Finite Multi-Coset Sampling and Sparse Arrays

Yuval Kochman and Gregory W. Wornell
Dept. Electrical Engineering and Computer Science
Massachusetts Institute of Technology
Cambridge, MA 02139, USA
Email: {yuvalko,gww}@mit.edu

Abstract—**Signals with sparse but otherwise unknown frequency content are well-represented by multi-coset samples, and efficient algorithms can be used to recover the underlying sparsity structure. While such sampling is usually analyzed over a sampling interval sufficiently large that edge effects can be ignored, in this work we develop how to take into account finite-window effects in system design. Such considerations are particularly important in the context of antenna arrays, and we analyze the associated redundancy. Additionally, we describe an efficient MIMO radar implementation of multi-coset arrays. As an example application of our results, we develop a natural two-stage architecture for direction-of-arrival estimation in sparse environments using a multi-coset array over the available aperture.**

I. INTRODUCTION

The famous Nyquist-Shannon sampling theorem [13] states that any signal of finite bandwidth B may be reconstructed from samples at uniform interval $1/2B$. However, this condition is by no means necessary. In [8], Landau shows that an average interval of $1/2\tilde{B}$ is indeed necessary for reconstruction from any sampling pattern (uniform or not), where $\tilde{B}$ is the Lebesgue measure of the spectral support of the signal. In general, $\tilde{B}$ may be much smaller than B; one may easily construct examples in which the Landau bound is achievable, even using uniform sampling. Over the years, various works have considered sampling patterns and reconstruction algorithms that approach the Landau bound. More recently, the considerable attention devoted to problems of compressive sensing of late has, in turn, has renewed interest in such sampling problems.

Feng and Bresler [6] show that for a class of sources (where the support is the union of frequency bands of equal width) the Landau rate may be achieved by *multi-coset sampling*, i.e., the union of a finite number of low-

rate uniform sampling patterns. Furthermore, it is shown that there exist inter-coset spacings that allow reconstruction for *any* signal in the class; thus, sampling may be *blind*, i.e., the locations of the support intervals need not be known in advance. If they are known at the time of reconstruction, then the Landau rate is achievable. If the support is not known even at that time, it may be recovered during a first stage of the reconstruction. For "most" signals this can still be performed at the Landau rate, and for the worst-case signal there is a factor-of-two rate penalty. Mishali and Eldar [11] use compressive sensing ideas [10] to suggest new spectral support recovery algorithms, as well as to prove that no sampling scheme can have worst-case performance better than the multi-coset one. Beyond multi-cosets, other techniques have been recently proposed for the same problem; see e.g. [12], [16].

In the notion of sampling rate, it is implicit that sampling takes place indefinitely, or at least over a window that includes sufficiently many samples that edge effects may be ignored. This is not a good assumption in all applications. Most notably, sampling and spectral estimation are integral to the array processing problems of imaging and source localization. In fact, for a fixed system operating wavelength and under a far-field assumption, the array plane and measurement azimuth are equivalent to the time and frequency domains, respectively. Thus, a scene in which no signal is arriving from most directions is equivalent to a spectrally-sparse signal, and array element placement may be seen as sampling (where the "Nyquist" spacing is half the wavelength). When designing an array, one is restricted both by the number of elements and by the aperture size.

The rest of the paper is organized as follows. In Section II we present the signal model of interest in the paper, along with some basic terminology we require. Section III contains the main results for signal reconstruction from sparse finite-window sampling. In Sec-

tion IV we connect these results with array processing and especially MIMO radar. Finally, in Section V we apply the results to the problem of direction-of-arrival estimation.

II. SIGNAL MODEL AND TERMINOLOGY

Let $x[n]$ be a complex-valued signal of interest with discrete-time Fourier transform $X(e^{j2\pi f})$. For any integers $K \leq M$. We say that $x[n]$ is (K, M)-sparse if $X(e^{j2\pi f}) = 0$ for all $f \notin \mathcal{F}$, where

$$\mathcal{F} = \left\{ \bigcap_{k=1}^{K} \left(\overline{f_k} - \frac{1}{2M}, \overline{f_k} + \frac{1}{2M} \right] \right\} \bmod (0, 1] \quad (1)$$

for some band centers $\{\overline{f_k}\}$. For our developments in this paper, we also need a slotted notion of sparsity. Define M_s (overlapping) frequency windows:

$$\mathcal{F}_m = \left(\frac{m-1}{M_s}, \frac{m+1}{M_s} \right] \bmod (0, 1], m = 0, \ldots, M_s - 1.$$

Let $\mathbf{q}$ be the *slotted support vector*, i.e., it contains the indices of windows $\mathcal{F}_m$ that intersect with the spectral support $\mathcal{F}$ of $x[n]$. If $\mathbf{q}$ has dimension K_s, we say that $x[n]$ has (K_s, M_s)-sparse slotted support.

Let

$$\lambda = \frac{M_s}{M} \quad (2)$$

be the band expansion factor. It is straightforward to show that any (K, M) sparse signal has (K_s, M_s)-sparse slotted support with

$$\frac{K_s}{M_s} \geq \frac{K}{M} \left(1 + \frac{2}{\lfloor \lambda \rfloor} \right). \quad (3)$$

For instance, substituting $\lambda = 2$, any (K, M)-sparse signal has a $(4K, 2M)$-sparse slotted support.

We are interested in reconstructing the sparse signal $x[n]$ over some window $|n| \leq W/2$ from a subset of its samples.[1] A reconstruction $\hat{x}[n]$ is ϵ-exact over W if

$$\sum_{-W/2}^{W/2} |\hat{x}[n] - x[n]|^2 \leq \epsilon \cdot \sum_{-W/2}^{W/2} |x[n]|^2. \quad (4)$$

We use some S sampling instances $n_1 \leq n_2 \leq \ldots \leq n_S$ inside an interval $|n| \leq \tilde{W}/2$. Let the sampling sparsity be the ratio between the number of samples used, and the number of samples in the sampling window:

$$\rho = \frac{S}{\tilde{W}}. \quad (5)$$

[1]We note that the problem of reconstructing a bandlimited signal from its Nyquist-spaced samples over a finite window is comparatively better understood. Indeed, this problem has a long history of attention, largely in a continuous-time setting.

Any sampling scheme may thus be characterized by the pair $(\tilde{W}, \rho)$. Our goal is to characterize the pairs needed for given (W, ϵ).

We choose to employ multi-coset sampling. In particular, $x[n]$ can be expressed in terms of its cosets

$$x^{(p)}[n] = x[nM_s + p], \ p = 0, \ldots, M_s - 1. \quad (6)$$

Let the cosets pattern be the L integers $0 \leq p_1 < \ldots < p_L \leq M_s - 1$, such that the sparse sampling pattern includes all of the cosets indexed by elements of $\mathbf{p}$. For a window $\tilde{W}$, this is $(\tilde{W}, \rho)$ sampling with

$$\rho = \frac{L}{M_s}. \quad (7)$$

For any slotted support vector $\mathbf{q}$ and cosets pattern $\mathbf{p}$, the *measurement matrix* $\mathbf{A} = \mathbf{A}(\mathbf{p}, \mathbf{q})$, which will be defined later, is an $L \times K_s$ sub-matrix of the M_s-dimensional discrete Fourier transform (DFT) matrix, i.e., its elements are of the form

$$A_{l,k} = e^{j2\pi p_l q_k / M_s}. \quad (8)$$

III. FINITE-WINDOW MULTI-COSET SAMPLING

In the limit of an infinite window, a (K, M)-sparse signal may be reconstructed from samples with sparsity $\rho = K/M$, corresponding to the Landau rate, as long as the spectral support is known at the time of recovery. If it is not known then, there is a set of signals (of zero probability under some mild randomness assumptions) that require a redundancy of factor 2 with respect to the Landau rate [6]. We show below that for a finite window the same is achievable with some additive redundancy $\tilde{W} - W$ which is fixed in W and finite for any accuracy $\epsilon > 0$. Reconstruction is performed using a variant of the time-domain scheme of [6]. In this variant, the decomposition of the signal into frequency bands is replaced by the decomposition into (twice as many) overlapping Nyquist bands, which form a complete orthormal set, and are defined by prototype Nyquist filter responses. The use of Nyquist filters is suggested in [3] in a different context.

In Section III-A we describe the reconstruction scheme for a known slotted support vector $\mathbf{q}$. In Section III-B we analyze the tradeoff between accuracy and redundancy. In Section III-C we address the recovery of the support vector from the samples, and in Section III-D we bring a design example with simulation results.

It should be noted that the signal model considered does not include measurement noise. In the presence of wide-band measurement noise, the signal cannot strictly satisfy a sparsity assumption of the form (1). While noise
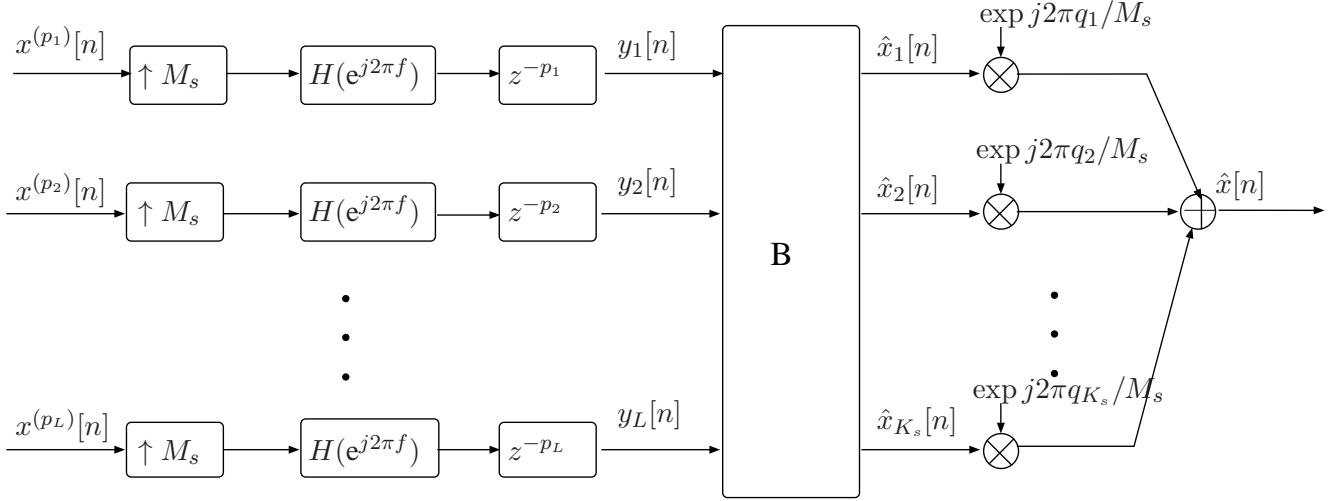
Fig. 1: Reconstruction Scheme.

is not explicitly considered, it may limit the performance of support recovery, as well as be folded into signal bands; see Section III-B and Section III-C in the sequel.

A. Reconstruction for Known Spectral Support

The scheme is depicted in Fig. 1. Each of the sampled cosets is interpolated by factor M_s, using a filter $h[n]$. Further, each filter output is delayed by the coset offset p_l. At each time instant, we form from the outputs $y_l[n]$ an L-dimensional vector $\boldsymbol{y}$. To that vector, we apply the $K_s \times L$ pseudo-inverse matrix

$$\mathbf{B} = (\mathbf{A}^\dagger \mathbf{A})^{-1} \mathbf{A}^\dagger \tag{9}$$

where $\mathbf{A} = \mathbf{A}(\mathbf{p}, \mathbf{q})$ is the measurement matrix (8). The outputs of this matrix are used to form signals, which are finally modulated and summed together.

We constrain the interpolation filter $h[n]$ to be a Nyquist filter of order M_s, i.e., $h[n \cdot M_s] = 0$ for all $n \neq 0$, which in the frequency domain is equivalent to the constraint

$$\sum_{m=0}^{M_s-1} H\left(e^{j2\pi(f-m/M_s)}\right) = 1. \tag{10}$$

As a consequence, for any $x[n]$ we have

$$X(e^{j2\pi f}) = \sum_{m=0}^{M-1} X(e^{j2\pi f}) H\left(e^{j2\pi(f-m/M_s)}\right). \tag{11}$$

Now suppose that $h[n]$ is a low-pass filter with zero response for any $|f| > 1/M_s$.[2] For a signal with

---

[2]This assumption is in contradiction with the finite-window assumption. We will later replace it with a less stringent requirement, which leads to implementable filters at the cost of limitations to the accuracy ϵ.

(K_s, M_s)-sparse slotted support

$$X(e^{j2\pi f}) = \sum_{k=1}^{K_s} X(e^{j2\pi f}) H\left(e^{j2\pi(f-q_k/M_s)}\right)$$
$$\triangleq \sum_{k=1}^{K_s} X_k\left(e^{j2\pi(f-q_k/M_s)}\right). \tag{12}$$

Thus, the signal can be decomposed as the sum of filtered versions of components corresponding to (overlapping) bands. It is apparent from the scheme that if $\hat{x}_k[n]$ equal $x_k[n]$ defined above, then indeed $\hat{x}[n]$ is a perfect reconstruction of $x[n]$.

In order to see why this holds, we use the same conditions on $h[n]$ to decompose the delayed interpolated cosets as

$$Y_l(e^{j2\pi f}) = e^{-j2\pi f p_l/M_s} H(e^{j2\pi f}) \sum_{m=0}^{M_s-1} X\left(e^{j2\pi(f-m/M_s)}\right)$$
$$= \sum_{k=1}^{K_s} A_{l,k} X_k\left(e^{j2\pi(f-q_k/M_s)}\right). \tag{13}$$

It follows that at each time instant,

$$\boldsymbol{y} = \mathbf{A}\mathbf{x}$$

where $\mathbf{x}$ is a K_s dimensional vector taken from the processes $x_k[n]$. For a full-rank $\mathbf{A}$ we thus have that the matrix output at each time instant is:

$$\hat{\mathbf{x}} = \mathbf{B}\boldsymbol{y} = \mathbf{B}\mathbf{A}\mathbf{x} = \mathbf{x}.$$

The rank of $\mathbf{A}$ depends upon the coset pattern $\mathbf{p}$ and the support vector $\mathbf{q}$. A pattern $\mathbf{p}$ of length L, which guarantees full rank for any $\mathbf{q}$ of length K_s, is called a

universal pattern. By [6], such pattern exists as long as $L \geq K_s$.

By taking $L = K_s$ and choosing a large band expansion factor λ (2), the sparsity ρ (7) may arbitrarily approach the Landau rate K/M. The scheme as presented here is still not applicable when considering the samples are only taken from a finite window, but as we will see in the sequel, the replacement of sharp band filters by more gracefully shaped Nyquist filters allows us to approach the infinite-window scheme at the cost of a small increase in the sampling window. However, the sequel also shows that a large λ is undesirable.

B. Redundancy Analysis

Using a finite filter $h[n]$ over a finite data window amounts to suffering from an aliasing effect. We choose the observation window to be

$$\tilde{W} = W + \delta \tag{14}$$

where δ is the length of the filter $h[n]$.[3]

The aliasing effect amounts to an interference term that is added to each term in the summations (12) and (13), with energy proportional to the side-lobe energy of $h[n]$. This is equivalent to having a non-truncated bandlimited filter $\bar{h}[n]$ operating on a signal that is not perfectly bandlimited. The later situation is analyzed in [18], where it is shown that the noise amplification due to passing these interference terms through the pseudo-inverse B can be uniformly bounded for a given pattern $\mathbf{p}$. It follows that the accuracy ϵ (4) can be bounded by[4]

$$\epsilon \leq c(M_s, K_s, \mathbf{p}) \, \mu(\delta) \tag{15}$$

or some constant c, where

$$\mu(\delta) = \frac{\int_{1/2M_s < |f| < 1/2} |H(e^{j2\pi f})|^2 df}{\int_{|f| < 1/2} |H(e^{j2\pi f})|^2 df} \tag{16}$$

is the out-of-band energy.

One of the ways to construct a "good" finite-duration Nyquist filter is to use a window function

$$h[n] = \mathrm{sinc}[nL] \, w[n], \tag{17}$$

where $w[n]$ is a window of length δ with good attenuation for $|f| > 1/2M_s$. The Kaiser window is known to be a good approximation to the optimal

---

[3]We assume a symmetric filter, thus it has $\delta/2$ taps at each side of the main tap; this amounts to a total of δ reconstructed samples for which the filter output is not as it should be according to the scheme, assuring that within W the outputs are indeed valid.

[4]A similar bound also holds for out-of-band measurement noise, which may be folded into the signal bands.

(in terms of out-of-band energy for a given main-lobe width) prolate spheroidal wave function window. The out-of-band energy decays exponentially with the filter length. Specifically, substituting $\Delta \omega = 2\pi/M_s$ in [14, eq. (7.93)], the out-of-band energy satisfies:

$$-10 \log_{10} \mu(\delta) \leq 14.3 \frac{\delta}{M_s} + 8. \tag{18}$$

Combining (14), (15), and (18), we obtain that for a given accuracy ϵ, an estimate of the window redundancy r is

$$r \triangleq \frac{\tilde{W} - W}{M_s} = 0.7 \log_{10} \left(\frac{c(M_s, K_s, \mathbf{p})}{\epsilon} \right) - 0.5. \tag{19}$$

Since the excess window needed is "per-band," choosing a large M_s, while helping the slotted support approach the signal spectral support thus lowering the multiplicative redundancy of ρ, results in higher additive redundancy. Investigating the dependence of the constant c upon the choice of M_s should shed more light on the optimization of the number of bands.

C. Support Recovery

The task of support recovery for the vector $\mathbf{q}$ corresponding to the overlapping bands $\mathcal{F}_m$ (2) is in principle no different than that of recovering the support of non-overlapping bands for the original infinite-window scheme. We therefore briefly summarize existing results. In the reconstruction scheme of Fig. 1, we may obtain the signals $\{y_l[n]\}$ without knowing the support. We therefore look for a $\mathbf{q}$ which allows us to satisfy (13). There is a unique solution if

$$L \geq 2K_s + 1. \tag{20}$$

If the K_s functions $\{X_k(e^{j2\pi f})\}$ form a linearly-independent set, then the condition for a unique solution becomes the more favorable

$$L \geq K_s + 1; \tag{21}$$

see [6], [1]. The condition (21) allows us to approach the Landau rate if a large band expansion factor λ (2) is chosen. Recovery may be performed in the time-domain using, e.g, MUSIC-like algorithms [6], [9], or in the frequency-domain using, e.g., basis pursuit [11].

Restricting the sampling window to a finite $\tilde{W}$ may affect the recovery process in the following ways. The out-of-band energy discussed above means that (13) is only approximately satisfied. However, in practice, even when using a long sampling window, wide-band noise is present. The recovery algorithms above have
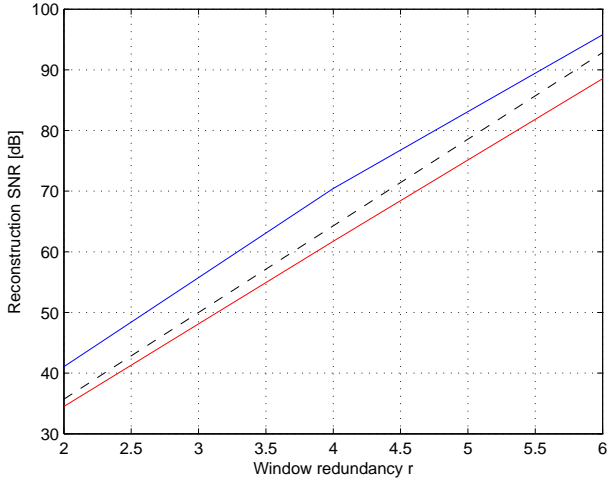
Fig. 2: Simulation results: reconstruction SNR $1/\epsilon$ as a function of the window redundancy. Upper solid curve: average SNR; lower solid curve: worst-case SNR; dashed curve: SNR predicted by (19) with $c = 1$.

good robustness properties, thus taking large enough $\tilde{W}$ such that the out-of-band components are not stronger than the noise suffices. Beyond this, the finite window has another fundamental efect. Linear dependence in the set $\{X_k(e^{j2\pi f})$ involves dependence at each time instance. When limiting the window, one has to consider the probability that even statistically-independent continuous-amplitude signals look "almost linearly dependent," limiting the performance of recovery using the more favorable cobdition (21), in the presence of some noise. This is an additional argument against using large δ, beyond the window redundancy.

D. Simulation Results

We consider a $(2, 7)$-sparse signal. If one were to use $M_s = 2$, by (3) the Nyquist-bank scheme requires $K_s = 6$, impeding any sparsity. We therefore take $\delta = 3$, resulting in $K_s = 10$ and $M_s = 21$. According to (21), we therefore take $L = 11$. The cosets pattern chosen is

$$\mathbf{p} = [1, 4, 6, 8, 9, 10, 11, 15, 17, 18, 20].$$

Fig. 2 shows the reconstruction SNR $1/\epsilon$ in dB as a function of the window redundancy (19). For each of the values $r = (2, 4, 6)$, 1000 trials were run, with a random stationary Gaussian signal bandlimited according to band centers $\bar{f}_1, \bar{f}_2$ uniformly and independently drawn at each trial. The average and worst-case SNR over these trials were measured. It can be appreciated both SNRs follows closely (19). The average performance is even better

than that of (19) with $c = 1$, which may be explained by the bound on the Kaiser out-of-band energy being rather loose (reflecting the attenuation of the highest side-lobe). A realistic estimate of the coefficient c is the gap between the average and worst-case performance, thus $c(10, 21, \mathbf{p}) \cong 5$.

IV. APPLICATION TO ARRAY PROCESSING AND MIMO RADAR

As discussed in the Section I, one important application of finite multi-coset sampling is array design. The number of samples is translated to the number of array elements that need to be placed, and the observation window translates to the required aperture. To make this concrete, assume some linear one-dimensional array with elements placed at a subset of the locations

$$t = n \cdot \frac{\lambda}{2}, \ |n| \leq \frac{\tilde{W}}{2}, \tag{22}$$

where λ is the length of waves hitting the array (assumed to be fixed, i.e., all radiation is monochromatic). Also assume that targets are in the far field and the scene is only described by a single angle (azimuth) θ; over these angles are some targets that transmit (or reflect) a fixed (over time) signal, which after the attenuation between the target and the array has complex amplitude $X(\theta)$. Under these conditions, the following Fourier relation holds:

$$X[n] \Leftrightarrow X\left(e^{j\pi \sin(\theta)}\right) \triangleq X(e^{j2\pi f}). \tag{23}$$

A sparse scene would be one where all of the incoming waves arrive from clustered angles θ such that $f = \sin(\theta)/2$ satisfies the sparsity condition (1). If the scene is known to be such, one may design a multi-coset array of sparsity ρ and aperture $\tilde{W}$ and use the reconstruction scheme to reconstruct (to accuracy ϵ) the signal of all the $\lambda/2$-spaced elements on an array of aperture W. To that data, any beamforming or other processing may be applied.

The sparsity assumption is more easily justified in the context of *radar*: incoming signals may be first distributed into range/doppler cells, and then only returns within the same cell need to be processed together by the sparse recovery algorithm. See, e.g., [15] for a recent tutorial on the application of compressive sensing ideas to radar.

The multi-coset structure is especially appealing in the context of MIMO radar. One of the most important features of a MIMO radar is the ability to create a *virtual array* using different TX and RX arrays, one of

them short and dense and the other wide and sparse. The total number of array elements is much smaller (order of square root) of the number of elements in the virtual array; see e.g. [7]. A multi-coset virtual array may be expressed as the convolution of an array of coset leaders (at locations $\mathbf{p}$) and a sparse uniform array with element spacing M_s. Thus, the multi-coset concept may be integrated in MIMO radar in order to further reduce the number of elements: one array will have $L \geq K_s + 1$ elements over aperture $M_s \lambda/2$, while the other will have $\tilde{W}/M_s$ elements over aperture $\tilde{W} \lambda/2$. The application of compressive sensing to MIMO radar has been suggested in some recent works (e.g., [2]), but to the best of our knowledge, no specific architecture was suggested.

V. DIRECTION-OF-ARRIVAL ESTIMATION

We now replace the spectral support model (1) by the assumption that the signal of interest is the sum of K complex exponentials and wideband noise:

$$x[n] = \sum_{k=1}^{K} a_k e^{j2\pi f_k n} + z[n], \qquad (24)$$

where $z[n]$ is additive white Gaussian noise. In the array setting, this would correspond to K plane waves coming from different angles (and may be generated from far-field point sources), with independent sensor noise. Many estimation strategies have been developed for this important problem, each offering different accuracy and resolution performance; see, e.g., [17]. In the context of this work, we ask what are the required number of samples and observation window in order to assure some target performance level is met.

We suggest a two-stage strategy. In the first stage, an ϵ-accurate reconstruction is obtained over a window W, and in the second, any direction-of-arrival estimation algorithm is applied to the signal over that window. The resulting performance is set by the algorithm. Care should be taken when evaluating it, since the ϵ-"noise" reflects aliasing, which is far from being white or additive.

For the first stage, note that without noise, $x[n]$ has $(2K, M_s)$-sparse slotted support for any M_s, thus one may use multi-coset sampling with the reconstruction scheme described in Section III-A. The larger M_s is, the sparser the array becomes, and in the limit of large W the optimal M_s is unbounded, which corresponds to the fact that the point-source scene is "infinitely sparse". however for a finite observation window, large M_s causes aliasing and limits the performance of support recovery, see Section III.

When choosing the number of cosets L, it may not be safe to assume independent bands and use (21). The reason is, that as for the model (24) two bands will seem dependent if the exponents appear at the same frequency, relative to the band, to the finite resolution offered by the aperture. However, the effect of the same exponent on two bands does yield independent functions, as the spectral lines appear $1/M_s$ apart. Consequently, at the worst case there are K dependent functions, and the condition is:

$$L \geq 3K + 1 = \frac{3K_s}{2} + 1. \qquad (25)$$

It is interesting to compare this approach with the recent work by Duarte and Baraniuk [5], where it is noted that straightforward application of compressive sensing methods to the DFT matrix of the observations yields poor results (see also [4] for a similar observation). Such work assumes random Gaussian measurement matrices, so that the number of measurements is not directly related to the number of time samples. However, replacing the measurements by samples taken from $x[n]$, the algorithm of [5] takes the following form.[5] The samples of $x[n]$ over the observation window are partitioned into existing samples and missing samples.

1) Initialize the missing samples to zero.
2) Form the "measurements" vector from the existing and missing samples.
3) Using this vector, estimate the signal frequencies.
4) Set the value of the missing samples according to the estimated frequencies.
5) Repeat steps 2-4 until convergence.

For the estimation in Step 3, it is suggested to use either traditional beam-forming, or the root-MUSIC algorithm. The problem with this algorithm, is that at first it attempts to reconstruct the missing samples using the parametric (sum of exponentials) model. In cases where the parameter estimation is hard (e.g. some exponentials are at very close frequencies), this might fail. The approach presented above, on the other hand, reconstructs the missing samples using a much milder model of frequency bands, and only later, equipped with all the samples, it uses the discrete-exponent assumption.

In order to demonstrate the advantage of the two-stage approach over the iterative algorithm of [5], we consider an example with two point sources, contaminated by white measurement noise with SNR of 20 dB. We draw the direction of one source at random, but set the angle

---

[5]Indeed, simulations show that replacing the random measurement matrices by random sampling times improves the performance.
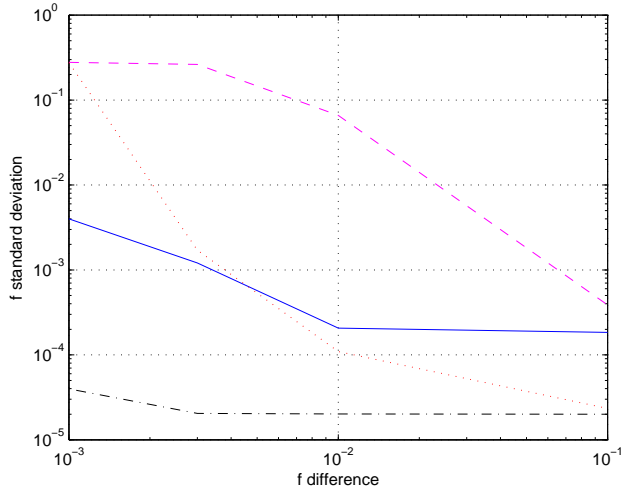
Fig. 3: Simulation results: direction-of-arrival estimation standard deviation as a function of angular difference. Ssolid line: two-stage algorithm; dashed line: iterative algorithm; dash-dotted and dotted lines: full-aperture and short uniform arrays, respectively.

between sources to a deterministic value. We measure the accuracy of angle estimation as a function of this angular distance.

According to (25), we must have $K_s = 7$. We choose $M_s = 50$, and an array of 10 periods. Thus, the array consists of 70 elements spread over an aperture $\tilde{W} = 500$. We compare the performance of the two-stage algorithm to the iterative algorithm [5] applied to randomly-selected 70 elements over the aperture. We also present two reference schemes: one having a full array of the aperture $\tilde{W} = 500$, and the other with aperture 70. In all cases, root-MUSIC estimation is applied.

The results are depicted in Fig. 3. It can be appreciated that the two-stage algorithm is fairly robust with respect to the change of angular difference. In comparison, placing the same number of elements close together yields better results when the targets are far apart, but poor resolution. The iterative algorithm does not perform as well, especially when the angular difference is small. While these results show the potential of our approach, they also show that it should still be improved in order to guarantee good performance for far-spread targets.

REFERENCES

[1] Y. Bresler. Spectrum-blind sampling and compressive sensing for continuous-index signals. In *Information Theory and Applications Workshop, 2008*, pages 547–554, 2008.

[2] C.-Y. Chen and P.P. Vaidyanathan. Compressed sensing in MIMO radar. In *42nd Asilomar Conference on Signals, Systems and Computers*, pages 41–44, 2008.

[3] Y. Chen, M. Mishali, Y. C. Eldar, and A. O. Hero. Modulated wideband converter with non-ideal lowpass filters. In *IEEE International Conference on Acoustics Speech and Signal Processing (ICASSP) 2010*, pages 3630–3633, 2010.

[4] Y. Chi, L. Scharf, A. Pezeshki, and R. Calderbank. The sensitivity to basis mismatch of compressed sensing in spectrum analysis and beamforming. In *Workshop on Defense Applications of Signal Processing (DASP), Lihue, HI*, October 2009.

[5] M. F. Duarte and R. G. Baraniuk. Spectral compressive sensing. http://www.math.princeton.edu/ mduarte/images/SCS-TSP.pdf, 2010.

[6] P. Feng and Y. Bresler. Spectrum-blind minimum-rate sampling and reconstruction of multiband signals. In *Proc. ICASSP-96*, volume 3, pages 1688–169, May 1996.

[7] K. W. Forsyth and D. W. Bliss. MIMO radar: Concepts, performance enhancements, and applications. In J. Li and P. Stoica, editors, *MIMO Radar Signal Processing*. Wiley, New Jersey, 2009.

[8] H. Landau. Necessary density conditions for sampling and interpolation of certain entire functions. *Acta math.*, pages 37–52, 1967.

[9] K. Lee and Y. Bresler. Subspace-augmented MUSIC for joint sparse recovery with any rank. In *IEEE Sensor Array and Multichannel Signal Processing Workshop (SAM)*, pages 205–208, 2010.

[10] Y. M. Lu and M. N. Do. A theory for sampling signals from a union of subspaces. *IEEE Trans. Signal Processing*, 56(6):2334–2345, 2007.

[11] M. Mishali and Y. C. Eldar. Blind multiband signal reconstruction: Compressed sensing for analog signals. *IEEE Trans. Signal Processing*, 57(3):993–1009, 2009.

[12] M. Mishali, Y. C. Eldar, and A. Elron. Xampling: Signal acquisition and processing in union of subspaces. 2009.

[13] H. Nyquist. Certain topics in telegraph transmission theory. *Trans. Amer. Inst. of Elect. Eng.*, 47:617–644, Apr. 1928.

[14] A. V. Oppenheim and R. W. Schafer. *Digital Signal Processing*. Prentice Hall, Englewood Cliffs, NJ, 1975.

[15] L. C. Potter, E. Ertin, J. T. Parker, and M. Cetin. Sparsity and compressed sensing in radar imaging. *Proceedings of the IEEE*, 98(6):1006 –1020, 2010.

[16] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R. G. Baraniuk. Beyond Nyquist: Efficient sampling of sparse, bandlimited signals. *IEEE Trans. Info. Theory*, 56(1):520–544, Jan. 2010.

[17] H. L. Van Trees. *Detection Estimation and Modulation theory (Part IV: Optimum Array Processing)*. Wiley, New York, 2002.

[18] R. Venkataramani and Y. Bresler. Perfect reconstruction formulas and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals. *IEEE Trans. Info. Theory*, 46(6):2173–2183, September 2000.