indyk, S. Muthukrishnan, and M. Strauss, "Near-Optimal Sparse Fourier Representations via Sampling," *In Proceedings of the 34th Annual ACM Symposium on Theory of Computing (STOC)*, pp.152-161, May 2002.