A PRACTICAL HIGH-DIMENSIONAL SPARSE FOURIER TRANSFORM

Shaogang Wang, Vishal M. Patel and Athina Petropulu

Department of Electrical and Computer Engineering
Rutgers, The State University of New Jersey, Piscataway, NJ 08854, USA

ABSTRACT

As compared to the FFT, the recently introduced Sparse Fourier Transform (SFT) achieves substantial reduction in the complexity of detecting frequencies in signals that are sparse in the frequency domain. However, the SFT requires the significant frequencies to be on the grid and the exact sparsity of the signal to be known. In this paper, we propose a framework that overcomes these issues. Our method makes use of a pre-permutation window to confine the leakage within finite frequency bins and the Neyman-Pearson criterion to detect weak signals without knowing the exact signal sparsity. Various numerical experiments and an application to radar target detection demonstrate the advantages of the proposed method.

***Index Terms*—** Multi-dimensional signal processing, sparse Fourier transform, detection and estimation.

1. INTRODUCTION

Conventional signal processing in radar, sonar, and communication systems usually involves applying the Discrete Fourier Transform (DFT) followed by a detection stage. The DFT is usually implemented via the Fast Fourier Transform (FFT). Recently, by leveraging the sparsity of a signal in the frequency domain, the Sparse Fourier Transform (SFT) [1, 2] can further reduce the complexity required to identify the underlying frequencies. Different versions of the SFT have been investigated in several applications including a fast Global Positioning System (GPS) receiver, wide-band spectrum sensing, light field reconstruction, etc. [3–5].

One of the constrains in the aforementioned SFT algorithms is the assumption that the signal frequencies are all on-grid. In reality, however, depending on the grid size, the signal frequencies can fall between grid points. The consequence of off-grid frequencies is leakage to other frequency bins, which essentially reduces the sparsity of the signal. Our early work to address these issues appeared in [6], where the Realistic Sparse Fourier Transform (RSFT) was proposed, i.e., a high dimensional extension of the SFT in [2], which allows off-grid frequencies owning to a *pre-permutation windowing* process.

The current literature on multi-dimensional extensions of the SFT [7–11] mainly addresses sample complexity, i.e., using the least number of time domain samples to reconstruct the signal frequencies. In order to detect the significant frequencies in an approximately sparse settings, arising for example when the signal is corrupted by additive noise, the aforementioned methods assume knowing the exact sparsity, and compare the frequency domain peaks with a predefined threshold. However, in many real applications, the exact signal sparsity may be either unknown or subject to change. For example, in the radar case, in which the sparsity of the received signal in some high dimensional space corresponds to the the number of targets to be detected, it is typically unknown and usually varies from time to time. Also, setting up an ideal detection threshold is not trivial in noisy cases since it relates to the tradeoff between probability of detection and false alarm rate.

In this work, we put the RSFT of [6] into a Neyman-Pearson (NP) detection framework. As it will be explained later, the RSFT involves two detection stages which are interconnected, i.e., the output of the first stage serves as the input of the second stage. Based on the signal model and other design specifications, we give the (asymptotically) optimal thresholds for the two detection stages. These thresholds are jointly found by formulating and solving an optimization problem, which minimizes the Signal to Noise Ratio (SNR) corresponding to the weakest sinusoid within the signal, and its constraints reflect the relationship between the signal SNR, probability of detection and false alarm rate for both stages. As a result, the detection of frequencies does not require the knowledge of the exact signal sparsity. Due to space constraints, some details of the work are deferred to an extended version of this paper [12].

Notation: We use lower-case (upper-case) bold letters to denote vectors (matrices). $(\cdot)^T$ and $(\cdot)^H$ respectively denote the transpose and conjugate transpose of a matrix or a vector. $\|\cdot\|$ is Euclidean norm for a vector. $[\mathbf{a}]_i$ is the i_{th} element of vector $\mathbf{a}$. All operations on indices in this paper are taken modulo N, denoted by $[\cdot]_N$. We use $\lfloor\cdot\rfloor$ to denote rounding to floor. $[S]$ refers to the set of indices $\{0, ..., S-1\}$, and $[S] \setminus a$ is for eliminating element a from set $[S]$. We use $\{0,1\}^B$ to denote the set of B-dimensional binary vectors. We use $\mathrm{diag}(\mathbf{v})$ to denote forming a diagonal matrix from the vector $\mathbf{v}$ and use $\mathbb{E}\{\cdot\}$ to denote expectation. The DFT of signal $\mathbf{s}$ is denoted as $\hat{\mathbf{s}}$. We also assume that the signal length in each dimension is an integer power of 2.

2. SIGNAL MODEL AND PROBLEM FORMULATION

We model the continuous time signal as a superposition of K sinusoids plus additive white noise. We sample the signal uniformly both in I and Q channels with a sampling frequency above the Nyquist rate. Assume the total sampling time is divided into T consecutive equal length segments, each of which contains N samples, and $K << N$ (i.e., the signal is sparse in the frequency domain). The signal samples over the time segment $s \in [T]$, can be expressed in vector form as

$$\mathbf{r}_s = \sum_{i\in[K]} b_{i,s}\mathbf{v}(\omega_i) + \mathbf{n}_s, \tag{1}$$

where $\mathbf{v}(\omega_i)$ denotes the i_{th} complex sinusoid with frequency $\omega_i \in [0, 2\pi)$, i.e., $\mathbf{v}(\omega_i) = [1 \quad e^{j\omega_i} \quad \cdots \quad e^{j(N-1)\omega_i}]^T$, and $b_{i,s}$ denotes the corresponding complex weight. We further assume that ω_i is unknown deterministic and constant during the whole process, and $b_{i,s}$

The work was supported by ARO grant W911NF-16-1-0126, NSF Grant ECCS-1408437, China Scholarship Council and Shanghai Institute of Spaceflight Electronics Technology.

takes random value for each segment. More specifically, we model $b_{i,s}$ as a circularly symmetric complex Gaussian process with distribution $b_{i,s} \sim \mathcal{CN}(0, \sigma_{bi}^2)$. Likewise, the noise $\mathbf{n}_s$ is distributed as $\mathbf{n}_s \sim \mathcal{CN}(\mathbf{0}, \sigma_n^2\mathbf{I})$, where $\mathbf{0}$ is an N-dimensional zero vector, and $\mathbf{I} \in \mathbb{R}^{N\times N}$ is the identity matrix. We also assume that each sinusoid and the noise are uncorrelated. In addition, the neighboring sinusoids are resolvable in the frequency domain, i.e., their frequency spacing is greater than $\eta_m \frac{2\pi}{N}$, where $\eta_m \in \mathbb{N}$ is a broadening parameter of a window that applies on $\mathbf{r}_s$.

Let $SNR_i = \sigma_{bi}^2/\sigma_n^2$ be the SNR of the i_{th} sinusoid. Let us define the *worst case SNR* as $SNR_{min} \triangleq \min_{i\in[K]}(SNR_i)$. Let P_d denote for the probability of detection for the sinusoid with SNR_{min}, and P_{fa} the corresponding probability of false alarm on each frequency bin.

We want to detect and estimate each ω_i from the input signal with certain requirements on SNR_{min}, P_d and P_{fa}. From a non-parametric and data-independent perspective, this is a classic *spectral analysis* and detection problem that can be solved by any FFT-based spectrum estimation method. For instance, this can be done by using the Bartlett method [13] followed by an NP detection procedure. In this case, the FFT computes the signal spectrum on N frequency bins, and the detection is carried on each bin to determine whether there exists a significant frequency or not. In what follows, we will solve this detection and estimation problem using the RSFT.

3. DETECTION IN THE RSFT

The RSFT is summarized in Algorithm 1. First, the pre-permutation windowing confines the leakage from off-grid frequencies within limited number of frequency bins. Then, the *permutation* procedure reorders the input data in the time domain, causing the frequencies to also reorder. The permutation causes closely spaced frequencies to appear in well separated locations with high probability. Then, a *flat-window* [2] is applied on the permuted signal for the purpose of extending a single frequency into a (nearly) boxcar, for a reason that will become apparent in the following. The windowed data are aliased, and the frequency domain equivalent of this *aliasing* is undersampling by N/B. The flat-window used at the previous step ensures that no peaks are lost due to the effective undersampling in the frequency domain. After this stage, an FFT of length B is employed. The permutation and the aliasing procedure effectively map the signal frequencies from the N-dimensional space into a reduced B-dimensional space, where the *first stage detection* procedure finds the significant frequencies' peaks, and subsequently their indices are *reverse mapped* into the original N-dimensional frequency space. However, the reverse mapping yields not only the true location of the significant frequencies, but also N/B ambiguous locations for each frequency. To remove the ambiguity, multiple iterations of processing with randomized permutation are performed. Finally, the *second stage detection* procedure locates the K most significant frequencies from the accumulated data for each iteration.

The RSFT algorithm discussed in this work is slightly different from that presented in [6]. In this paper, we apply the iteration from the pre-permutation windowing to the accumulation stage for each different data segment, while in [6], we assume that the input data is the same for each iteration. The main reason for this change is to reduce the variance of the estimation. The major difference from [6] is that we apply NP criteria in both detection stages, which obviates the need for knowing the signal sparsity; in our previous work [6], the first and second stage detection was performed by counting respectively the dK and K highest peaks in the frequency domain,

Algorithm 1 RSFT algorithm

Input: complex signal $\mathbf{r}_s, s \in [T]$ in any fixed dimension
Output: $\mathbf{o}$, sparse frequency locations of input signal

procedure RSFT($\mathbf{r}_s$)
  Generate a set of $\sigma_s, s \in [T]$ randomly for each dimension
  $\bar{\mathbf{a}} \leftarrow \mathbf{0}$
  for $s \leftarrow 0$ **to** T **do**
    Pre-Permutation windowing: $\mathbf{y} \leftarrow \mathbf{W}\mathbf{r}_s$
    Permutation: $\mathbf{p} \leftarrow \mathbf{P}_{\sigma_s}\mathbf{y}$
    Flat windowing: $\mathbf{z} \leftarrow \overline{\mathbf{W}}\mathbf{p}$
    Aliasing: $\mathbf{f} \leftarrow \mathrm{Aliasing}(\mathbf{z})$
    N-D FFT: $\hat{\mathbf{f}} \leftarrow \mathrm{NDFFT}(\mathbf{f})$
    First stage detection: $\mathbf{c} \leftarrow \mathrm{NPdet1}(|\hat{\mathbf{f}}|^2)$
    Reverse mapping: $\mathbf{a} \leftarrow \mathrm{Reverse}(\mathbf{c})$
    Accumulation: $\bar{\mathbf{a}} \leftarrow \bar{\mathbf{a}} + \mathbf{a}$
  end for
  Second stage detection: $\mathbf{o} \leftarrow \mathrm{NPdet2}(\bar{\mathbf{a}})$
  return $\mathbf{o}$
end procedure

where d is a number empirically determined by the pre-permutation windows in each dimension.

In the following, we investigate the two stages of detection separately, then summarize the solution into an optimization problem. The analysis is done in one dimension, while the high-dimensional extension is straightforward.

First Stage Detection: The first stage detection is performed on each data segment. After pre-permutation windowing, permutation and flat-windowing, the input signal can be expressed as $\mathbf{z} = \overline{\mathbf{W}}\mathbf{P}_{\sigma_s}\mathbf{W}\mathbf{r}_s$, where σ_s is the permutation parameter for the s_{th} segment and is assumed uniformly distributed; $\mathbf{P}_{\sigma_s}$ is the permutation matrix, and $[\mathbf{P}_{\sigma_s}\mathbf{x}]_i = [\mathbf{x}]_{\sigma_s i}$; $\mathbf{W} = \mathrm{diag}(\mathbf{w})$, $\overline{\mathbf{W}} = \mathrm{diag}(\bar{\mathbf{w}})$, where $\mathbf{w}$ and $\bar{\mathbf{w}}$ are pre-permutation window and flat-window, respectively.

The time domain aliasing can be described as $\mathbf{f} = \mathbf{V}_{\sigma_s}\mathbf{r}_s$, where $\mathbf{V}_{\sigma_s} = \sum_{i\in[L]} \overline{\mathbf{W}}_i\mathbf{P}_{\sigma_s}\mathbf{W}$, and $L = N/B$, $\overline{\mathbf{W}}_i$ is the i_{th} sub-matrix of $\overline{\mathbf{W}}$, which is comprised of the iB_{th} to the $((i+1)B-1)_{th}$ rows of $\overline{\mathbf{W}}$. Applying FFT on $\mathbf{f}$, the output of the k_{th} entry can be expressed as $[\hat{\mathbf{f}}]_k = \mathbf{u}_k^H\mathbf{V}_{\sigma_s}\mathbf{r}$, $k \in [B]$, where $\mathbf{u}_k = [1 \quad e^{jk\Delta\omega_B} \quad \cdots \quad e^{jk(B-1)\Delta\omega_B}]^T$, and $\Delta\omega_B = 2\pi/B$. Substituting (1) into $\mathbf{f}$, and taking out of the summation the m_{th} sinusoid, which we assume is the weakest sinusoid with SNR equals to SNR_{min}, we get

$$[\hat{\mathbf{f}}]_k = b_m\mathbf{u}_k^H\mathbf{V}_{\sigma_s}\mathbf{v}(\omega_m) + \sum_{j\in[K]\setminus m}\Big(b_j\mathbf{u}_k^H\mathbf{V}_{\sigma_s}\mathbf{v}(\omega_j)\Big) + \mathbf{u}_k^H\mathbf{V}_{\sigma_s}\mathbf{n}. \tag{2}$$

Since $[\hat{\mathbf{f}}]_k$ is a linear combination of $b_i, [\mathbf{n}]_j, i \in [K], j \in [N]$, it holds that $[\hat{\mathbf{f}}]_k \sim \mathcal{CN}(0, \sigma_{fk}^2)$, where

$$\sigma_{fk}^2 = \sigma_{bm}^2\alpha(k, \sigma_s, \omega_m) + \sum_{j\in[K]\setminus m}\big(\sigma_{bj}^2\alpha(k, \sigma_s, \omega_j)\big) + \sigma_n^2\beta(\sigma_s),$$

and $\alpha(k, \sigma_s, \omega) = |\mathbf{u}_k^H\mathbf{V}_{\sigma_s}\mathbf{v}(\omega)|^2$, $\beta(\sigma_s) = \|\overline{\mathbf{W}}\mathbf{P}_{\sigma_s}\mathbf{w}\|^2$. It is easy to see that σ_{fk}^2 is a summation of weighted variances of each signal and noise components.

We now investigate the p_{th} bin where ω_m is mapped to. The permutation modularly dilates the highest gridded peak of ω_m, i.e.,

$\lfloor \frac{\omega_m}{\Delta\omega_N} \rfloor$, $\Delta\omega_N = 2\pi/N$ by σ_s, while the following aliasing and FFT stages rescale the frequency location by B/N. Thus, we can localize p as $p(\omega_m, \sigma_s) = \lfloor \frac{B}{N}[\sigma_s \lfloor \frac{\omega_m}{\Delta\omega_N} \rfloor]_N \rfloor$. Since the signal is sparse in the frequency domain, we assume that only ω_m maps to bin p, and the side-lobes (leakage) are far below the noise level, owning to the two stages of windowing, which attenuates the leakage down to a desired level. Thus, the effect of leakage from other sinusoids can be ignored. Thus, we can approximate the variance of $[\hat{\mathbf{f}}]_p$ as $\sigma_{fp}^2 \approx \sigma_{bm}^2 \alpha(p, \sigma_s, \omega_m) + \sigma_n^2 \beta(\sigma_s)$. The bin $u \in [B]$, to which no significant frequency is mapped, contains only noise, and the corresponding variance for $[\hat{\mathbf{f}}]_u$ is $\sigma_{fu}^2 \approx \sigma_n^2 \beta(\sigma_s)$. Hence, the hypothesis test for the first stage detection on $[\hat{\mathbf{f}}]_j, j \in [B]$ is formulated as

- $\mathcal{H}0$: no significant frequency is mapped to bin j;
- $\mathcal{H}1$: at least one significant frequency is mapped to bin j with the worst case SNR equals to SNR_{min}.

By applying a likelihood ratio test (LRT) on each bin in $[B]$, we establish the relationship of SNR_{min}, threshold γ, expected probability of detection $\bar{P}_d$ and false alarm rate $\bar{P}_{fa}$ in the first stage of detection. Specifically,

$$\bar{P}_{fa} = e^{-\frac{\gamma}{\sigma_n^2 \bar{\beta}}}, \quad \bar{P}_d(\omega_m) = \bar{P}_{fa}^{\frac{\bar{\beta}}{\bar{\alpha}(p,\omega_m) SNR_{min} + \bar{\beta}}}, \tag{3}$$

where $\bar{P}_d(\omega_m) = \mathbb{E}\{\tilde{P}_d(\sigma_s, \omega_m)\}$, $\bar{P}_{fa} = \mathbb{E}\{\tilde{P}_{fa}(\sigma_s)\}$, $\bar{\alpha}(p, \omega_m) = \mathbb{E}\{\tilde{\alpha}(p, \omega_m, \sigma_s)\}$, and $\bar{\beta} = \mathbb{E}\{\tilde{\beta}(\sigma_s)\}$. The expectation is taken with respect to the permutation σ_s, and $\tilde{P}_d(\sigma_s, \omega_m)\}$, $\tilde{P}_{fa}(\sigma_s)$ are the probability of detection for the weakest frequency and false alarm rate on each bin, respectively. The derivation of (3) is straightforward, since the two competing statistics are all zero-mean circularly symmetric Gaussian scalars with different variances.

Second Stage Detection: Let $\mathbf{c}_{\sigma_s} \in \{0,1\}^B$ denote the output of the first stage detection for the s_{th} segment, with permutation parameter σ_s. Each entry in $\mathbf{c}_{\sigma_s}$ is a Bernoulli random variable, i.e., for $j \in [B]$, $[\mathbf{c}_{\sigma_s}]_j \sim \text{Bernoulli}\left(\tilde{P}_{fa}(\sigma_s)\right)$ under $\mathcal{H}0$, and $[\mathbf{c}_{\sigma_s}]_j \sim \text{Bernoulli}\left(\tilde{P}_d(\omega_m, \sigma_s)\right)$ under $\mathcal{H}1$. Note that under $\mathcal{H}1$, we assume that $[\mathbf{c}_{\sigma_s}]_j$ corresponds to the weakest sinusoid. For the other $K-1$ co-existing sinusoids, since their SNR may be greater than SNR_{min}, their probability of detection may also be greater than $\tilde{P}_d(\omega_m, \sigma_s)$.

The reverse mapping stage hashes the B-dimensional $\mathbf{c}_{\sigma_s}$ back to the N-dimensional $\mathbf{a}_{\sigma_s}$. After accumulation of T iterations, each entry in the accumulated output is summation of T Bernoulli variables with different success rate. On defining $\bar{\mathbf{a}}$ as the accumulated output, then for its $i_{th}, i \in [N]$ entry, we have

$$[\bar{\mathbf{a}}]_i = \sum_{i \in \mathcal{R}(j,\sigma_s), s \in [T]} [\mathbf{c}_{\sigma_s}]_j, \tag{4}$$

where $\mathcal{R}(j, \sigma_s)$ denotes for the reverse mapping of the j_{th} entry in $[B]$ to the i_{th} entry in $[N]$, with the permutation parameter σ_s. Note that in (4), each term inside the sum corresponds to a different segment, i.e., $[\mathbf{c}_{\sigma_s}]_j$ is from the s_{th} segment. Since σ_s is drawn randomly for each segment, j may take different values, and relates to i via $\mathcal{R}(j, \sigma_s)$.

Now, the hypothesis test for the second stage detection on $[\bar{\mathbf{a}}]_i, i \in [N]$ is formulated as

- $\overline{\mathcal{H}}0$: no significant frequency exists.
- $\overline{\mathcal{H}}1$: there exists a significant frequency, whose SNR is at least SNR_{min}.

Under both hypothesis, as $T \to \infty$, $[\bar{\mathbf{a}}]_i$ follows a normal distribution. Specifically, under $\overline{\mathcal{H}}1$,

$$[\bar{\mathbf{a}}]_i \sim N(\mu_{a1}(\omega_m), \sigma_{a1}^2(\omega_m)), \tag{5}$$

where $\mu_{a1}(\omega_m) = T\bar{P}_d(\omega_m), \sigma_{a1}^2(\omega_m) \le T\bar{P}_d(\omega_m)(1-\bar{P}_d(\omega_m))$. Under $\overline{\mathcal{H}}0$,

$$[\bar{\mathbf{a}}]_i \sim N(\mu_{a0}(\omega_m), \sigma_{a0}^2(\omega_m)), \tag{6}$$

where $\mu_{a0}(\omega_m) = F\eta_p\bar{P}_d(\omega_m) + (T-F)\bar{P}_{fa}$, $\sigma_{a0}^2(\omega_m) \le F\eta_p\bar{P}_d(\omega_m)(1-\eta_p\bar{P}_d(\omega_m)) + (T-F)\bar{P}_{fa}(1-\bar{P}_{fa})$, and $F = \frac{TK\eta_m}{B}$, where η_m is the broadening parameter of the pre-permutation window $\mathbf{w}$. $\eta_p \in [1, \frac{1}{\bar{P}_d(\omega_m)}]$ is a calibration parameter of the probability of detection for the other $K-1$ co-existing sinusoids.

To derive (5) and (6), we first explore the properties of mapping and reverse mapping, which guarantees that a mapped location can be reverse mapped (with ambiguities), and two distinct locations with the same permutation parameter are mapped to distinct locations. With these properties, we notice that under $\overline{\mathcal{H}}1$, assuming that $[\bar{\mathbf{a}}]_i$ corresponds to the weakest sinusoid, each term inside the sum of (4) is distributed as $[\mathbf{c}_{\sigma_s}]_j \sim \text{Bernoulli}\left(\tilde{P}_d(\omega_m, \sigma_s)\right), s \in [T]$. Under $\overline{\mathcal{H}}0$, the sum in (4) is composed of F Bernoulli variables with success rate $\tilde{P}_d(\omega_m, \sigma_s)$, and $T-F$ Bernoulli variables with success rate $\tilde{P}_{fa}(\sigma_s)$. The parameter η_m reflects the fact that sparsity is affected by the pre-permutation windowing. And since we assume that $\mathbf{v}(\omega_m)$ has the minimum SNR, i.e., SNR_{min}, other sinusoids with higher SNR will have larger $\bar{P}_d$. Hence, we multiply $\bar{P}_d(\omega_m)$ with η_p to calibrate the success rate of $[\mathbf{c}_{\sigma_s}]_j$ under $\mathcal{H}1$. Next, by applying Lyapunov Central Limit Theory [14], it is easy to show that $[\bar{\mathbf{a}}]_i$ approaches normal distribution asymptotically.

The Optimization Problem: Based on the above discussion, the optimal threshold design can be solved by the following optimization problem, i.e.,

$$\begin{aligned} &Minimize_{\{\mu, \bar{P}_{fa}, \bar{P}_d\}} \quad SNR_{min} \\ &s.t. \;\; \bar{P}_d(\omega_m) = \bar{P}_{fa}^{\frac{\bar{\beta}}{\bar{\alpha}(p,\omega_m) SNR_{min} + \bar{\beta}}}, P_{fa} = \int_{\mu}^{\infty} g_{a_0}(u) du \\ &P_d = \int_{\mu}^{\infty} g_{a_1}(u) du, \mu \in [T], 0 \le \bar{P}_{fa} \le 1, 0 \le \bar{P}_d \le 1, \end{aligned} \tag{7}$$

where $g_{a_0}(u), g_{a_1}(u)$ are the asymptotic PDF of $[\bar{\mathbf{a}}]_i$ (which corresponds to the weakest sinusoid) under $\overline{\mathcal{H}}0$ and $\overline{\mathcal{H}}1$, respectively. We take the upper bounds of the variances in both distributions. It can be shown numerically that the actual variances are close to their upper bounds; the proof is omitted due to lack of space. Since both of them are Normal distributions, with fixed threshold μ, we can solve for $\bar{P}_d(\omega_m), \bar{P}_{fa}$, and then compute the SNR_{min}. By exhaustive search for $\mu \in [T]$, the minimum worst case SNR, i.e., SNR_{min}^*, can be found, and the corresponding μ^* is the optimal threshold for the second stage of detection. The optimal threshold for the first stage of detection, i.e., γ^*, can thus be calculated via (3).

Remark 1. For the second stage detection, the LRT is obtained based on two Normal distributions. The test statistic under $\overline{\mathcal{H}}1$ is "stable", because it only depends on $\bar{P}_d(\omega_m)$. On the other hand, under $\overline{\mathcal{H}}0$, the distribution depends on the number of co-existing sinusoids, as well as on each sinusoid's SNR. A larger K and a higher SNR will "push" the distribution under $\overline{\mathcal{H}}0$ closer to the distribution under $\overline{\mathcal{H}}1$, hence causing a degradation of the detection performance. In order to compensate for this, a larger SNR_{min} is required.

Remark 2. In (6), we set a parameter η_p to calibrate the distribution of $[\bar{\mathbf{a}}]_i$ under $\overline{\mathcal{H}}0$. By setting η_p as 1 or $\frac{1}{P_d(\omega_m)}$, we can get respectively the lower and upper bound of SNR^*_{min} for the variation of SNR of other co-existing sinusoids. If K is the maximum budget of signal sparsity, the optimal thresholds found by solving (7) provides the optimal thresholds for the worst case. If the actual signal sparsity were less than K, P_{fa} would be lower than the expected value, while P_d would be unchanged according to Remark 1.

4. NUMERICAL RESULTS

In this section, we verify our theoretical findings and compare to the detection method in the SFT via simulations. We use the following parameters: $N = 1024, T = 50, B = 64, \eta_m = 3, P_d = 0.9, P_{fa} = 10^{-6}, \omega_m = 64.5\Delta\omega_N \approx 0.4$. As pre-permutation window, we adopt a Dolph-Chebyshev window with $40dB$ attenuation. We use the same window to convolve with a boxcar whose bandwidth is $1/B$ in the frequency domain to create the flat window.

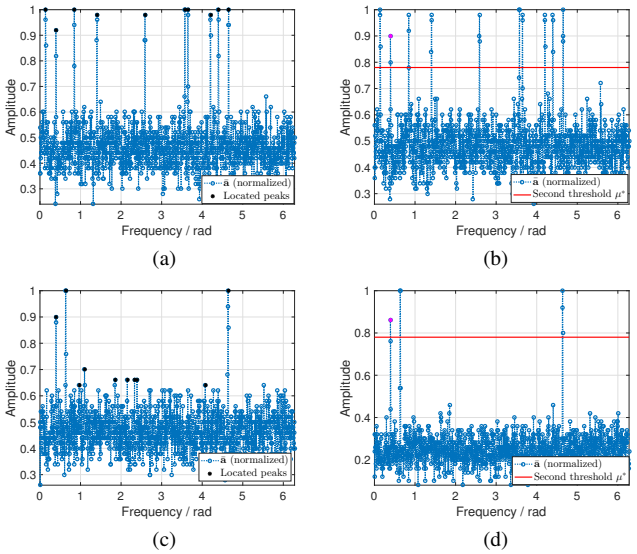

Fig. 1. **Detection with Unknown Signal Sparsity.** The optimized thresholding approach is robust to the unknown sparsity. The magenta dots in (b) and (c) correspond to ω_m. (a) Second stage detection by counting peaks when $K = 10$. (b) Second stage detection by thresholding when $K = 10$. (c) Second stage detection by counting 10 peaks, while the true sparsity is 3. (d) Second stage detection with the threshold optimized for $K = 10$, while the true sparsity is 3.

To show that the proposed method is robust to the signal sparsity, we do not assume that we know K, instead, a guess for K, i.e., $K = 10$ is used. Fig. 1 shows the detection with counting-K-peaks and optimal thresholding, respectively. When the guess for K is correct, both methods can locate all the significant frequencies, as shown in Fig. 1 (a) and (b). However, when the true sparsity is 3, the counting-K-peaks will result into many false alarms, while the optimal thresholding method, which is optimal for $K = 10$, can still recover the correct significant frequencies (see Fig. 1 (d)). Moreover, compared to (b), it yields similar P_d for the weakest signal but better P_{fa}, since the noise level is much lower than expected.

For the high dimensional setting, we use the radar example in [6], in which the signal from 4 targets lies in a 3-D range, velocity and direction of arrival (DOA) space after applying DFT in each dimension. The significant frequencies are not assumed on-grid points in that 3-D space. Moreover, the signal has a dynamic range of $20dB$. In this scenario, it is difficult for the SFT to determine the number of peaks to be counted since the leakage from stronger targets destroys the sparsity of the signal. To recover all the targets, the number of peaks to be counted needs to be gradually increased until all the targets are observed in the recovered 3-D space [6]. However, in reality, we do not know how many targets to recover, while SNR_{min}, which corresponds to the target with smallest radar cross section (RCS) at the largest unambiguous range, is easy to determine at the testing phase of the radar. With knowledge of SNR_{min} and a worst case signal sparsity K, we can find the detection thresholds for the RSFT with a fixed P_{fa} and maximized P_d (by recasting (7) for maximizing P_d with fixed SNR_{min} and P_{fa}). Fig. 2 shows our detection results where it can be seen that the RSFT can recover all the targets when the thresholds are determined by our optimization procedure for $SNR_{min} = -20dB, P_{fa} = 10^{-6}, K = 10$, while the SFT recovers only the stronger targets (Targets 1 and 2) and their side-lobes due to high level of leakage and incorrect knowledge of the signal sparsity. We should emphasize that in a real system, determining the number of peaks to be counted for the SFT-based method lacks a theoretical foundation, while the thresholding approach in the RSFT is consistent with the conventional FFT-based processing, both of which are based on the NP criterion.

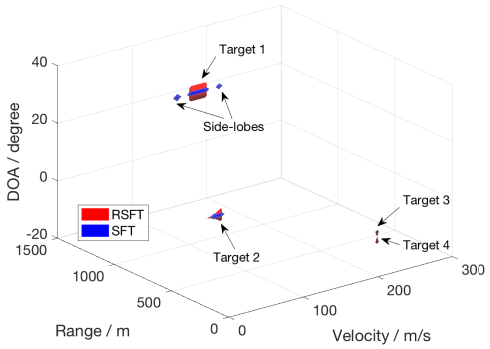

Fig. 2. **Radar Targets Reconstruction via 3-D SFT and the RSFT.** The SFT-based processing recovers the stronger targets (Targets 1 and 2) and their side-lobes by counting 50 frequency peaks in the first stage of detection, while the RSFT-based method recovers the main-lobes from all the targets with our detection framework.

5. CONCLUSION

In this paper, we have addressed the detection problem in the RSFT. By incorporating the NP criterion, we have shown that the asymptotically optimal thresholding can be achieved in the RSFT on the assumed signal model. Moreover, it has been shown that our detection framework is robust to the incomplete knowledge of signal sparsity, and thus is more applicable to the real world problems. Some of the interesting properties of the RSFT have also been revealed by our analysis, such as the performance of detection not only relies on the frequency under examination, but also depends on other co-existing significant frequencies, which is very different from the traditional FFT-based processing.

6. REFERENCES

[1] 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*. ACM, 2012, pp. 563–578.

[2] 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*. 2012, SODA '12, pp. 1183–1194, SIAM.

[3] Haitham Hassanieh, Fadel Adib, Dina Katabi, and Piotr Indyk, "Faster gps via the sparse fourier transform," in *Proceedings of the 18th annual international conference on Mobile computing and networking*. ACM, 2012, pp. 353–364.

[4] Haitham Hassanieh, Lixin Shi, Omid Abari, Ezz Hamed, and Dina Katabi, "Ghz-wide sensing and decoding using the sparse fourier transform," in *INFOCOM, 2014 Proceedings IEEE*. IEEE, 2014, pp. 2256–2264.

[5] Lixin Shi, Haitham Hassanieh, Abe Davis, Dina Katabi, and Fredo Durand, "Light field reconstruction using sparsity in the continuous fourier domain," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 1, pp. 12, 2014.

[6] Shaogang Wang, Vishal M. Patel, and Athina Petropulu, "RSFT: a realistic high dimensional sparse fourier transform and its application in radar signal processing," in *Milcom 2016 Track 1 - Waveforms and Signal Processing.*, Baltimore, USA, Nov. 2016, Avaliable at: http://www.rci.rutgers.edu/ vmp93/Conference_pub/ MILCOM2016_SFT.pdf.

[7] André Rauh and Gonzalo R Arce, "Sparse 2d fast fourier transform," *Proceedings of the 10th International Conference on Sampling Theory and Applications*, 2013.

[8] Badih Ghazi, Haitham Hassanieh, Piotr Indyk, Dina Katabi, Erik Price, and Lixin Shi, "Sample-optimal average-case sparse fourier transform in two dimensions," in *Communication, Control, and Computing (Allerton), 2013 51st Annual Allerton Conference on*. IEEE, 2013, pp. 1258–1265.

[9] Frank Ong, Sameer Pawar, and Kannan Ramchandran, "Fast and efficient sparse 2d discrete fourier transform using sparse-graph codes," *arXiv preprint arXiv:1509.05849*, 2015.

[10] 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*. IEEE, 2014, pp. 514–523.

[11] Haitham Hassanieh, Maxim Mayzel, Lixin Shi, Dina Katabi, and Vladislav Yu Orekhov, "Fast multi-dimensional nmr acquisition and processing using the sparse fft," *Journal of Biomolecular NMR*, pp. 1–11, 2015.

[12] S. Wang, V. M. Patel, and A. Petropulu, "An Efficient High-Dimensional Sparse Fourier Transform," *ArXiv e-prints*, Oct. 2016.

[13] Petre Stoica and Randolph L Moses, *Spectral analysis of signals*, pp. 49–50, Pearson/Prentice Hall Upper Saddle River, NJ, 2005.

[14] Robert B Ash and Catherine Doleans-Dade, *Probability and measure theory*, p. 309, Academic Press, 2000.