Blind Multiband Signal Reconstruction: Compressed Sensing for Analog Signals

Moshe Mishali, *Student Member, IEEE*, and Yonina C. Eldar, *Senior Member, IEEE*

***Abstract*—We address the problem of reconstructing a multiband signal from its sub-Nyquist pointwise samples, when the band locations are unknown. Our approach assumes an existing multi-coset sampling. To date, recovery methods for this sampling strategy ensure perfect reconstruction either when the band locations are known, or under strict restrictions on the possible spectral supports. In this paper, only the number of bands and their widths are assumed without any other limitations on the support. We describe how to choose the parameters of the multi-coset sampling so that a unique multiband signal matches the given samples. To recover the signal, the continuous reconstruction is replaced by a single finite-dimensional problem without the need for discretization. The resulting problem is studied within the framework of compressed sensing, and thus can be solved efficiently using known tractable algorithms from this emerging area. We also develop a theoretical lower bound on the average sampling rate required for blind signal reconstruction, which is twice the minimal rate of known-spectrum recovery. Our method ensures perfect reconstruction for a wide class of signals sampled at the minimal rate, and provides a first systematic study of compressed sensing in a truly analog setting. Numerical experiments are presented demonstrating blind sampling and reconstruction with minimal sampling rate.**

***Index Terms*—Landau–Nyquist rate, multiband, multiple measurement vectors (MMV), nonuniform periodic sampling, sparsity.**

I. Introduction

THE well-known Whittaker, Kotelńikov, and Shannon (WKS) theorem links analog signals with discrete representations, allowing signal processing in a digital framework. The theorem states that a real-valued signal bandlimited to B Hertz can be perfectly reconstructed from its uniform samples if the sampling rate is at least $2B$ samples per second. This statement highlights the basic ingredients of a sampling theorem. First, a signal model which appears in practical applications and relates to physical properties. In this case, the maximal frequency B parameterizes a bandlimited signal model. Second, a minimal rate requirement to allow perfect reconstruction, referred to as the Nyquist rate in the WKS setting. Finally, a design of sampling and reconstruction stages which is based on the model parameters.

Manuscript received September 11, 2007; revised November 24, 2008. First published January 13, 2009; current version published February 13, 2009. The associate editor coordinating the review of this manuscript was Dr. Zoran Cvetkovic.

The authors are with the Technion—Israel Institute of Technology, Haifa 32000, Israel (e-mail: moshiko@tx.technion.ac.il; yonina@ee.technion.ac.il).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TSP.2009.2012791

In this paper, we consider the class of multiband signals, whose frequency support resides within several continuous intervals, spread over a wide spectrum. This model naturally arises when observing the sum of a few narrowband transmissions which are modulated by relatively high carriers, with different frequencies. When the band locations and their widths are known prior to sampling, the signal model defines a subspace of possible inputs. Landau [1] developed a minimal rate requirement for an arbitrary sampling method that allows perfect reconstruction in this setting. The Landau rate equals the sum of the band widths, which is below the corresponding Nyquist rate. Several works proposed sampling and reconstruction stages, designed according to the specific spectral support of the input signal. Low-rate uniform sampling was studied in [2] for a real bandpass signal, whereas [3] suggested periodic nonuniform sampling at the same average sampling rate. Lin and Vaidyanathan [4] extended this approach to multiband signals. These methods allow exact recovery at rates approaching that derived by Landau.

A much more challenging problem is to design a *blind* sampling and reconstruction system, that does not rely on knowledge of the band locations. Blindness is important whenever detecting the spectral support, prior to sampling, is impossible or too expensive to implement. Reconstruction under partial knowledge of the support was addressed in a series of conference papers [5]–[7]. These works do not assume the exact support but require that the band locations obey a certain mathematical condition. Sampling is carried out by a multi-coset strategy, independent of the band locations. In order to recover the signal, [5], [6] identified similarities with direction of arrival problems, implying potential use of techniques from this field, such as MUSIC [8]. An alternative method was proposed in [7] under the assumption that the support obeys a strict mathematical condition. Similar results were derived in [9]. This approach, however, suffers from two main drawbacks. First, as detailed in Section V-B, if the support violates the condition of [7], then the samples match many possible input signals, thus hindering any chance for recovery. Since this condition depends on the band locations which are unknown, it cannot be validated in practice. The difficulty arises form the fact that this requirement is not directly related to simple physical properties of a multiband model (i.e., the number of bands and their widths). Second, even when the condition is met, exact recovery is not guaranteed. This is a result of the fact that the proposed sampling rate approaches that dictated by Landau. As we prove in Section III, perfect blind reconstruction requires a higher sampling rate.

In subsequent publications, Herley and Wong [10] suggested a half-blind system. Similar ideas were later suggested in [11].

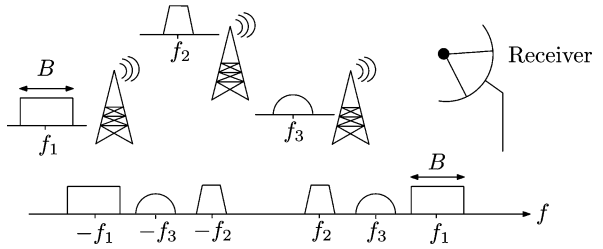

Fig. 1. Three RF transmissions with different carriers f_i. The receiver sees a multiband signal (bottom drawing). In this example the real-valued signal contains $N = 6$ bands, and the modulation techniques of the transmitters determine the maximal expected width B.

In this case, the signal is sampled using a multi-coset sampling strategy, independent of the band locations. However, recovery is based on knowledge of the spectral support. This approach guarantees perfect (nonblind) recovery under the same condition of [7]. Thus, to date, there exist no spectrum-blind method that is guaranteed to recover a multiband signal (with arbitrary band positions) at rates lower than the Nyquist rate.

A. Main Contributions

Following the ingredients of the WKS theorem, we begin by formulating a multiband model $\mathcal{M}$ which contains signals with at most N bands, whose widths do not exceed B. Fig. 1 depicts a typical communication application that obeys the model $\mathcal{M}$. Once the signal model is established, we develop a minimal rate requirement for blind perfect reconstruction of signals from $\mathcal{M}$ with an arbitrary system. As we show, the lower bound is twice the Landau rate and no more than the Nyquist rate.

The body of this paper focuses on spectrum-blind reconstruction (SBR) assuming a preceding multi-coset stage which satisfies the minimal rate requirement. Our first main result describes a parameter selection for the sampling stage such the every signal in $\mathcal{M}$ matches a unique digital representation. This selection relies on the known values N, B, but does not impose any restrictions on the band locations. We show that this choice is also valid for known-spectrum reconstruction with half the sampling rate.

Our approach to blind recovery is to first determine the spectral support from the given samples. Once the support is recovered, the reconstruction has a closed form solution, which motivates this paradigm. Theoretical tools are developed in order to transform the continuous support recovery problem into a finite-dimensional program without any discretization. We then prove that the solution can be obtained by finding the unique sparsest solution matrix from multiple measurement vectors. This set of operations is grouped under a block we name *Continuous to Finite* (CTF). This block is the cornerstone of two SBR algorithms we develop. One is entitled SBR4 and uses a single instance of the CTF. SBR4 guarantees perfect reconstruction when sampling at twice the minimal rate. The other is referred to as SBR2 and involves a bi-section process and several uses of the CTF block. SBR2 works at the minimal rate, but does not guarantee exact recovery for certain multiband signals. Other differences between the algorithms are also discussed. Both SBR4 and SBR2 can easily be implemented in DSP processors or in software environments.

The CTF block requires finding a sparsest solution matrix which is an NP-hard problem [13]. Several suboptimal efficient methods have been developed for this problem in the compressed sensing (CS) literature [13], [14]. In our algorithms, any of these techniques can be used. However, when using a suboptimal technique in SBR4, exact recovery is no longer guaranteed. Fortunately, numerical experiments on random constructions of multiband signals from $\mathcal{M}$ show that both SBR4 and SBR2 maintain a satisfactory exact recovery for minimal sampling rate even when suboptimal implementations of the CTF block are used. In addition, SBR4 is accompanied with a nice feature which indicates the recovery success, allowing to switch between suboptimal techniques in case of a failure. While discussing experimental results, we also address stability under measurement noise and emphasize the advantages of SBR4 and SBR2 over a simple inexact discretization-based reconstruction.

B. Compressed Sensing for Analog Signals

In his seminal work on compressed sensing [15], Donoho posed the ultimate goal of merging compression and sampling. In such an approach, the actual information contents dictate the sampling rate, rather than the dimensions of the ambient space in which the signal resides. This goal triggered a surge of research on various mathematical and algorithmic aspects of sensing sparse signals, which were mainly studied for discrete and finite vectors. The present work takes a first step towards sensing in a truly continuous environment. An adaptation of CS results to continuous signals was also considered in a series of conferences papers (see [16] and [17] and the references therein). However, these papers did not address the case of multiband signals. In [16], a vector signal was assumed, but the generalization to continuous signals was not studied, while in [17] an underlying discrete model was used so that the continuous signal is a linear combination of a finite number of known functions. Here, there is no discrete model as the signals are treated in a continuous framework without any discretization.[1]

A second aspect that distinguishes our work from mainstream CS papers relates to tractability of the recovery. From a sampling viewpoint, perfect reconstruction is a desired goal even when yielding a difficult recovery stage (i.e., an NP-hard problem). The WKS theorem, for example, suggests an ideal lowpass filter for reconstruction, whose practical implementation is highly demanding. The approach in compressive sensing promotes the opposite viewpoint and strives for tractable recovery algorithms at the expense of suboptimality in the sampling rate [13]–[15], [24].

C. Outline

The paper is organized as follows. In Section II, we formulate the signal model and the reconstruction goals. The minimal rate theorem for blind reconstruction is stated and proved in Section III. A brief overview of multi-coset sampling is presented in Section IV. We develop our main theoretical results on spectrum-blind reconstruction and present the CTF block in

[1]Since submission of the report [18] and the conference version [19], we have further expanded the analog CS framework in several follow-up papers; see [20]–[23].

TABLE I
NOTATION

$x(t)$	continuous-time signal with finite energy
$\|x(t)\|$	standard L_2 norm of $x(t)$, $\|x(t)\|^2 = \int_{-\infty}^{+\infty} \|x(t)\|^2 \mathrm{d}t$
$X(f)$	Fourier transform of $x(t)$ (assumed to exist)
$a[n]$	bounded energy sequence
$\|a[n]\|$	standard ℓ_2 norm of $a[n]$, $\|a[n]\|^2 = \sum_n \|a[n]\|^2$
z^*	conjugate of the complex number z
$\mathbf{v}$	vector
$\mathbf{v}_i$ or $\mathbf{v}(i)$	ith entry of $\mathbf{v}$
$\mathbf{v}(f)$	vector that depends on a continuous parameter f
$\mathbf{A}$	matrix
$\mathbf{A}_{ik}$	ikth entry of $\mathbf{A}$
$\mathbf{A}_i$	ith column of $\mathbf{A}$
$\mathbf{A}^T, \mathbf{A}^H$	transpose and the conjugate-transpose of $\mathbf{A}$
$\mathbf{A} \succeq 0$	Hermitian positive semi-definite (PSD) matrix
$\mathbf{A}^\dagger$	Moore-Penrose pseudo-inverse of $\mathbf{A}$
S	finite or countable set
S_i	ith element of S
$\|S\|$	cardinality of a finite set S
$\mathcal{T}$	infinite uncountable set
$\lambda(\mathcal{T})$	the Lebesgue measure of $\mathcal{T} \subseteq \mathbb{R}$
$I(\mathbf{v})$	support of $\mathbf{v}$: $I(\mathbf{v}) = \{k \,\|\, \mathbf{v}_k \neq 0\}$ indicates the locations of the nonzero entries of $\mathbf{v}$
$\|\mathbf{v}\|$	ℓ_2 norm of $\mathbf{v}$, $\|\mathbf{v}\|^2 = \sum_i \|\mathbf{v}_i\|^2$
$\|\mathbf{v}\|_0$	ℓ_0 pseudonorm of $\mathbf{v}$, $\|\mathbf{v}\|_0 = \|I(\mathbf{v})\|$
$I(\mathbf{A})$	support of $\mathbf{A}$: $I(\mathbf{A}) = \{k \,\|\, (\mathbf{A}^T)_k \neq \mathbf{0}\}$ indicates the rows of $\mathbf{A}$ that are non-identically zero
$\mathbf{A}_S$	column restriction of $\mathbf{A}$ to S: the columns of $\mathbf{A}$ indicated by S, $(\mathbf{A}_S)_i = (\mathbf{A})_{S_i}$, $1 \leq i \leq \|S\|$
$\mathbf{A}^S$	row restriction of $\mathbf{A}$ to S: the rows of $\mathbf{A}$ indicated by S, $(\mathbf{A}^S)^T = (\mathbf{A}^T)_S$

Section V. Based on these results, in Section VI, we design and compare the SBR4 and SBR2 algorithms. Numerical experiments are described in Section VII.

II. Signal Model and Problem Formulation

A. Multiband Signal Model

Common notation, as summarized in Table I, is used throughout the paper. It is assumed that $x(t)$ is a continuous function of t, and that the Fourier transform

$$X(f) = \int_{-\infty}^{+\infty} x(t) \exp(-j2\pi ft)\,\mathrm{d}t \tag{1}$$

is (piecewise) continuous in f.

The following multiband model is the prime focus of our work.

Definition 1: The set $\mathcal{M}$ contains all complex-valued signals $x(t)$ satisfying the following:

1) $x(t)$ is bandlimited to $\mathcal{F} = [0, 1/T]$, namely $X(f) = 0, f \notin \mathcal{F}$;
2) $X(f)$ is further restricted to have a support on no more than N disjoint intervals (bands) in $\mathcal{F}$; and
3) the band widths do not exceed B.

Fig. 2 depicts a typical spectral support of a multiband signal $x(t) \in \mathcal{M}$. The Nyquist rate corresponding to any $x(t) \in \mathcal{M}$ is $1/T$. Note that each band is uniquely represented by its edges $[a_i, b_i]$; however, the band positions of $x(t) \in \mathcal{M}$ are arbitrary, and in particular, unknown in advance.

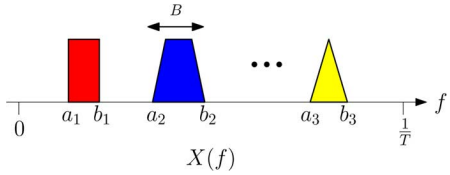

Fig. 2. Typical spectrum support of $x(t) \in \mathcal{M}$.

Although our interest is mainly in signals $x(t) \in \mathcal{M}$, our results are applicable to a broader class of signals, as explained in the relevant sections. In addition, the results of the paper are easily adopted to real-valued signals supported on $[-1/2T, +1/2T]$. The required modifications are detailed in Appendix A.

B. Problem Formulation

We wish to perfectly reconstruct $x(t) \in \mathcal{M}$ from a set of pointwise samples under two constraints. One is blindness, so that the information about the band locations is not used while acquiring the samples and neither can it be used in the reconstruction process. The other is that the sampling rate required to guarantee perfect reconstruction should be minimal.

This problem is solved if either of the constraints is removed. Without the rate constraint, the WKS theorem allows perfect blind-reconstruction for every signal $x(t)$ bandlimited to $\mathcal{F}$ from its uniform samples $x(nT)$ at the Nyquist rate. Alternatively, if the exact number of bands and their locations are known, then under mild conditions on the band positions the method of [4] allows perfect reconstruction at the minimal sampling rate provided by Landau's theorem [1].

III. Minimal Sampling Rate

We begin by quoting Landau's theorem for the minimal sampling rate of an arbitrary sampling method that allows known-spectrum perfect reconstruction. We then develop a similar theorem for blind reconstruction, which proves that twice the Landau rate is required to allow for exact recovery.

A. Known Spectrum Support

Consider the space of bandlimited functions whose Fourier transform is restricted to a known support $\mathcal{T} \subseteq \mathcal{F}$:

$$\mathcal{B}_\mathcal{T} = \{x(t) \in L^2(\mathbb{R}) \,|\, \mathrm{supp} X(f) \subseteq \mathcal{T}\}. \tag{2}$$

A set $R = \{r_n\}$ is called *a sampling set* for $\mathcal{B}_\mathcal{T}$ if the sequence of samples $x_R[n] = x(r_n)$ is stable, namely there exist constants $\alpha > 0$ and $\beta < \infty$ such that

$$\alpha\|x - y\|^2 \leq \|x_R - y_R\|^2 \leq \beta\|x - y\|^2 \quad \forall x, y \in \mathcal{B}_\mathcal{T}. \tag{3}$$

Landau [1] proved that if R is a sampling set for $\mathcal{B}_\mathcal{T}$ then it must have a density $D^-(R) \geq \lambda(\mathcal{T})$, where

$$D^-(R) = \lim_{r\to\infty} \inf_{y\in\mathbb{R}} \frac{|R \cap [y, y + r]|}{r} \tag{4}$$

is the lower Beurling density, and $\lambda(\mathcal{T})$ is the Lebesgue measure of $\mathcal{T}$. The numerator in (4) counts the number of points from

R in every interval of width r of the real axis.[2] The Beurling density (4) reduces to the usual concept of *average* sampling rate for uniform and periodic nonuniform sampling. The Landau rate, $\lambda(\mathcal{T})$, is the minimal average sampling rate for $\mathcal{B}_{\mathcal{T}}$.

B. Unknown Spectrum Support

Consider the set $\mathcal{N}_\Omega$ of signals bandlimited to $\mathcal{F}$ with bandwidth occupation no more than $0 < \Omega < 1$, so that

$$\lambda(\mathrm{supp}X(f)) \leq \frac{\Omega}{T} \quad \forall x(t) \in \mathcal{N}_\Omega. \tag{5}$$

The Nyquist rate for $\mathcal{N}_\Omega$ is $1/T$. Note that $\mathcal{N}_\Omega$ is not a subspace so that the Landau theorem does not apply here. Nevertheless, it is intuitive to argue that the minimal sampling rate for $\mathcal{N}_\Omega$ cannot be below Ω/T as this value is the Landau rate had the spectrum support been known.

A blind sampling set R for $\mathcal{N}_\Omega$ is a stable sampling set whose design does not assume knowledge of $\mathrm{supp}X(f)$. Similarly to (3) the stability of R requires the existence of $\alpha > 0$ and $\beta < \infty$ such that

$$\alpha\|x-y\|^2 \leq \|x_R - y_R\|^2 \leq \beta\|x-y\|^2 \quad \forall x, y \in \mathcal{N}_\Omega. \tag{6}$$

Theorem 1 (Minimal Sampling Rate): Let R be a blind sampling set for $\mathcal{N}_\Omega$. Then

$$D^-(R) \geq \min\left\{\frac{2\Omega}{T}, \frac{1}{T}\right\}. \tag{7}$$

Proof: The set $\mathcal{N}_\Omega$ is of the form

$$\mathcal{N}_\Omega = \bigcup_{\mathcal{T}\in\Gamma} \mathcal{B}_{\mathcal{T}} \tag{8}$$

where $\mathcal{B}_{\mathcal{T}}$ is defined by (2) and

$$\Gamma = \{\mathcal{T} \,|\, \mathcal{T} \subseteq \mathcal{F}, \lambda(\mathcal{T}) \leq \Omega/T\}. \tag{9}$$

Clearly, $\mathcal{N}_\Omega$ is an uncountable union of subspaces. Sampling signals that lie in a union of subspaces has been recently treated in [12]. For every $\gamma, \theta \in \Gamma$ define the subspaces

$$\mathcal{B}_{\gamma,\theta} = \mathcal{B}_\gamma + \mathcal{B}_\theta = \{x + y \,|\, x \in \mathcal{B}_\gamma, y \in \mathcal{B}_\theta\}. \tag{10}$$

Since R is a sampling set for $\mathcal{N}_\Omega$, (6) holds for some constants $\alpha > 0, \beta < \infty$. It was proved in ([12], Proposition 2) that (6) is valid if and only if

$$\alpha\|x-y\|^2 \leq \|x_R - y_R\|^2 \leq \beta\|x-y\|^2 \quad \forall x, y \in \mathcal{B}_{\gamma,\theta} \tag{11}$$

holds for every $\gamma, \theta \in \Gamma$. In particular, R is a sampling set for every $\mathcal{B}_{\gamma,\theta}$ with $\gamma, \theta \in \Gamma$.

Now, the space $\mathcal{B}_{\gamma,\theta}$ is of the form (2) with $\mathcal{T} = \gamma \cup \theta$. Applying Landau's theorem for each $\gamma, \theta \in \Gamma$ results in

$$D^-(R) \geq \lambda(\gamma \cup \theta) \quad \forall \gamma, \theta \in \Gamma. \tag{12}$$

[2]The numerator is not necessarily finite but as the sampling set is countable the infimum takes on a finite value.

Choosing

$$\gamma = \left[0, \frac{\Omega}{T}\right], \quad \theta = \left[\frac{1-\Omega}{T}, \frac{1}{T}\right] \tag{13}$$

we have that for $\Omega \leq 0.5$,

$$D^-(R) \geq \lambda(\gamma \cup \theta) = \lambda(\gamma) + \lambda(\theta) = \frac{2\Omega}{T}. \tag{14}$$

If $\Omega \geq 0.5$, then $\gamma \cup \theta = \mathcal{F}$ and

$$D^-(R) \geq \lambda(\gamma \cup \theta) = \frac{1}{T}. \tag{15}$$

Combining (14) and (15) completes the proof. ■

In [12], the authors consider minimal sampling requirements for a union of shift-invariant subspaces, with a particular structure of sampling functions. Specifically, they view the samples as inner products with sampling functions of the form $\{\psi_k(t-m)\}_{1\leq k\leq K, m\in\mathbb{Z}}$, which includes multi-coset sampling. Theorem 1 extends this result to an arbitrary pointwise sampling operator. In particular, it is valid for aperiodic sampling sets that are not covered by [12].

An immediate corollary of Theorem 1 is that if $\Omega > 0.5$ then uniform sampling at the Nyquist rate and reconstruction with an ideal lowpass filter satisfies the requirements of our problem formulation: sampling and reconstruction are independent of the band locations, and the sampling rate is minimal. Since $\mathcal{M}$ is contained in the space of bandlimited signals, this choice provides perfect reconstruction for every $x(t) \in \mathcal{M}$. Therefore, in the sequel we assume that $\Omega \leq 0.5$ so that the minimal sampling rate is twice the Landau rate.

It is easy to see that $\mathcal{M}$ is a true subset of $\mathcal{N}_\Omega$ for $\Omega = NBT$. The Proof of Theorem 1 can be adapted to show that sampling at an average rate of $2NB$ is necessary to allow blind perfect reconstruction for all signals from $\mathcal{M}$ (for $\Omega = NBT < 0.5$). For known spectral support of $x(t) \in \mathcal{M}$, the Landau rate is NB.

Both Landau's theorem and Theorem 1 state a lower bound but do not provide a method to achieve the bound. An important distinction between the two theorems is that for known spectrum the set $\mathcal{B}_{\mathcal{T}}$ is linear (namely, closed under linear combinations) and known reconstruction schemes [4], [11] are also linear. Consequently, the stability condition (3) suffices to ensure stability of the reconstruction. In contrast, $\mathcal{N}_\Omega$ is a nonlinear set (as is $\mathcal{M}$), and our proposed reconstruction scheme is also nonlinear. Theorem 1 states the necessary density for a blind stable sampling set while a higher density may be required to ensure stable reconstruction.

The rest of the paper is devoted to developing a reconstruction method that approaches the minimal sampling rate of Theorem 1 while ensuring perfect reconstruction. We address stability in Section VII-C.

IV. Universal Sampling

This section reviews multi-coset sampling which is used in our development. We also briefly explain the fundamentals of known-spectrum reconstruction as derived in [11].

A. Multi-Coset Sampling

Uniform sampling of $x(t)$ at the Nyquist rate results in samples $x(nT)$ that contain all the information about $x(t)$. Multi-coset sampling is a selection of certain samples from this grid. The uniform grid is divided into blocks of L consecutive samples. A constant set C of length p describes the indexes of p samples that are kept in each block. The set $C = \{c_i\}_{i=1}^{p}$ is referred to as the sampling pattern where

$$0 \le c_1 < c_2 < \cdots < c_p \le L-1. \tag{16}$$

Define the ith sampling sequence for $1 \le i \le p$ as

$$x_{c_i}[n] = \begin{cases} x(nT), & n = mL + c_i, m \in \mathbb{Z} \\ 0, & \text{otherwise.} \end{cases} \tag{17}$$

The sampling stage can be implemented by p uniform sampling sequences with period $1/LT$, where the ith sampling sequence is shifted by $c_i T$ from the origin.

Direct calculations link the known discrete-time Fourier transform $X_{c_i}(e^{j2\pi fT})$ of $x_{c_i}[n]$ to the unknown Fourier transform $X(f)$ of $x(t)$:

$$\begin{aligned} X_{c_i}(e^{j2\pi fT}) &= \sum_{n=-\infty}^{+\infty} x_{c_i}[n] \exp(-j2\pi fnT) \\ &= \frac{1}{LT} \sum_{r=0}^{L-1} \exp\left(j\frac{2\pi}{L} c_i r\right) X\left(f + \frac{r}{LT}\right) \end{aligned} \tag{18}$$

for every $1 \le i \le p$, and every f in the interval

$$\mathcal{F}_0 = \left[0, \frac{1}{LT}\right). \tag{19}$$

It is convenient to express (18) in matrix form as

$$\mathbf{y}(f) = \mathbf{A}\mathbf{x}(f) \quad \forall f \in \mathcal{F}_0 \tag{20}$$

where $\mathbf{y}(f)$ is a vector of length p whose ith element is $X_{c_i}(e^{j2\pi fT})$, $\mathbf{A}$ is a matrix with ikth element given by

$$\mathbf{A}_{ik} = \frac{1}{LT} \exp\left(j\frac{2\pi}{L} c_i k\right) \tag{21}$$

and the vector $\mathbf{x}(f)$ contains L unknowns for each f

$$\mathbf{x}_i(f) = X\left(f + \frac{i}{LT}\right), \quad 0 \le i \le L-1, \quad f \in \mathcal{F}_0. \tag{22}$$

Fig. 3 illustrates the relation between $X(f)$ and $\mathbf{x}(f)$. The problem of recovering $x(t)$ is equivalent to determining $\mathbf{x}(f)$ from $\mathbf{y}(f)$ for all $f \in \mathcal{F}_0$.

Dealing with real-valued multiband signals requires simple modifications to (20). These adjustments are detailed in Appendix A.

The average sampling rate of a multi-coset sampling set is

$$\frac{1}{T_{\text{AVG}}} = \frac{p}{LT} \tag{23}$$

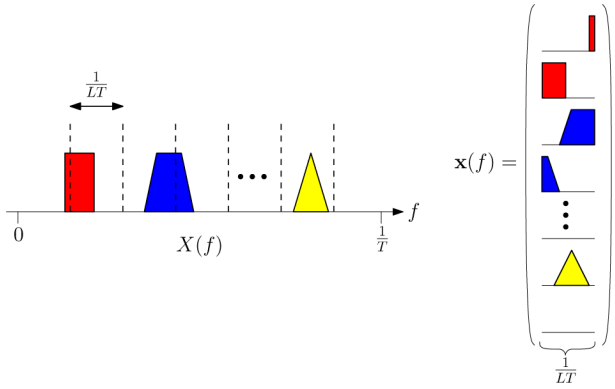

Fig. 3. Vector $\mathbf{x}(f)$ is constructed by slicing $X(f)$ into L equal-length intervals, and rearranging them in vector form.

which is lower than the Nyquist rate for $p < L$. However, an average sampling rate above the Landau rate is not sufficient for known-spectrum reconstruction. Additional conditions are needed as explained in the next section.

B. Known-Spectrum Reconstruction and Universality

Presentation of the reconstruction is simplified using CS sparsity notation. A vector $\mathbf{v}$ is called K-sparse if the number of nonzero values in $\mathbf{v}$ is not greater than K, $\|\mathbf{v}\|_0 \le K$. We also use the following definition of the Kruskal-rank of a matrix [25]:

Definition 2: The Kruskal-rank of $\mathbf{A}$, denoted $\sigma(\mathbf{A})$, is the maximal number q such that every set of q columns of $\mathbf{A}$ is linearly independent.

Observe that for every $f \in \mathcal{F}_0$ the system (20) has fewer equations than unknowns. However, when incorporating the structure of $\mathbf{x}(f)$ as prior, the number of unknowns may be reduced. As evident from Fig. 3, every band can contribute several nonzero values to $\mathbf{x}(f)$. Thus, in order to match the number of equations to the number of unknowns, it is necessary that

$$\mathbf{x}(f) \text{ is } p\text{-sparse} \quad \forall f \in \mathcal{F}_0. \tag{24}$$

In [11] it is assumed that the information about the band locations is available for reconstruction, so that the support set $I(\mathbf{x}(f))$ is known for every $f \in \mathcal{F}_0$. In this setting, a sufficient condition for a unique solution in (20) can be formulated using the Kruskal-rank of $\mathbf{A}$:

$$\mathbf{x}(f) \text{ is } \sigma(\mathbf{A})\text{-sparse} \quad \forall f \in \mathcal{F}_0. \tag{25}$$

The known-spectrum reconstruction of [11] basically restricts the columns of $\mathbf{A}$ to $I(\mathbf{x}(f))$ and inverts the resulting matrix in order to recover $\mathbf{x}(f)$.

A sampling pattern C that yields a fully Kruskal-rank $\mathbf{A}, \sigma(\mathbf{A}) = p$, is called universal and corresponds to $\sigma(\mathbf{A}) = p$ [11]. Therefore, the set of signals that are consistent with (25) is the broadest possible if a universal sampling pattern is used. As we show later, choosing $L \le 1/BT, p \ge N$ and a universal pattern C makes (25) valid for every signal $x(t) \in \mathcal{M}$.

Finding a universal pattern C, namely one that results in a fully Kruskal-rank $\mathbf{A}$, is a combinatorial problem. Several specific constructions of sampling patterns are proved to be universal [11], [26]. In particular, choosing a prime integer L renders every pattern universal [26]. In addition, a random pattern C drawn uniformly among all $\binom{L}{p}$ combinations is universal with high probability; see [24] for details.

V. Spectrum-Blind Reconstruction

In this section, we develop the theory needed for SBR. These results are then used in the next section to construct two efficient algorithms for blind signal reconstruction.

A. Setting Up the Sampling Stage

In the blind setting, our only information is that $\mathbf{x}(f)$ is sparse since every band contributes only a few nonzero values. However, the support $I(\mathbf{x}(f))$ is unknown. The following theorem from the CS literature is used to provide a sufficient condition for a unique solution in (20).

Theorem 2: Suppose $\bar{\mathbf{x}}$ is a solution of $\mathbf{y} = \mathbf{A}\mathbf{x}$. If $\|\bar{\mathbf{x}}\|_0 \leq \sigma(\mathbf{A})/2$ then $\bar{\mathbf{x}}$ is the unique sparsest solution of the system.

Theorem 2 and its proof are given in [13], [27] with a slightly different notation of $\mathrm{Spark}(\mathbf{A})$ instead of the Kruskal-rank of $\mathbf{A}$. Note that the condition of the theorem is not necessary as there are examples in which the sparsest solution $\bar{\mathbf{x}}$ of $\mathbf{y} = \mathbf{A}\mathbf{x}$ is unique while $\|\bar{\mathbf{x}}\|_0 > \sigma(\mathbf{A})/2$.

Using Theorem 2, it is evident that perfect reconstruction is possible for every signal satisfying

$$\mathbf{x}(f) \text{ is } \frac{\sigma(\mathbf{A})}{2}\text{-sparse} \quad \forall f \in \mathcal{F}_0. \tag{26}$$

As before, choosing a universal pattern makes the set of signals that conform with (26) the widest possible. Note that a factor of two distinguishes between the sufficient conditions (25) and (26) reflecting the missing information on the nonzero locations of $\mathbf{x}(f)$.

For a given number of sampling cosets p and a universal pattern C, (26) implies that certain spectral supports can be recovered. However, these supports are defined in (26) using the unknown values of $\mathbf{x}(f)$. The following theorem proposes a parameter selection for the multi-coset strategy, based on the known values N, B, T, such that every $x(t) \in \mathcal{M}$ can be perfectly reconstructed. The theorem is valid for both known and blind reconstruction with a slight difference resulting from the factor of two in the sufficient conditions.

Theorem 3 (Uniqueness): Let $x(t) \in \mathcal{M}$ be a multiband signal. If:

1) $L \leq 1/BT$;
2) $p \geq N$ for nonblind reconstruction or $p \geq 2N$ for blind;
3) C is a universal pattern

then, for every $f \in \mathcal{F}_0$, the vector $\mathbf{x}(f)$ is the unique N-sparse solution of (20).

Proof: If $L \leq 1/BT$ then for the ith band $\mathcal{T}_i = [a_i, b_i]$,

$$\lambda(\mathcal{T}_i) \leq B \leq \frac{1}{LT}, \quad 1 \leq i \leq N. \tag{27}$$

Therefore, $f \in \mathcal{T}_i$ implies

$$f + \frac{k}{LT} \notin \mathcal{T}_i \quad \forall k \neq 0. \tag{28}$$

According to (22) and Fig. 3 for every $f \in \mathcal{F}_0$ the vector $\mathbf{x}(f)$ takes the values of $X(f)$ on a set of L points spaced by $1/LT$. Consequently, the number of nonzero values in $\mathbf{x}(f)$ is not greater than the number of bands, namely $\mathbf{x}(f)$ is N-sparse.

Since C is a universal pattern, $\sigma(\mathbf{A}) = p$. This implies that conditions (25) and (26) are satisfied. ■

The parameter selection of Theorem 3 required to ensure uniqueness for every $x(t) \in \mathcal{M}$ is not restrictive as the minimal sampling rate is still allowed. Indeed, from (23):

$$\frac{1}{T_{\mathrm{AVG}}} = \frac{p}{LT} \geq \begin{cases} NB & \text{nonblind reconstruction} \\ 2NB & \text{blind reconstruction.} \end{cases} \tag{29}$$

Equality is achieved when p is minimal and $L = 1/BT$ is an integer.

B. Reconstruction Paradigm

The goal of our reconstruction scheme is to perfectly recover $x(t)$ from the set of sequences $x_{c_i}[n], 1 \leq i \leq p$, or equivalently, to reconstruct $\mathbf{x}(f)$ for every $f \in \mathcal{F}_0$ from the input data $\mathbf{y}(f)$. Although the problem is similar to that described in the CS literature, here finding the unique sparse vector must be solved for each value f in the continuous interval $\mathcal{F}_0$, which clearly cannot be implemented.

A straightforward approach is to find the sparsest vector $\mathbf{x}(f)$ over a dense grid of $\mathcal{F}_0$ and then approximate the solution over the entire continuous interval $\mathcal{F}_0$. However, this discretization strategy cannot guarantee perfect reconstruction.

Our reconstruction paradigm is targeted at finding the set

$$S = \bigcup_{f \in \mathcal{F}_0} I(\mathbf{x}(f)) \tag{30}$$

referred to as the *diversity* of $x(t)$. With the knowledge of S perfect reconstruction of $\mathbf{x}(f)$ is possible for every $f \in \mathcal{F}_0$ by noting that (20) can be written as

$$\mathbf{y}(f) = \mathbf{A}_S \mathbf{x}^S(f). \tag{31}$$

If the diversity set of $x(t)$ satisfies

$$|S| \leq \sigma(\mathbf{A}) \tag{32}$$

then the $p \times |S|$ matrix $\mathbf{A}_S$ has full column rank and

$$\mathbf{A}_S^{\dagger}\mathbf{A}_S = \left(\mathbf{A}_S^H \mathbf{A}_S\right)^{-1} \mathbf{A}_S^H \mathbf{A}_S = \mathbf{I}. \tag{33}$$

Multiplying both sides of (31) by $\mathbf{A}_S^{\dagger}$ results in

$$\mathbf{x}^S(f) = \mathbf{A}_S^{\dagger}\mathbf{y}(f) \quad \forall f \in \mathcal{F}_0. \tag{34}$$

From (30)

$$\mathbf{x}_i(f) = 0 \quad \forall f \in \mathcal{F}_0, i \notin S. \tag{35}$$

Thus, once S is known, and as long as (32) holds, perfect reconstruction can be obtained by (34)–(35). We point out that once the set S is recovered there is no essential difference between known and blind reconstruction.

This recovery paradigm is also used in [7]. However, it is assumed that the band locations obey $q \stackrel{\triangle}{=} |S| \leq p-1$. Since S of (30) involves the unknown $\mathbf{x}(f)$, this condition cannot be checked in practice, as is true for (24), (25) and (26). Moreover, when $q \geq p$, it is well known that (20) has many nonunique solutions ([28], Theorem 2). Therefore, $q \leq p-1$ is necessary but imposes a strict limitation on the possible band locations. Our model $\mathcal{M}$ and Theorem 3 circumvent these limitations by relying on the values N, B, T, which are known in practical applications (see Fig. 1). We also note that it is now clear why exact recovery is only almost surely guaranteed in [7]. Using $q \leq p-1$ represents a support measure q/LT with sampling rate approaching that of Landau. Theorem 1 proves, however, that exact recovery necessitates sampling at twice the Landau rate.

As we shall see later, (32) is implied by the condition required to transform the problem into a finite dimensional one. Furthermore, the following proposition shows that for $x(t) \in \mathcal{M}$, (32) is implied by the parameter selection of Theorem 3.

Proposition 1: If $L \leq 1/BT$ then $|S| \leq 2N$. If in addition $p \geq 2N$ and C is universal then for every $x(t) \in \mathcal{M}$, the set S satisfies (32).

Proof: The bands are continuous intervals upper bounded by B. From (22) and Fig. 3, it follows that $\mathbf{x}(f)$ is constructed by dividing $\mathcal{F}$ into L equal intervals of length $1/LT$. Therefore, if L is limited according to Theorem 3, then each band can either be fully contained in one of these intervals or it can be split between two consecutive intervals. Since the number of bands is no more than N it follows that $|S| \leq 2N$. With the additional conditions, we have that $\sigma(\mathbf{A}) = p \geq 2N \geq |S|$. ∎

C. Formulation of a Finite-Dimensional Problem

The set of (20) consists of an infinite number of linear systems because f is a continuous variable. Furthermore, the diversity set S given in (30) involves a union over the same continuous variable. The SBR algorithms developed in the next section show that S can be recovered exactly using only one finite-dimensional problem. We now develop the underlying theory behind these algorithms.

Consider a given $\mathcal{T} \subseteq \mathcal{F}_0$. Multiplying each side of (20) by its conjugate transpose, we have

$$\mathbf{y}(f)\mathbf{y}^H(f) = \mathbf{A}\mathbf{x}(f)\mathbf{x}^H(f)\mathbf{A}^H \quad \forall f \in \mathcal{T}. \tag{36}$$

Integrating both sides over the continuous variable f gives

$$\mathbf{Q} = \mathbf{A}\mathbf{Z}_0\mathbf{A}^H \tag{37}$$

with the $p \times p$ matrix

$$\mathbf{Q} = \int_{f\in\mathcal{T}} \mathbf{y}(f)\mathbf{y}^H(f)\,\mathrm{d}f \succeq 0 \tag{38}$$

and the $L \times L$ matrix

$$\mathbf{Z}_0 = \int_{f\in\mathcal{T}} \mathbf{x}(f)\mathbf{x}^H(f)\,\mathrm{d}f \succeq 0. \tag{39}$$

Define the diversity set of the interval $\mathcal{T}$ as

$$S_{\mathcal{T}} = \bigcup_{f\in\mathcal{T}} I(\mathbf{x}(f)). \tag{40}$$

Now,

$$(\mathbf{Z}_0)_{ii} = \int_{f\in\mathcal{T}} |\mathbf{x}_i(f)|^2 \mathrm{d}f. \tag{41}$$

Since $X(f)$ is piecewise continuous in f, $(\mathbf{Z}_0)_{ii} = 0$ if and only if $\mathbf{x}_i(f) = 0, \forall f \in \mathcal{T}$, which implies that $S_{\mathcal{T}} = I(\mathbf{Z}_0)$.

The next proposition sets the conditions for $\mathbf{Z}_0$ to be determined by the known matrix $\mathbf{Q}$. The proposition is stated for general matrices $\mathbf{Q}, \mathbf{A}$.

Proposition 2: Suppose a $p \times p$ matrix $\mathbf{Q} \succeq 0$ and a $p \times L$ matrix $\mathbf{A}$ are given. Let $\mathbf{Z}$ be an $L \times L$ matrix satisfying

$$\mathbf{Q} = \mathbf{A}\mathbf{Z}\mathbf{A}^H \tag{42a}$$

$$\mathbf{Z} \succeq 0 \tag{42b}$$

$$|I(\mathbf{Z})| \leq \sigma(\mathbf{A}). \tag{42c}$$

Then, $\mathrm{rank}(\mathbf{Z}) = \mathrm{rank}(\mathbf{Q})$. In addition, there can be only a single matrix $\mathbf{Z}$ satisfying (42a)–(42b) and

$$|I(\mathbf{Z})| \leq \frac{\sigma(\mathbf{A})}{2}. \tag{42d}$$

Proof: Suppose $\mathbf{Z}$ satisfies (42a)–(42c). Define $r_Q = \mathrm{rank}(\mathbf{Q}), r_Z = \mathrm{rank}(\mathbf{Z})$. Since $\mathbf{Z} \succeq 0$ it can be decomposed as $\mathbf{Z} = \mathbf{P}\mathbf{P}^H$ with $\mathbf{P}$ of size $L \times r_Z$ having orthogonal columns. From (42a)

$$\mathbf{Q} = (\mathbf{A}\mathbf{P})(\mathbf{A}\mathbf{P})^H. \tag{43}$$

It can be easily be concluded that $I(\mathbf{Z}) = I(\mathbf{P})$, and thus $|I(\mathbf{P})| \leq \sigma(\mathbf{A})$. The following lemma whose proof is given in Appendix B ensures that the matrix $\mathbf{A}\mathbf{P}$ of size $p \times r_Z$ also has full column rank.

Lemma 1: For every two matrices $\mathbf{A}, \mathbf{P}$, if $|I(\mathbf{P})| \leq \sigma(\mathbf{A})$ then $\mathrm{rank}(\mathbf{P}) = \mathrm{rank}(\mathbf{A}\mathbf{P})$.

Since for every matrix $\mathbf{M}$ it is true that $\mathrm{rank}(\mathbf{M}) = \mathrm{rank}(\mathbf{M}\mathbf{M}^H)$, (43) implies $r_Z = r_Q$.

For the second part of Proposition 2 suppose that $\mathbf{Z}, \tilde{\mathbf{Z}}$ both satisfy (42a), (42b), and (42d). The previous arguments ensure that

$$\mathrm{rank}(\mathbf{Z}) = \mathrm{rank}(\tilde{\mathbf{Z}}) = r_Q. \tag{44}$$

Following the earlier decompositions, we write

$$\begin{aligned} \mathbf{Z} &= \mathbf{P}\mathbf{P}^H, \quad I(\mathbf{Z}) = I(\mathbf{P}) \\ \tilde{\mathbf{Z}} &= \tilde{\mathbf{P}}\tilde{\mathbf{P}}^H, \quad I(\tilde{\mathbf{Z}}) = I(\tilde{\mathbf{P}}). \end{aligned} \tag{45}$$

In addition,

$$|I(\mathbf{P})| \leq \frac{\sigma(\mathbf{A})}{2}, \quad |I(\tilde{\mathbf{P}})| \leq \frac{\sigma(\mathbf{A})}{2}. \tag{46}$$

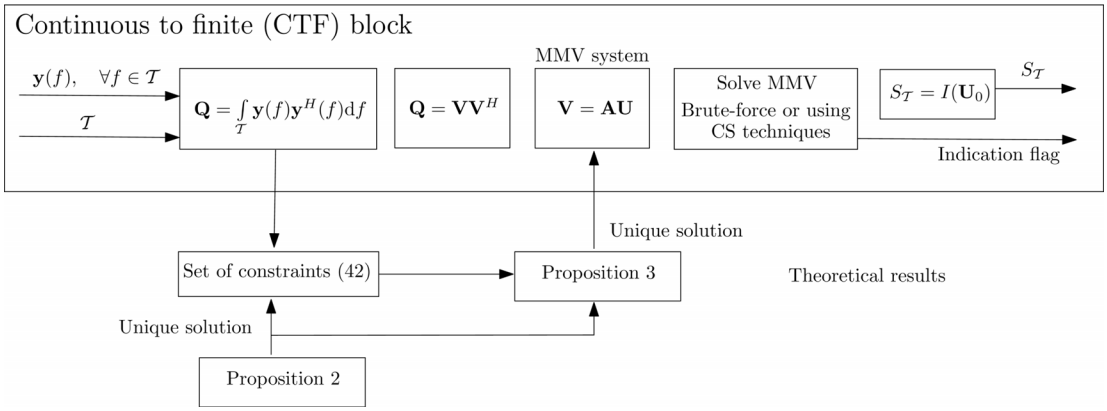

Fig. 4. CTF block determines the diversity set S_T of $\mathbf{x}(f)$ in a given interval $\mathcal{T}$.

From (42a),

$$\mathbf{Q} = (\mathbf{AP})(\mathbf{AP})^H = (\mathbf{A}\tilde{\mathbf{P}})(\mathbf{A}\tilde{\mathbf{P}})^H \tag{47}$$

which implies that

$$\mathbf{A}(\mathbf{P} - \tilde{\mathbf{P}}\mathbf{R}) = 0 \tag{48}$$

for some unitary matrix $\mathbf{R}$. It is easy to see that (46) results in $|I(\tilde{\mathbf{P}}\mathbf{R})| \leq \sigma(\mathbf{A})/2$. Therefore, the matrix $\mathbf{P} - \tilde{\mathbf{P}}\mathbf{R}$ has at most $\sigma(\mathbf{A})$ rows that are nonidentically zero. Applying Lemma 1 to (48) results in $\mathbf{P} = \tilde{\mathbf{P}}\mathbf{R}$. Substituting this into (45), we have that $\mathbf{Z} = \tilde{\mathbf{Z}}$. ■

The following proposition shows how to construct the matrix $\mathbf{Z}$ by finding the sparsest solution of a linear system.

Proposition 3: Consider the setting of Proposition 2 and assume $\mathbf{Z}$ satisfies (42a), (42b), and (42d). Let $r = \mathrm{rank}(\mathbf{Q})$ and define a matrix $\mathbf{V}$ of size $p \times r$ using the decomposition $\mathbf{Q} = \mathbf{V}\mathbf{V}^H$, such that $\mathbf{V}$ has r orthogonal columns. Then the linear system

$$\mathbf{V} = \mathbf{A}\mathbf{U} \tag{49}$$

has a unique sparsest solution matrix $\mathbf{U}_0$. Namely, $\mathbf{V} = \mathbf{A}\mathbf{U}_0$ and $|I(\mathbf{U}_0)|$ is minimal. Moreover, $\mathbf{Z} = \mathbf{U}_0\mathbf{U}_0^H$.

Proof: Substitute the decomposition $\mathbf{Q} = \mathbf{V}\mathbf{V}^H$ into (42a) and let $\mathbf{Z} = \mathbf{P}\mathbf{P}^H$. The result is $\mathbf{V} = \mathbf{A}\mathbf{P}\mathbf{R}$ for some unitary $\mathbf{R}$. Therefore, the linear system (49) has a solution $\mathbf{U}_0 = \mathbf{P}\mathbf{R}$. It is easy to see that $I(\mathbf{U}_0) = I(\mathbf{P}) = I(\mathbf{Z})$, thus (42d) results in $|I(\mathbf{U}_0)| \leq \sigma(\mathbf{A})/2$. Applying Theorem 2 to each of the columns of $\mathbf{U}_0$ provides the uniqueness of $\mathbf{U}_0$. It is trivial that $\mathbf{Z} = \mathbf{U}_0\mathbf{U}_0^H$. ■

Using the same arguments as in the proof it is easy to conclude that $I(\mathbf{Z}) = I(\mathbf{U}_0)$, so that $S_{\mathcal{T}}$ can be found directly from the solution matrix $\mathbf{U}_0$. In particular, we develop the *Continuous to Finite* (CTF) block which determines the diversity set $S_{\mathcal{T}}$ of a given frequency interval $\mathcal{T}$. Fig. 4 presents the CTF block that contains the flow of transforming the continuous linear system (20) on the interval $\mathcal{T}$ into the finite dimensional problem (49) and then to the recovery of $S_{\mathcal{T}}$. The role of Propositions 2 and 3 is also illustrated.

The advantage of our approach is that by defining the recovery problem within the CS framework, we may use theoretical results and algorithms from this field. Specifically, the linear system (49) is referred to as *multiple measurement vectors* (MMV) in the CS literature. Proposition 3 guarantees that the CTF produces an MMV with a unique sparsest solution. It is however a well-known CS result that finding $\mathbf{U}_0$ is solvable but NP-hard [13]. Several suboptimal efficient algorithms for finding $\mathbf{U}_0$ are given in [13], [14], [20], [29], and [30]. Some of them can indicate a success recovery of $\mathbf{U}_0$. We explain which class of algorithms has this property in Section IV-A.

VI. SBR Algorithms

The theoretical results developed in the previous section are now used in order to construct the diversity set S which enables the recovery of $x(t)$ via (34)–(35).

We begin by defining two conceptual set of signals $\mathcal{A}_K, \mathcal{B}_K$. These sets do not represent a valid signal model since their definition involves the unknown size $|S|$. Specifically, an input signal $x(t)$ cannot be determined to lie in either set, when band locations are unknown. We use $\mathcal{A}_K, \mathcal{B}_K$ only for the theoretical analysis of our algorithms. Recovery within $\mathcal{M}$, which parameterizes signals according to their physical properties, is then accomplished based on deriving appropriate relations to $\mathcal{A}_K, \mathcal{B}_K$.

For a given multi-coset system with p sampling channels, SBR4 guarantees perfect reconstruction for every signal in $\mathcal{A}_K$ and every $K \leq \sigma(\mathbf{A})/2$. SBR2 is proved to recover signals from $\mathcal{B}_K$ for the same range of K, except for certain examples which are detailed in the text. Universal patterns, $\sigma(\mathbf{A}) = p$, are assumed throughout this section as they lead to the largest sets $\mathcal{A}_K, \mathcal{B}_K$ for a given sampling rate. We show that $\mathcal{M} \subseteq \mathcal{A}_{2N}$ which implies that SBR4 can recover every signal in $\mathcal{M}$ when $p \geq 4N$. This requirement on p translates to twice the minimal sampling rate of Theorem 1. To improve on this result, it is proved that $\mathcal{M} \subseteq \mathcal{B}_N$, thus recovery from minimal sampling rate is possible with SBR2 up to the above exceptions. Fig. 5 depicts the relation between the various signal classes.

A. The SBR4 Algorithm

Let L be an integer and define the signal class

$$\mathcal{A}_K = \{\mathrm{supp}X(f) \subseteq \mathcal{F} \text{ and } |S| \leq K\} \tag{50}$$

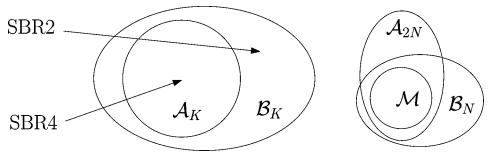

Fig. 5. SBR algorithms are developed to allow recovery within the signal sets $\mathcal{A}_K, \mathcal{B}_K$. The relation to the model $\mathcal{M}$ is drawn on the right.

with S given by (30). In words, $\mathcal{A}_K$ is the class of signals that upon slicing $X(f)$ into L intervals, only K (or less) out of them are nonidentically zero. For brevity, the dependency on the fixed value L is omitted. We point out that $\mathcal{A}_K$ is a set of signals with a specific property and the definition does not enforce multi-coset sampling. It can be verified that the minimal rate requirement for $\mathcal{A}_K$ is $2K/LT$.

Consider a multi-coset sampling with a period L that coincides with the value defining $\mathcal{A}_K$. Let $\mathcal{T} = \mathcal{F}_0$, and observe that $p \geq 2K$ ensures that all the conditions of Proposition 2 are valid for every $x(t) \in \mathcal{A}_K$. Thus, applying the CTF block on $\mathcal{T} = \mathcal{F}_0$ results in a unique sparsest solution $\mathbf{U}_0$, with $S = I(\mathbf{U}_0)$. The reconstruction of the signal is then carried out by (34)–(35). We note that $\mathcal{A}_K \subseteq \mathcal{A}_p$ for $p \geq 2K$ which implies (32).

Algorithm 1, named SBR4, follows the steps of the CTF block in Fig. 4 to recover the diversity set S from $\mathbf{y}(f)$, for any $x(t) \in \mathcal{A}_K$. Clearly, SBR4 guarantees perfect reconstruction of signals in $\mathcal{A}_K$ from samples at the minimal rate $2K/LT$. The algorithm also outputs an indication flag which we discuss later on.

Algorithm 1: SBR4

Input: $\mathbf{y}(f)$, **Assume**: $\sigma(\mathbf{A}) = p$
Output: the set S, flag
1: Set $\mathcal{T} = \mathcal{F}_0$
2: Compute the matrix $\mathbf{Q}$ by (38)
3: Decompose $\mathbf{Q} = \mathbf{V}\mathbf{V}^H$ according to Proposition 3
4: Solve $\mathbf{V} = \mathbf{A}\mathbf{U}$ for the sparsest solution $\mathbf{U}_0$
5: $S = I(\mathbf{U}_0)$
6: flag $= \{|S| \leq p/2\}$
7: **return** S, flag

It is easy to see that flag $= 1$ for every signal in $\mathcal{A}_K$, since SBR4 guarantees perfect reconstruction within this class. However, when a suboptimal algorithm is used to solve the MMV in step 4, exact recovery cannot be guaranteed and SBR4 may end with flag $= 0$. This indication for $x(t) \in \mathcal{A}_K$ means a failure with the particular MMV method. A different MMV approach can subsequently be used instead.

Existing algorithms for MMV systems can be classified into two groups. The first group contains algorithms that seek the sparsest solution matrix $\mathbf{U}_0$, e.g., convex relaxations from the basis pursuit family [13], [30], [31] or matching pursuit [14], [29], [32] with a termination criterion based on the residual. The other group contains methods that approximate a sparse solution according to user specification, e.g., matching pursuit with a predetermined number of iterations [14]. Using a technique from the latter group neutralizes the indication flag as the approximation is always sparse. Therefore, this set of algorithms should be avoided if an indication is desired.

An important advantage of algorithm SBR4 is that the correlation matrix $\mathbf{Q}$ can be computed in the time domain from the known sequences $x_{c_i}[n], 1 \leq i \leq p$ as described in [5]. Reconstruction based on time-domain computations is demonstrated in Section VII-E.

Proposition 1 shows that for $x(t) \in \mathcal{M}$ the set S satisfies $|S| \leq 2N$ if $L \leq 1/BT$. Thus, under this condition on L we have $\mathcal{M} \subseteq \mathcal{A}_{2N}$, which in turn implies $p = 4N$ as a minimal value for p. Consequently, SBR4 guarantees perfect reconstruction for $\mathcal{M}$ when choosing $L \leq 1/BT$ and $p \geq 4N$. The corresponding sampling rate is

$$\frac{p}{LT} \geq \frac{4N}{T\frac{1}{BT}} = 4NB. \tag{51}$$

In contrast, it follows from Theorem 3 that sampling at rate $2NB$ is sufficient for uniqueness of the solution. The reason for the factor of two in the sampling rate is that $\mathbf{x}(f)$ is N-sparse for each specific f; however, when combining the frequencies, the maximal size of S is $2N$. The SBR2 algorithm, developed in the next section, capitalizes on this difference to regain the factor of two in the sampling rate, and thus achieves the minimal rate, at the expense of a more complicated reconstruction.

B. The SBR2 Algorithm

To reduce the sampling rate, we introduce a set $\mathcal{B}_K$ for which SBR2 allows perfect reconstruction if $p \geq 2K$, and then prove that $\mathcal{M} \subseteq \mathcal{B}_N$.

Consider a partition of $\mathcal{F}_0$ into M consecutive intervals defined by

$$0 = \bar{d}_1 < \bar{d}_2 < \cdots < \bar{d}_{M+1} = \frac{1}{LT}. \tag{52}$$

For a given partition $\bar{D} = \{\bar{d}_i\}$ and integer L we define the set

$$\mathcal{B}_{K,\bar{D}} = \{\text{supp}X(f) \subseteq \mathcal{F} \text{ and } |S_{[\bar{d}_i,\bar{d}_{i+1}]}| \leq K, 1 \leq i \leq M\} \tag{53}$$

where the dependency on the fixed value of L is omitted for brevity. Next, We define the set $\mathcal{B}_K$ as

$$\mathcal{B}_K = \bigcup_{\bar{D}} \mathcal{B}_{K,\bar{D}} \tag{54}$$

which is the union of $\mathcal{B}_{K,\bar{D}}$ over all choices of partition sets $\bar{D}$ and integers M. It can be verified that both $\mathcal{B}_{K,\bar{D}}$ and $\mathcal{B}_K$ require minimal sampling rate of $2K/LT$. Clearly, if $p \geq 2K$

then we can perfectly reconstruct every $x(t) \in \mathcal{B}_{K,\bar{D}}$ by applying the CTF block to each of the intervals $[\bar{d}_i, \bar{d}_{i+1}]$. As a consequence, for $x(t) \in \mathcal{B}_K$, if the specific partition $\bar{D}$ which corresponds to $x(t)$ is found by any means, then exact recovery is accomplished within the appropriate set $\mathcal{B}_{K,\bar{D}}$.

The following proposition shows that if the parameters are chosen properly, then $\mathcal{M} \subseteq \mathcal{B}_N$.

Proposition 4: If L, p, C are selected according to Theorem 3 then $\mathcal{M} \subseteq \mathcal{B}_N$.

Proof: In the Proof of Theorem 3 we showed that under the conditions of the theorem, $\mathbf{x}(f)$ is N-sparse for every $f \in \mathcal{F}_0$. The proof then follows from the following lemma.

Lemma [11]: If $x(t)$ is a multiband signal with N bands sampled by a multi-coset system then there exists a partition set $\bar{D} = \{\bar{d}_i\}$ with $M = 2N + 1$ intervals such that $I(\mathbf{x}(f))$ is a constant set over the interval $[\bar{d}_i, \bar{d}_{i+1}]$ for $1 \le i \le M$.

Lemma 2 implies that $|S_{[\bar{d}_i, \bar{d}_{i+1}]}| \le N$ for every $1 \le i \le M = 2N + 1$ which means that $x(t) \in \mathcal{B}_{N,\bar{D}} \subset \mathcal{B}_N$. ■

We now assume that $x(t) \in \mathcal{B}_K \cap \mathcal{A}_{2K}$ and $p \ge 2K$. Observe that $x(t) \in \mathcal{M}$ lies in $\mathcal{B}_N$ according to Proposition 4 and $\mathcal{M} \subseteq \mathcal{A}_{2N}$ from Proposition 1, which motivates this approach. We aim at finding a partition D such that $x(t) \in \mathcal{B}_{K,D}$. The existence of at least one such partition $\bar{D}$ is ensured since $x(t) \in \mathcal{B}_K$. Once D is found,

$$\hat{S} = \bigcup_{i=0}^{|D|-1} S_{[d_i, d_{i+1}]} \tag{55}$$

is equal to S. The CTF block is used to recover the diversity set of each subinterval $[d_i, d_{i+1}]$ in the union (55). Condition (32) is implied by $x(t) \in \mathcal{A}_{2K}$ and thus exact recovery is obtained by (34)–(35). It can be verified that Propositions 2–3 guarantee a unique diversity set for each subinterval of (55).

In order to find D we suggest a bi-section process on $\mathcal{F}_0$. We initialize $D = [0, 1/LT], \mathcal{T} = \mathcal{F}_0$ and seek $S_{\mathcal{T}}$. If $|S_{\mathcal{T}}| > K$, then we halve $\mathcal{T}$ into $\mathcal{T}_1$ and $\mathcal{T}_2$ and recursively determine $S_{\mathcal{T}_1}$ and $S_{\mathcal{T}_2}$. The bi-section process is continued until ultimately each subinterval corresponds to a diversity set of size K at the most. To ensure termination, an interval width of less than ϵ is not halved. The set $\hat{S}$ is then determined according to (55).

Consider an interval $\mathcal{T} \subseteq \mathcal{F}_0$. The matrix $\mathbf{Z}_0$ of (39) satisfies the constraints (42a)–(42b). Since $x(t) \in \mathcal{A}_{2K}$ and $p \ge 2K$ (42c) is also valid. However, the last constraint (42d) of Proposition 2 is not guaranteed as it requires a stronger condition $|S_{\mathcal{T}}| \le K \le p/2$. This condition is satisfied for each subinterval of the correct $\bar{D}$. However, during the execution of SBR2, it is likely to encounter $\mathcal{T}$ with $|S_{\mathcal{T}}| > K$. This typically happens on initialization when $\mathcal{T} = \mathcal{F}_0$. To avoid the unpredicted output of the CTF block in these cases, we detect them in advance. We suggest to approximate $|S_{\mathcal{T}}| = |I(\mathbf{Z}_0)|$ by $\mathrm{rank}(\mathbf{Q})$, and solve the MMV system for the sparsest solution only if $\mathrm{rank}(\mathbf{Q}) \le p/2$. This approximation is motivated by the fact that for any $\mathbf{Z} \succeq 0$ it is true that $\mathrm{rank}(\mathbf{Z}) \le |I(\mathbf{Z})|$. From Proposition 2 we have that $\mathrm{rank}(\mathbf{Z}_0) = \mathrm{rank}(\mathbf{Q})$ which results in

$$\mathrm{rank}(\mathbf{Q}) \le |I(\mathbf{Z}_0)|. \tag{56}$$

However, only special multiband signals result in strict inequality in (56). Therefore, an interval $\mathcal{T}$ that produces $\mathrm{rank}(\mathbf{Q}) > p/2$ is halved. Otherwise, we apply the CTF block for this $\mathcal{T}$ assuming that (56) holds with equality. These reconstruction steps are detailed in Algorithm 2, named SBR2.

Algorithm 2: SBR2

Input: $\mathcal{T}$, **Initialize:** $\mathcal{T} = \mathcal{F}_0$, **Assume:** $\sigma(\mathbf{A}) = p$
Output: a set $\hat{S}$
1: **if** $\lambda(\mathcal{T}) \le \epsilon$ **then**
2: **return** $\hat{S} = \{\}$
3: **end if**
4: Compute the matrix $\mathbf{Q}$ by (38)
5: **if** $\mathrm{rank}(\mathbf{Q}) \le (p/2)$ **then**
6: Decompose $\mathbf{Q} = \mathbf{V}\mathbf{V}^H$
7: Solve $\mathbf{V} = \mathbf{A}\mathbf{U}$ for the sparsest solution $\mathbf{U}_0$
8: $\hat{S} = I(\mathbf{U}_0)$
9: $\mathsf{flag} = \{|\hat{S}| \le (p/2)\}$
10: **end if**
11: **if** $(\mathrm{rank}(\mathbf{Q}) > (p/2))$ **or** $(\mathsf{flag} = 0)$ **then**
12: split $\mathcal{T}$ into two equal width intervals $\mathcal{T}_1, \mathcal{T}_2$
13: $\hat{S}^{(1)} = \mathbf{SBR2}(\mathcal{T}_1)$
14: $\hat{S}^{(2)} = \mathbf{SBR2}(\mathcal{T}_2)$
15: $\hat{S} = \hat{S}^{(1)} \cup \hat{S}^{(2)}$
16: **end if**
17: **return** $\hat{S}$

Two error sources may lead SBR2 to produce $\hat{S} \ne S$ and fail the recovery of $x(t)$. One reason this can happen is if strict inequality holds in (56) for some interval $\mathcal{T}$. In this scenario step 7 is executed even though $\mathbf{Z}_0$ does not satisfy (42d). For example, a signal $x(t)$ with two equal width bands $[a_1, a_1 + W]$ and $[a_2, a_2 + W]$ such that

$$\left\lfloor \frac{a_1}{LT} \right\rfloor = \left\lfloor \frac{a_2}{LT} \right\rfloor = \gamma \tag{57}$$

and $\gamma + W \in \mathcal{F}_0$. If $x(t)$ also satisfies

$$X(f - a_1) = X(f - a_2) \quad \forall f \in [0, W] \tag{58}$$

then it can be verified that $|I(\mathbf{Z}_0)| = 2$ while $\mathrm{rank}(\mathbf{Z}_0) = \mathrm{rank}(\mathbf{Q}) = 1$ on the interval $\mathcal{T} = [\gamma, \gamma + W]$. This is of course a rare special case. Another reason is a signal for which the algorithm reached the termination step 1 for some small enough interval. This scenario can happen if two or more points of $\bar{D}$ reside in an interval width of ϵ. As an empty set $\hat{S}$ is returned for this interval, the final output may be missing some of the elements of S. Clearly, the value of ϵ influences the number of cases of this type. We note that since we do not rely on $D = \bar{D}$ the missing values are typically recovered from other intervals. Fortunately, both these sources of error are very uncommon (in the sense that random signals from $\mathcal{M}$ do not have these exceptional structures with high probability).

So far, we have shown that when using brute-force optimal algorithms for the MMV in Step 7, exact recovery is ensured

TABLE II
SPECTRUM-BLIND RECONSTRUCTION METHODS FOR MULTI-BAND SIGNALS

	WKS theorem	This paper	
Applicability	$\operatorname{supp} X(f) \subseteq \mathcal{F}$	$x(t) \in \mathcal{A}_K$	$x(t) \in \mathcal{B}_K \cap \mathcal{A}_{2K}$
Sampling strategy	Uniform	Multi-coset	Multi-coset
# Uniform sequences	1	$p \geq 2K$	$p \geq 2K$
Implied rate	Nyquist	$2K/LT$	$2K/LT$
Reconstruction method	Ideal low pass	SBR4	SBR2
Fully-blind	Yes	Yes	Yes
Time complexity	constant	1 MMV system	bi-section + finite # of MMV
Indication	No	for $x(t) \in \mathcal{A}_K$ only	No
For signals in $\mathcal{M}$			
Sampling setup	Uniform	Theorem 3, $p \geq 4N$	Theorem 3
Implied rate	Nyquist	$4NB$	$2NB$
Achieve the minimal rate	No	No	Yes
Perfect reconstruction	Yes	Yes	No
Indication	No	Yes	No

for every $x(t) \in \mathcal{B}_K \cap \mathcal{A}_{2K}$ up to the above exceptional examples. Consequently, perfect reconstruction within $\mathcal{M}$ is ensured if $p \geq 2N$ and up to the same exceptions. The choice $p \geq 2N$ allows to reach the minimal rate, in contrast to the requirement $p \geq 4N$ of SBR4.

The analysis of SBR2 becomes more complicated when incorporating suboptimal algorithms. Besides the increased number of iterations, suboptimal MMV solutions contribute another source for errors $\hat{S} \neq S$. Interestingly, we found in simulations that a suboptimal algorithm may fail to recover $\mathbf{U}_0$ for some $\mathcal{T}$, and still succeed for each of its subsections, even though $|S_{\mathcal{T}}| \leq p/2$. We therefore also halves every interval for which the CTF produces an indication $\mathsf{flag} = 0$. Evidently, the CTF indication plays a crucial role in refining the partition D and in compensating for the practical limitations of MMV techniques.

The complexity of SBR2 is dictated by the number of iterations of the bi-section process, which is also affected by the behavior of the MMV algorithm that is used. Numerical experiments in Section VII show that empirically SBR2 converges sufficiently fast for practical usage.

Finally, we emphasize that SBR2 does not provide an indication whether $\hat{S} = S$ even for $x(t) \in \mathcal{M}$ since there is no way to know in advance if $x(t)$ is a signal of the special type that SBR2 cannot recover.

C. Comparison Between SBR4 and SBR2

Table II compares the SBR algorithms of the previous sections. The WKS theorem is also included in the comparison, since sampling at the Nyquist rate allows blind recovery of signals from $\mathcal{M}$. The upper part of the table considers the properties of each method with respect to the applicable signal set. Observe that an indication of correct recovery of S is available only for SBR4 and only if the signals are known to lie in $\mathcal{A}_K$. In addition, since $\mathcal{A}_K$ is a true subset of $\mathcal{B}_K \cap \mathcal{A}_{2K}$, SBR2 is applicable to a wider set of signals for the same average sampling rate.

Considering signals from our model $\mathcal{M}$ allows to design a fully-blind system based on the known model-parameters N, B. The specific behavior for this scenario is compared in the bottom part of Table II. In particular, SBR4 requires twice the minimal rate and guarantees perfect reconstruction when using a brute-force MMV method. It is important to note that suboptimal MMV algorithms typically require an increase in sampling rate to allow recovery.

Fig. 6 summarizes the design flow of a fully blind system. This figure together with Table II can be used in order to decide on the most appropriate implementation of the system. Clearly, for $\Omega > 0.5$ it should be preferred to sample at the Nyquist rate and to reconstruct with an ideal lowpass filter. For $\Omega \leq 0.5$ we have to choose between SBR4 and SBR2 according to the preferred complexity.

The tradeoff presented here between complexity and sampling rate also exists in the known-spectrum reconstruction. Sampling at the minimal rate of Landau requires a reconstruction that consists of piecewise constant filters [11]. The number of pieces and the reconstruction complexity grow with L. This complexity can be prevented by doubling the value of p which also doubles the average sampling rate according to (23). Then, (34)–(35) are used to reconstruct the signal by only one inversion of a known matrix [5].

VII. NUMERICAL EXPERIMENTS

We now provide several experiments in which we examine the following aspects of the reconstruction of $x(t) \in \mathcal{M}$ using algorithms SBR4 and SBR2:

1) recovery for different sampling rates where the MMV produced by the CTF block is solved using either an optimal (exponential time) method or a suboptimal (polynomial time) technique;
2) reconstruction with measurement noise;
3) comparison with a discretization-based method;
4) sampling and reconstruction in the time domain.

A. Setup

To test our results, we would like to sample and perfectly reconstruct a continuous time signal $x(t) \in \mathcal{M}$. However, an experiment of this type requires implementing both stages in hardware. In this section, we describe computer simulations for which several simplifications are required.

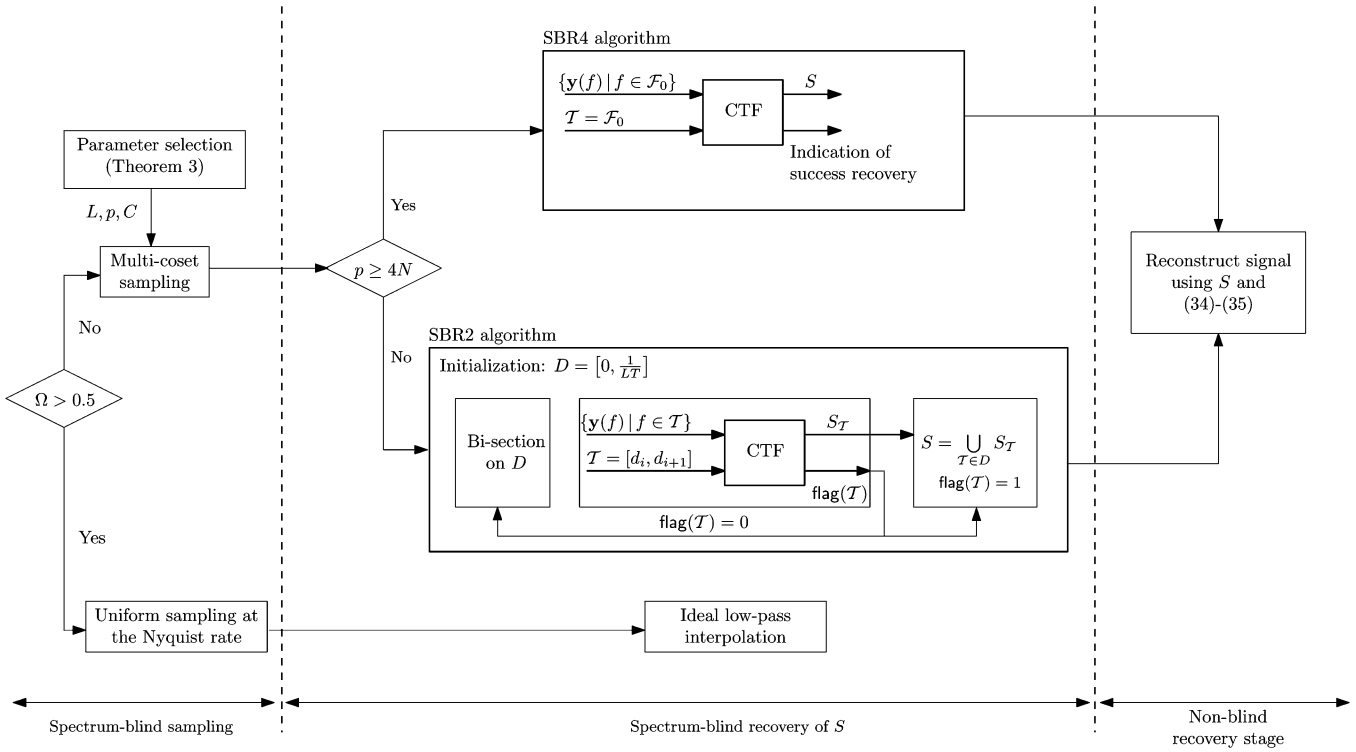

Fig. 6. Spectrum-blind reconstruction scheme.

1) A multiband signal $x(t)\in\mathcal{M}$ has an infinite support in the time-domain as follows from the uncertainty principle. Since only finite length signals can be simulated, we prefer to generate the test case signals directly in the frequency domain, where the support of $X(f)$ is confined to the finite interval $\mathcal{F}$. In the last experiment we qualitatively examine our method on a truncated length signal in the time domain.
2) The interval $\mathcal{F}$ is finite but continuous and thus uncountable. A digital representation of $X(f)$ on $\mathcal{F}$ is possible only over a finite grid. We choose a dense grid to represent the continuous case. Note that the model remains continuous, and only the experiments necessitate discretization. To clarify, we compare our approach with a method that also discretizes the model.
3) In practice, only a finite (though extensive) set of signals from $\mathcal{M}$ can be simulated. We choose to simulate random realizations of $x(t)\in\mathcal{M}$.

The setup described hereafter is used as a basis for all the experiments. The parameter values are chosen differently for each simulation.

Consider an example of the class $\mathcal{M}$ where $\mathcal{F}=[0,20\text{ GHz}]$ with exactly N bands of width B each. In order to evaluate the reconstruction of our algorithms, 500 test cases from this class were generated randomly according to the following steps:

1) draw $\{a_i\}_{i=1}^N$ uniformly at random from $[0,20\text{ GHz}-B]$;
2) set $b_i=a_i+B$ for $1\leq i\leq N$, and ensure that the bands do not overlap;
3) generate an equal-spaced grid over $\mathcal{F}$ whose density is 100 grid points per interval length B;
4) generate $X(f)$ on the chosen grid by

$$X(f)=\begin{cases}E_i(A_R(f)+jA_I(f)), & f\in\bigcup_{i=1}^N[a_i,b_i]\\ 0, & \text{otherwise.}\end{cases}$$

For every f, the values of $A_R(f)$ and $A_I(f)$ are drawn independently from a normal distribution with zero mean and unit variance. The constants E_i are chosen uniformly from the interval [1,5], independently for each band.

A single multi-coset system is designed by selecting the values of L,p,C according to Theorem 3. The multi-coset system is not changed during the experiment.

An experiment is conducted by recovering the unknown vectors $\mathbf{x}(f)$ from the measurement vectors $\mathbf{y}(f)$ of (20) over the chosen grid. To simulate different sampling rates, SBR4 and SBR2 are executed using only the first $r\leq p$ cosets (while ignoring the rest). The reconstruction is initially carried out for a small value of r and then repeated while gradually increasing r until the maximal value of p. For each value of r, the empirical success rate is calculated as the ratio of simulations in which the signal $X(f)$ was recovered up to machine precision over the chosen grid. The empirical recovery rate reported in our results is only an average measure of recovery success of a *typical* $x(t)\in\mathcal{M}$. The reader is referred to [24] and [28] for further discussion on this issue. In addition, we collect run time data to qualitatively compare the reconstruction complexity.

B. Recovery for Different MMV Solvers

We set $N=3$ and $B=1.05$ GHZ and select the largest possible value for L satisfying the requirement of Theorem 3:

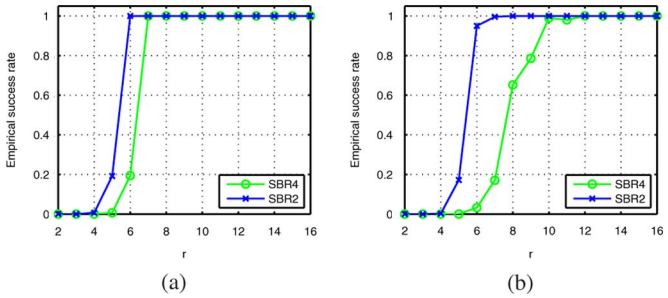

Fig. 7. Empirical recovery rate for different sampling rates when the MMV solution is obtained (a) by a brute-force method or (b) by the M-OMP algorithm.

$$L = 19 \leq \frac{1}{BT} = 19.04. \tag{59}$$

In addition, we choose $p = 16 \geq 2N = 6$. The sampling pattern C is selected uniformly at random among the $\binom{19}{16}$ combinations. Since L is prime, C is guaranteed to be universal [26]. Therefore, for $6 \leq r \leq p$ a unique solution is guaranteed by Theorem 3. Note that in this setting, $r = 6$ corresponds to a sampling rate that is slightly more than the minimal rate of Theorem 1, i.e., $6/(LT) = 6.315$ GHz $> 2NB = 6.3$ GHz.

Fig. 7 depicts the empirical recovery rate for $2 \leq r \leq p$ and for two selections of an MMV solver. One is a brute-force method [13] and the other uses multi-orthogonal matching pursuit (M-OMP) [14]. Recall that $|S| \leq 2N$ from Proposition 1. However, in this experiment a typical signal has $|S| = 2N = 6$ since $1/LT \approx B$. As the results show, for SBR4 a high-recovery rate (empirically 100%) is achieved with $r \geq |S| + 1$ for the brute-force solver, or with $r \geq 2|S| = 12$ for M-OMP. These thresholds conform with known results regarding these recovery methods [14], [28].

The average run time is reported in Fig. 8, in which it is apparent that recovery based on M-OMP is faster. For M-OMP, the run time of SBR2 is approximately one order above SBR4 in the range $6 \leq r \leq 9$. The reason is that in this range the recovery rate of M-OMP is low (as the recovery curve of SBR4 testifies). Consequently, the recursion depth of SBR2 grows as it is difficult to find a suitable partition set D. For the brute-force solver, the number of rows in $\mathbf{U}_0$ that are nonidentically zero dictates the run time, thus SBR2 is sometimes even faster than SBR4 for this solver.

C. Performance With Measurement Noise

Algorithms SBR4 and SBR2 are readily generalized to deal with noisy measurements by ignoring the noise eigenvectors in $\mathbf{Q}$. Specifically, in step 3 of SBR4 the matrix $\mathbf{V}$ is chosen as the set of eigenvectors of $\mathbf{Q}$ whose eigenvalues are above a certain threshold ($5 \cdot 10^{-4}$ is used in the simulation). While this method may not be optimal for noise rejection, it is quite simple. A detailed analysis of methods for noise suppression is beyond the scope of this paper. For SBR2, the same approach is used for computing $\mathrm{rank}(\mathbf{Q})$ in step 5 and for the decomposition in step 6.

To simulate the noisy measurements, we create a noise signal $N(f)$ over the entire grid, such that both real and imaginary parts of $N(f)$ are drawn independently from a Gaussian distribution with zero mean and unit variance. The noise energy $\|N(f)\|$ is scaled to a desired signal-to-noise ratio, defined as $\mathrm{SNR(dB)} = 10\log_{10}(\|X(f)\|/\|N(f)\|)$. The empirical recovery rate is reported as the fraction of simulations in which the set S of the noise-free $X(f)$ was recovered correctly. The results, depicted in Fig. 9, show a similar behavior for both algorithms, while SBR2 is slightly more sensitive since failures from subintervals may propagate to the final recovered support. Here we used only M-OMP.

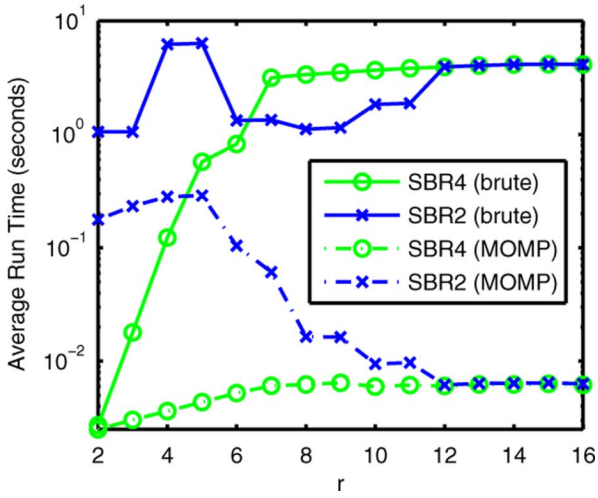

Fig. 8. Average run time for SBR4 and SBR2 with different MMV solvers.

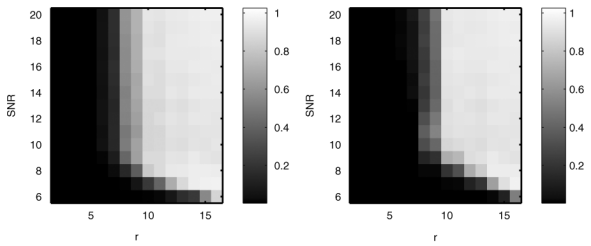

Fig. 9. Image intensity represents the empirical recovery rate for different sampling rates r and for several SNR ratios. The setup of this experiment is the same as of the first experiment, i.e., $N = 3, B = 1.05$ GHz, $L = 19, p = 16$.

The stability of SBR4 and SBR2 is dictated by the underlying CS algorithms that is used to recover $\mathbf{U}_0$. Consequently, a robust algorithm, such as M-OMP which was studied in [29], implies robustness of the entire recovery procedure. This demonstrates that formulating the recovery problem within the CS framework, as implied by Proposition 3, is beneficial since other robust methods can be adapted, e.g., [15], [24], and [33].

D. Comparison With Discretization

We now describe a discretization approach for blind recovery and compare it with SBR4 and SBR2. In the computer simulations our approach is also performed over a discrete grid of $\mathcal{F}_0$. We show that our method outperforms discretization in terms of recovery rate and run time in simulations. In addition, we detail the prominent benefits of SBR4 and SBR2 when reconstruction is carried out in hardware.

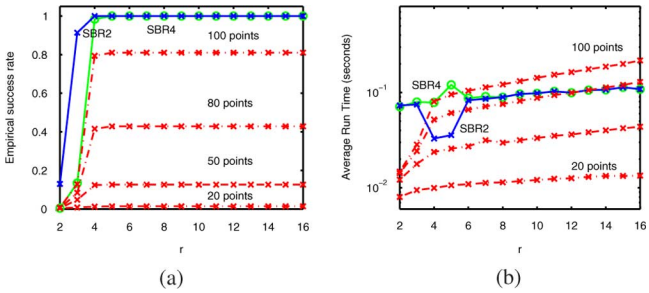

Fig. 10. Comparison with discretization: (a) Empirical recovery rate and (b) run time for brute-force solvers.

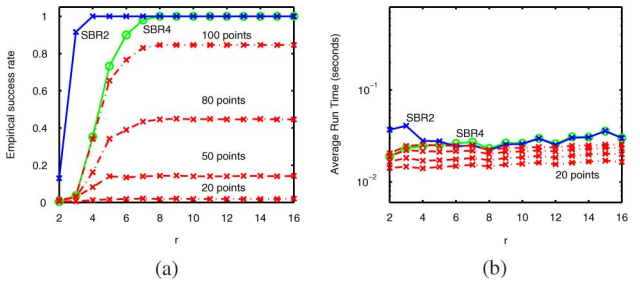

Fig. 11. Comparison with discretization: (a) Empirical recovery rate and (b) run time for M-OMP solvers.

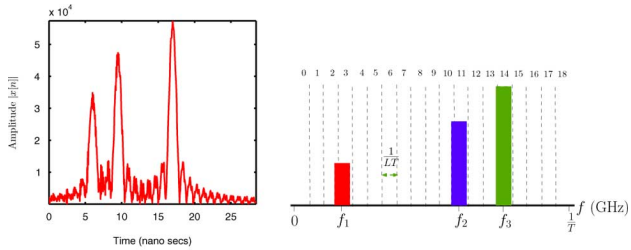

Fig. 12. Amplitude of the Nyquist rate sequence $x[n] = x(nT)$ is plotted in (a) for $0 \le n \le 569$. In this example $E_i = \{1, 2, 3\}$ and $t_i = \{6, 9.5, 17\}$ ns. The support of $X(f)$ is centered around $f_i = \{3.4, 12, 15.2\}$ GHz as shown in (b). The diversity set is $S = \{2, 3, 10, 11, 13, 14\}$.

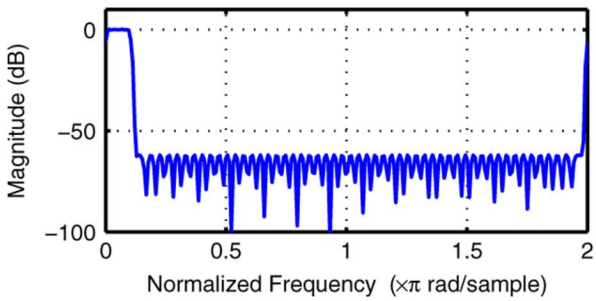

Fig. 13. Frequency response of $h[n]$ (magnitude only) for $N_{\mathrm{TAPS}} = 191$.

A naïve approach to overcome the continuous nature of (20) is to discretize the frequency interval $\tau = \mathcal{F}_0$ to an equally spaced finite grid $\tilde{\tau} = \{f_i\}_{i=1}^{G} \subset \mathcal{F}_0$ and to solve only for $\mathbf{x}(f_i)$. The parameter selection of Theorem 3 guarantees uniqueness of the N-sparse vectors $\mathbf{x}(f_i)$. The resulting finite dimensional problem can be solved within the regular CS framework. The missing values of $\mathbf{x}(f)$ between the grid points, that is for $f \in (f_i, f_{i+1})$, can be interpolated if additional smoothness assumptions on $X(f)$ are used. Alternatively, we can approximate the set S of (30) by $\tilde{S} = \cup_{i=1}^{G} I(\mathbf{x}(f_i))$ and then apply (34)–(35). Since our model and experiments do not assume smoothness, we compare algorithms SBR4 and SBR2 with the latter method.

Consider a computer experiment with $B = 10$ MHz and N, L, p as before. This selection produces multiband signals with narrow bands (relative to the slice width $1/LT$). In Figs. 10 and 11, we compare recovery rates and run times of SBR4, SBR2 with discretization. Two significant differences between these approaches are evident. First, for a comparable run time, SBR4 and SBR2 achieve higher recovery rate. Second, to improve the recovery rate of discretization, a dense grid of $G > 100$ is required which increases run time beyond that of SBR4, SBR2. A dramatic increase of G is typically required when B is only an upper-bound and not the exact band width as in simulations. We note that recovery rate of discretization decreases for G higher than few hundreds since machine precision limits the matrix dimension that can be used in matching pursuit. In contrast, 200 000 frequency bins are used in this simulation for SBR4 and SBR2. Evidently, a high frequency resolution hardly affects the recovery and run time of our SBR algorithms.

Besides improved performance in simulations, our approach has obvious advantages in hardware implementation. Discretization requires working in the frequency domain and necessitates additional assumptions on the content of $X(f)$ to approximate a solution. For SBR4, $\mathbf{Q}$ can be computed in analog hardware and in the time-domain (as discussed in the sequel). Moreover, in noisy environments discretization is inherently problematic since the noise amplitude can be locally higher than the signal and typically the recovered support is incorrect. In contrast, the SBR algorithms are based on the aggregated energy of the signal in each spectrum slice and thus are more effective in noisy environments. This distinction also holds for computer simulations with measurement noise.

E. Sampling and Reconstruction in the Time-Domain

Consider the following choice of $x(t) \in \mathcal{M}$ with $N = 3, B = 1.05$ GHz, $1/T = 20$ GHz:

$$x(t) = \sum_{i=1}^{N} \sqrt{E_i B_i}\,\mathrm{sinc}(B(t - t_i))\exp(j2\pi f_i t) \qquad (60)$$

where $\mathrm{sinc}(x) = \sin(\pi x)/(\pi x)$. The ith band of $x(t)$ defined in (60) has L_2-energy E_i, exact support width B, time offset t_i and carrier frequency f_i. The exact values appear in Fig. 12. The signal is observed on a time frame of 28.5 ns starting at $t = 0$, which corresponds to $28.5 \cdot 10^{-9}/T = 570$ Nyquist rate samples.

In order to sample $x(t)$, we set $L = 19, p = 8 \ge 2N = 6$ and a random sampling pattern C. The sampling sequences $x_{c_i}[n]$, given by (17), are padded by $L - 1$ zeros between the nonzero samples. Interpolating values for the $L - 1$ zeros is carried out by a convolution [5]

$$\tilde{x}_{c_i}[n] = x_{c_i}[n] * h[n] \qquad (61)$$

with an appropriate lowpass filter $h[n]$. We used Matlab command $\mathsf{fircls1}(N_{\mathrm{TAPS}}, 1/L, 0.02, 0.008)$ to create a

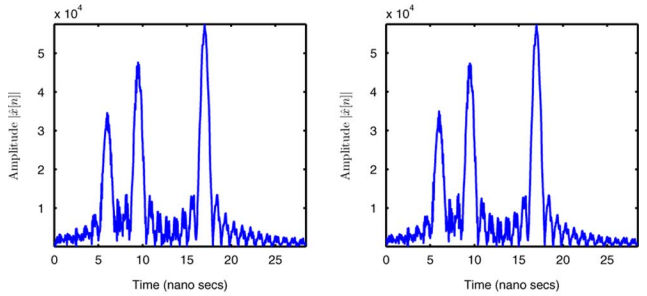

Fig. 14. Amplitude of the reconstructed Nyquist rate sequence $\hat{x}[n]$ is plotted for time points $t = nT, 0 \leq n \leq 569$ and for (a) $N_{\mathrm{TAPS}} = 191$ and (b) $N_{\mathrm{TAPS}} = 381$. The relative reconstruction error $\|x[n] - \hat{x}[n]\|/\|x = x[n]\|$ is 2:5%; 1:7% respectively.

real-valued (nonideal) low-pass filter $h_r[n]$ of length N_{TAPS}, which is frequency-shifted to obtain the complex filter $h[n] = h_r[n]\exp(j\pi n/L)$. Fig. 13 depicts the frequency response of $h[n]$.

To recover the diversity set S, we follow the steps of Fig. 4, where $\mathbf{Q}$ is computed in the time-domain by [5]

$$\mathbf{Q}_{ik} = \langle \tilde{x}_{c_i}[n], \tilde{x}_{c_k}[n] \rangle = \sum_{n=-\infty}^{+\infty} \tilde{x}^*_{c_i}[n]\tilde{x}_{c_k}[n]. \tag{62}$$

In our experiment, the infinite sum is truncated to $0 \leq n \leq 569$, thus producing an approximation $\tilde{\mathbf{Q}}$ with several negligible eigenvalues. The corresponding eigenspaces are removed from $\tilde{\mathbf{Q}}$ as in Section VII-C. Applying the CTF block to the dominant eigenvectors of $\tilde{\mathbf{Q}}$ outputs $\hat{S} = \{2, 3, 10, 11, 13, 14\}$ which is the exact diversity set of $x(t)$.

Once S is recovered, the remaining steps coincide with the known-spectrum recovery of [5]. Specifically, $\mathbf{A}_S$ is inverted and the sampling sequences $x_{c_i}[n]$ are combined accordingly. The reconstruction of the Nyquist rate sequence $\hat{x}[n]$ requires another lowpass filtering [5]. We used the nonideal filter $h[n]$ for this purpose. Fig. 14 depicts the recovered signal for two selections of N_{TAPS}. Additional analog lowpass filter is required to obtain the continuous-time $\hat{x}(t)$.

VIII. Conclusion

In this paper, we suggested a method to reconstruct a multiband signal from its samples when the band locations are unknown. Our development enables a fully spectrum-blind system where both the sampling and the reconstruction stages do not require knowledge of the spectral support.

We address the fundamental aspects of multiband sampling: a signal model, a minimal rate requirement and explicit sampling and reconstruction stages. The proposed signal model is parameterized by the number of bands and their widths and relates to known physical properties of these signals in practical applications. We do not assume the spectral support nor any restriction on the possible band locations.

Our main contributions describe the setup of the sampling stage and concrete blind reconstruction algorithms. We derive a clear connection between signal model and the choice of parameters for the multi-coset strategy. A unique solution is guaranteed under this choice without any further mathematical restrictions on the band locations. The blind recovery is accomplished by a finite dimensional problem which is formulated within the framework of compressed sensing. This result makes no use of discretization. The SBR algorithms developed here suggest a tradeoff between recovery complexity and sampling rate.

In addition, we proved a lower bound on the sampling rate for spectrum-blind reconstruction. One of the algorithms we proposed indeed approaches this minimal rate for a wide class of multiband signals from our model.

Numerical experiments demonstrated various practical aspects regarding the blind recovery, such as the influence of suboptimal CS algorithms, stability in noisy environment and a comparison with discretization techniques.

The machinery developed in this work extends compressed sensing beyond the finite framework on which it has been focused. In particular, multiband signals are only a representative example for sparsity of analog signals. Subsequent publications since submission of this paper already incorporate this approach to develop low-rate sampling strategies for other classes of sparse analog signals.

Appendix A
Real-Valued Signals

In order to treat real-valued signals the following definitions replace the ones given in the paper. The class $\mathcal{M}$ is changed to contain all real-valued multiband signals bandlimited to $\mathcal{F} = [-1/2T, 1/2T]$ with no more than N bands, where each band width is upper bounded by B as before. The Fourier transform is conjugate symmetric for real-valued signals, and thus N is always even. The Nyquist rate remains $1/T$ and the minimal rate requirement is $2NB$ (if less than the Nyquist rate).

Repeating the calculations of [11] that lead to (18), it can be seen that several modifications are required. To form $\mathbf{x}(f)$, the interval $\mathcal{F}$ is still divided into L equal intervals. However, a slightly different treatment is given for odd and even values of L, because of the aliasing of positive and negative frequencies. Define the set of L consecutive integers:

$$K = \begin{cases} \left\{-\frac{L-1}{2}, \ldots, \frac{L-1}{2}\right\}, & \text{odd } L \\ \left\{-\frac{L}{2}, \ldots, \frac{L}{2} - 1\right\}, & \text{even } L \end{cases} \tag{63}$$

and redefine the interval $\mathcal{F}_0$ by

$$\mathcal{F}_0 = \begin{cases} \left[-\frac{1}{2LT}, \frac{1}{2LT}\right], & \text{odd } L \\ \left[0, \frac{1}{LT}\right], & \text{even } L. \end{cases} \tag{64}$$

The vector $\mathbf{x}(f)$ is then given by

$$\mathbf{x}_i(f) = X(f + K_i/LT), \quad 0 \leq i \leq L-1, f \in \mathcal{F}_0. \tag{65}$$

The dimensions of $\mathbf{A}$ remain $p \times L$ with ikth entry

$$\mathbf{A}_{ik} = \frac{1}{LT}\exp\left(j\frac{2\pi}{L}c_i K_k\right). \tag{66}$$

The definition of $\mathbf{y}(f)$ remains the same with respect to $\mathcal{F}_0$ defined in (64). The results of the paper are thus extended to real-valued multiband signals since (20) is now valid with respect to these definitions of $\mathbf{x}(f)$, $\mathbf{A}$, and $\mathcal{F}_0$.

Note that, filtering out the negative frequencies of the given real-valued signal $x(t)$ results in a complex-valued version

whose support is $NB/2$. Treating the latter in the regular complex-valued framework leads to a Nyquist rate $1/2T$ and also halves the minimal rate requirement. Nevertheless, the information rate is the same for both versions, as each sample of a complex-valued signal is represented by two real numbers.

APPENDIX B
PROOF OF LEMMA 1

Let $r = \mathrm{rank}(\mathbf{P})$. Reorder the columns of $\mathbf{P}$ so that the first r columns are linearly independent. This operation does not change the rank of $\mathbf{P}$ nor the rank of $\mathbf{AP}$. Define

$$\mathbf{P} = [\mathbf{P}^{(1)} \quad \mathbf{P}^{(2)}] \tag{67}$$

where $\mathbf{P}^{(1)}$ contains the first r columns of $\mathbf{P}$ and the rest are contained in $\mathbf{P}^{(2)}$. Therefore,

$$r \geq \mathrm{rank}(\mathbf{AP}) = \mathrm{rank}\Big(\mathbf{A}\left[\mathbf{P}^{(1)} \quad \mathbf{P}^{(2)}\right]\Big) \geq \mathrm{rank}\Big(\mathbf{AP}^{(1)}\Big). \tag{68}$$

The inequalities result from the properties of the rank of concatenation and of multiplication of matrices. So it is sufficient to prove that $\mathbf{AP}^{(1)}$ has full column rank.

Let $\boldsymbol{\alpha}$ be a vector such that $\mathbf{AP}^{(1)}\boldsymbol{\alpha} = \mathbf{0}$. It remains to prove that this implies $\boldsymbol{\alpha} = \mathbf{0}$. Since $I(\mathbf{P}^{(1)}) \subseteq I(\mathbf{P})$ the vector $\mathbf{P}^{(1)}\boldsymbol{\alpha}$ is k-sparse for $k = |I(\mathbf{P})|$. However, $\sigma(\mathbf{A}) \geq k$ and its null space cannot contain a k-sparse vector unless it is the zero vector. Since $\mathbf{P}^{(1)}$ contains linearly independent columns this implies $\boldsymbol{\alpha} = \mathbf{0}$.

ACKNOWLEDGMENT

The authors would like to thank Prof. M. Horowitz for fruitful discussions. The authors also acknowledge conversations with Prof. E. Candès regarding stability aspects of nonlinear reconstruction methods. Finally, the authors would like to thank the anonymous referees and the past Editor-in-Chief Prof. A.-J. van der Veen, for their constructive comments which helped improve the quality of the paper.

During the review of this paper, we were provided with the thesis [9]. This work, which unfortunately was never cited and is available only on microfilms, follows the ideas of [5]–[7]. The current paper differs from [9] at the early step of defining the signal model. Consequently, our theorems and algorithms do not require the restriction on band locations of [7], which also appears in [9].

REFERENCES

[1] H. J. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Math.*, vol. 117, pp. 37–52, Feb. 1967.

[2] R. G. Vaughan, N. L. Scott, and D. R. White, "The theory of bandpass sampling," *IEEE Trans. Signal Process.*, vol. 39, no. 9, pp. 1973–1984, Sep. 1991.

[3] A. Kohlenberg, "Exact interpolation of band-limited functions," *J. Appl. Phys.*, vol. 24, pp. 1432–1435, Dec. 1953.

[4] Y.-P. Lin and P. P. Vaidyanathan, "Periodically nonuniform sampling of bandpass signals," *IEEE Trans. Circuits Syst. II*, vol. 45, no. 3, pp. 340–351, Mar. 1998.

[5] P. Feng and Y. Bresler, "Spectrum-blind minimum-rate sampling and reconstruction of multiband signals," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing (ICASSP)*, May 1996, vol. 3, pp. 1688–1691.

[6] Y. Bresler and P. Feng, "Spectrum-blind minimum-rate sampling and reconstruction of 2-d multiband signals," in *Proc. IEEE Int. Conf. Image Process.*, Sep. 1996, vol. 1, pp. 701–704.

[7] R. Venkataramani and Y. Bresler, "Further results on spectrum blind sampling of 2D signals," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 4–7, 1998, vol. 2, pp. 752–756.

[8] R. Schmidt, "Multiple emitter location and signal parameter estimation," *IEEE Trans. Antennas Propagat.*, vol. 34, no. 3, pp. 276–280, Mar. 1986.

[9] P. Feng, "Universal minimum-rate sampling and spectrum-blind reconstruction for multiband signals," Ph.D. dissertation, University of Illinois at Urbana-Champaign, Urbana–Champaign, IL, 1997.

[10] C. Herley and P. W. Wong, "Minimum rate sampling and reconstruction of signals with arbitrary frequency support," *IEEE Trans. Inf. Theory*, vol. 45, no. 5, pp. 1555–1564, Jul. 1999.

[11] R. Venkataramani and Y. Bresler, "Perfect reconstruction formulas and bounds on aliasing error in sub-Nyquist nonuniform sampling of multiband signals," *IEEE Trans. Inf. Theory*, vol. 46, no. 6, pp. 2173–2183, Sep. 2000.

[12] Y. M. Lu and M. N. Do, "A theory for sampling signals from a union of subspaces," *IEEE Trans. Signal Process.*, vol. 56, no. 6, pp. 2334–2345, Jun. 2007.

[13] J. Chen and X. Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *IEEE Trans. Signal Process.*, vol. 54, no. 12, pp. 4634–4643, Dec. 2006.

[14] S. F. Cotter, B. D. Rao, K. Engan, and K. Kreutz-Delgado, "Sparse solutions to linear inverse problems with multiple measurement vectors," *IEEE Trans. Signal Process.*, vol. 53, no. 7, pp. 2477–2488, Jul. 2005.

[15] D. L. Donoho, "Compressed sensing," *IEEE Trans. Inf. Theory*, vol. 52, no. 4, pp. 1289–1306, Apr. 2006.

[16] J. A. Tropp, M. B. Wakin, M. F. Duarte, D. Baron, and R. G. Baraniuk, "Random filters for compressive sampling and reconstruction," in *Proc. IEEE Int. Conf. Acoutics, Speech, Signal Processing (ICASSP)*, May 2006, vol. 3, p. III.

[17] J. N. Laska, S. Kirolos, M. F. Duarte, T. S. Ragheb, R. G. Baraniuk, and Y. Massoud, "Theory and implementation of an analog-to-information converter using random demodulation," in *Proc. IEEE Int. Symp. Circuits System*, May 2007, pp. 1959–1962.

[18] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," Electr. Eng. Dept., Technion—Israel Institute of Technology, Haifa, CCIT Rep. No. 639, Sep. 2007.

[19] M. Mishali and Y. C. Eldar, "Spectrum-blind reconstruction of multiband signals," in *Proc. IEEE Int. Conf. Acoustics, Speech, Signal Processing, 2008 (ICASSP)*, pp. 3365–3368.

[20] M. Mishali and Y. C. Eldar, "Reduce and boost: Recovering arbitrary sets of jointly sparse vectors," *IEEE Trans. Signal Process.*, vol. 56, no. 10, pp. 4692–4702, Oct. 2008.

[21] Y. C. Eldar, "Compressed sensing of analog signals," *IEEE Trans. Signal Process.*, Jun. 2008, arXiv.org 0806.3332, submitted for publication.

[22] Y. C. Eldar and M. Mishali, "Robust recovery of signals from a union of subspaces," *IEEE Trans. Inf. Theory*, Jul. 2008, arXiv.org 0807.4581, submitted for publication.

[23] Y. C. Eldar, "Uncertainty relations for analog signals," *IEEE Trans. Inf. Theory.*, arXiv.org 0809.3731, submitted for publication.

[24] E. J. Candès, J. Romberg, and T. Tao, "Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information," *IEEE Trans. Inf. Theory*, vol. 52, no. 2, pp. 489–509, Feb. 2006.

[25] J. B. Kruskal, "Three-way arrays: Rank and uniqueness of trilinear decompositions, with application to arithmetic complexity and statistics," *Linear Algebra Appl.*, vol. 18, no. 2, pp. 95–138, 1977.

[26] T. Tao, "An uncertainty principle for cyclic groups of prime order," *Math. Res. Lett.*, vol. 12, no. 1, pp. 121–127, 2005.

[27] D. L. Donoho and M. Elad, "Maximal sparsity representation via $\ell 1$ minimization," *Proc. Natl. Acad. Sci.*, vol. 100, pp. 2197–2202, Mar. 2003.

[28] D. Baron, M. B. Wakin, M. F. Duarte, S. Sarvotham, and R. G. Baraniuk, Distributed Compressed Sensing, 2005 [Online]. Available: http://www.dsp.ece.rice.edu/cs/DCS112005.pdf
[29] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Process. (Special Issue on Sparse Approximations in Signal and Image Processing)*, vol. 86, pp. 572–588, Apr. 2006.
[30] J. A. Tropp, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Process. (Special Issue on Sparse Approximations in Signal and Image Processing)*, vol. 86, pp. 589–602, Apr. 2006.
[31] S. S. Chen, D. L. Donoho, and M. A. Saunders, "Atomic decomposition by basis pursuit," *SIAM J. Sci. Comput.*, vol. 20, no. 1, pp. 33–61, 1999.
[32] S. G. Mallat and Z. Zhang, "Matching pursuits with time-frequency dictionaries," *IEEE Trans. Signal Process.*, vol. 41, no. 12, pp. 3397–3415, Dec. 1993.
[33] E. J. Candès, J. Romberg, and T. Tao, "Stable signal recovery from incomplete and inaccurate measurements," *Comm. Pure Appl. Math.*, vol. 59, no. 8, pp. 1207–1223, Mar. 2006.

Moshe Mishali (S'07) received the B.Sc. degree in electrical engineering (*summa cum laude*) from the Technion—Israel Institute of Technology, Haifa, in 2000, where he is currently working towards the Ph.D. degree in electrical engineering.

From 1996 to 2000, he was a member of the Technion Program for Exceptionally Gifted Students. Since 2006, he has been a Research Assistant and Project Supervisor with the Signal and Image Processing Lab, Electrical Engineering Department, Technion. His research interests include theoretical aspects of signal processing, information theory, and statistical signal processing.

Mr. Mishali received the Hershel Rich Innovation Award in 2008.

Yonina C. Eldar (S'98-M'02-SM'07) received the B.Sc. degree in physics in 1995 and the B.Sc. degree in electrical engineering in 1996 both from Tel-Aviv University (TAU), Tel-Aviv, Israel, and the Ph.D. degree in electrical engineering and computer science from the Massachusetts Institute of Technology (MIT), Cambridge, in 2001.

From January 2002 to July 2002, she was a Postdoctoral Fellow at the Digital Signal Processing Group at MIT. She is currently an Associate Professor in the Department of Electrical Engineering at the Technion—Israel Institute of Technology, Haifa, Israel. She is also a Research Affiliate with the Research Laboratory of Electronics at MIT. Her research interests are in the general areas of signal processing, statistical signal processing, and computational biology.

Dr. Eldar was in the program for outstanding students at TAU from 1992 to 1996. In 1998, she held the Rosenblith Fellowship for study in Electrical Engineering at MIT, and in 2000, she held an IBM Research Fellowship. From 2002 to 2005, she was a Horev Fellow of the Leaders in Science and Technology program at the Technion and an Alon Fellow. In 2004, she was awarded the Wolf Foundation Krill Prize for Excellence in Scientific Research, in 2005 the Andre and Bella Meyer Lectureship, in 2007 the Henry Taub Prize for Excellence in Research, and in 2008 the Hershel Rich Innovation Award, the Award for Women with Distinguished Contributions, and the Muriel & David Jacknow Award for Excellence in Teaching. She is a member of the IEEE Signal Processing Theory and Methods technical committee and the Bio Imaging Signal Processing technical committee, an Associate Editor for the IEEE TRANSACTIONS ON SIGNAL PROCESSING, the *EURASIP Journal of Signal Processing*, and the *SIAM Journal on Matrix Analysis and Applications*, and on the Editorial Board of *Foundations and Trends in Signal Processing*.