Reconciling Compressive Sampling Systems for Spectrally Sparse Continuous-Time Signals

Citation for published version:
Lexa, MA, Davies, ME & Thompson, JS 2012, 'Reconciling Compressive Sampling Systems for Spectrally Sparse Continuous-Time Signals' IEEE Transactions on Signal Processing, vol 60, no. 1, pp. 155-171. DOI: 10.1109/TSP.2011.2169408

Digital Object Identifier (DOI):
10.1109/TSP.2011.2169408

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
IEEE Transactions on Signal Processing

Reconciling Compressive Sampling Systems for Spectrally-sparse Continuous-time Signals

Michael A. Lexa* *Member, IEEE*, Mike E. Davies *Member, IEEE* and
John S. Thompson *Member, IEEE*

Abstract

The Random Demodulator (RD) and the Modulated Wideband Converter (MWC) are two recently proposed compressed sensing (CS) techniques for the acquisition of continuous-time spectrally-sparse signals. They extend the standard CS paradigm from sampling discrete, finite dimensional signals to sampling continuous and possibly infinite dimensional ones, and thus establish the ability to capture these signals at sub-Nyquist sampling rates. The RD and the MWC have remarkably similar structures (similar block diagrams), but their reconstruction algorithms and signal models strongly differ. To date, few results exist that compare these systems, and owing to the potential impacts they could have on spectral estimation in applications like electromagnetic scanning and cognitive radio, we more fully investigate their relationship in this paper. Specifically, we show that the RD and the MWC are both based on the general concept of random filtering, but that the sampling functions characterising the systems differ significantly. We next demonstrate a previously unreported model sensitivity the MWC has to short duration signals that resembles the known sensitivity the RD has to nonharmonic tones. We also show that block convolution is a fundamental aspect of the MWC, allowing it to successfully sample and reconstruct block-sparse (multiband) signals. This aspect is lacking in the RD, but we use it to propose a new CS based acquisition system for continuous-time signals whose amplitudes are block sparse. The paper includes detailed time and frequency domain analyses of the RD and the MWC that differ, sometimes substantially, from published results.

I. Introduction

The theory of compressed sensing (CS) says that if a signal is sufficiently sparse with respect to some basis or frame, then it can be faithfully reconstructed from a small set of linear, nonadaptive measurements, even if the support of the signal is unknown [1]–[3]. When the signal belongs to a finite

The authors are with the Institute of Digital Communications (IDCOM) & the Joint Research Institute for Signal and Image Processing, School of Engineering, The University of Edinburgh, Alexander Graham Bell Building, The King's Buildings, Mayfield Road, Edinburgh EH9 2JL, United Kingdom; Phone: ML +44 (0)131-650-5659, MD +44 (0)131-650-5795, JT +44 (0)131-650-5585; Fax: +44 (0)131-650-6554; Email: {michael.lexa, mike.davies, john.thompson}@ed.ac.uk.

dimensional space, this statement means that it can be reconstructed from a set of measurements whose cardinality may be significantly less than the space's dimension. It also implies that the measurement process is described by an underdetermined linear system of equations, or equivalently, a rectangular matrix with more columns than rows. The fundamental work of Candes, Romberg, and Tao [4] and Donoho [1] established sufficient conditions upon such sensing matrices, that if satisfied, allow the stable inversion of the linear system. A key aspect of CS, and one which plays an important role in this paper, is that sensing matrices drawn at random[1] often satisfy these conditions.

Conceptually, CS theory has three main thrusts: (1) the development of recovery methods that efficiently and faithfully reconstruct the original signal from its compressed samples, (2) the investigation of new signal models that effectively represent signal sparsity or other signal structure, and (3) the creation of new sampling (measurement) mechanisms that acquire signals in a compressed manner. All three lines of research are intertwined and all need to be considered when designing a sampling system. The first concerns the reconstruction process and asks how one specifically reconstructs the original signal from the CS measurements (see, e.g., [1], [4], [5]). The second concerns the examination of different signal classes of interest and asks if there exists a structured representation that can be exploited [6]–[8]. The third concerns the design of the physical sampling system and asks how one devises a system to acquire CS measurements [7], [9]–[11]. This paper concerns sampling systems and signal models.

Several CS based signal acquisition systems have been proposed for both continuous (analogue) and discrete signals. For example, the single-pixel camera [12] is a novel compressive imaging system, where light is projected onto a random basis using a micro-mirror device, and then the projected image is captured by a single photo-diode (the single "pixel"). Other examples include random filtering [13] and random convolution [9] that advocate random linear filtering and low rate sampling as a means to collect CS measurements. In these cases, "random" filters are linear filters whose impulse responses are realisations of particular random processes.

Along the same lines, the *Random Demodulator* (RD) [10], [14], [15] and the *Modulated Wideband Converter* (MWC) [11], [16]–[18] have recently been proposed as CS sampling systems that target continuous-time spectrally-sparse signals. The RD is a single channel, uniform sub-Nyquist sampling strategy for acquiring *sparse multitone signals*; the MWC is a multi-channel, uniform sub-Nyquist sampling strategy for acquiring *sparse multiband signals*. (Precise definitions for these two signal classes are provided in Section III.) The RD and the MWC have tremendous potential impact because of the longstanding, proven usefulness of spectral signal models in many engineering and scientific

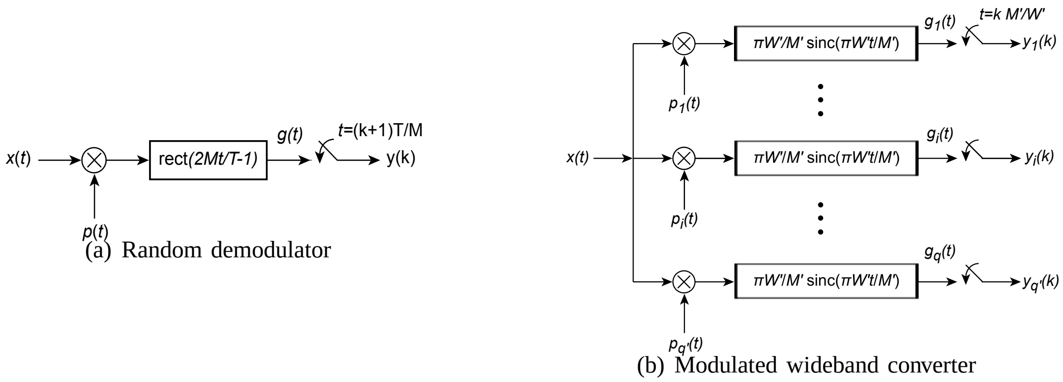
[1]There are several ways to construct viable random sensing matrices. For example, its entries could simply be independent and identically distributed realisations of a a zero mean, unit variance Gaussian random variable.

applications (e.g. electromagnetic scanning, cognitive radio, radar, and medical imaging). Perhaps owing to the near coincidental emergence of these systems, however, few results exist to date that reconcile their remarkably similar structures (see Figure 1) with their different reconstruction algorithms. In fact, the current literature paints a somewhat artificial dividing line between the RD and the MWC, this works preferring to focus primarily on one scheme rather than drawing connections between them.

In this paper, we offer new insights into the relationship between the RD and the MWC that complement the original works of Tropp et al. [10] and Mishali and Eldar [11]. We apply tools from modern sampling theory and classical Fourier analysis and show that the RD and the MWC are two manifestations of the same CS sampling approach, namely random filtering/convolution [9], [13]. This fact reflects the systems' similar structure. At the same time, we show that the sampling functions characterising the systems strongly distinguish the two schemes. For example, the RD's sampling functions have finite support in time and infinite support in frequency, whereas the MWC's sampling functions have infinite support in time and finite support in frequency. In Section III, we discuss two signal model sensitivities exhibited by the RD and the MWC. In particular, we demonstrate the MWC's sensitivity to short duration multiband signals that is the counterpart to the known RD sensitivity to nonharmonic multitone signals. In each case, a perturbation to the signal model triggers a possible loss in signal sparsity and endangers CS reconstruction. In Section IV, we highlight the MWC's use of block convolution as a principle processing step that enables it to successfully sample and recover "block" sparse signals, i.e. signals whose nonzero components are grouped together. The RD does not use block convolution, hence signal reconstruction for it can become computationally expensive for these types of signals. Extending the idea, we lastly propose a new CS based sampling system and show through an example that it can successfully sample and reconstruct continuous-time signals that are block sparse in the time domain.

The main contribution of this paper is the recognition that both systems are based on the underlying concept of random filtering, yet each implements the concept differently because of the different signal classes they target. The insights regarding model sensitivities, the differences in sampling functions, and the MWC's use of block convolution build a better understanding of these systems and allows further application of these ideas to new signal classes, like the new acquisition system for time block-sparse signals proposed in Section IV.

To be clear, we do not discuss the conditions of successful reconstruction, nor implementation issues in this paper. The original works of Tropp et al. [10] and Mishali and Eldar [11], and even some subsequent scholarship [16]–[18], extensively investigate these issues. Some of the reconstruction conditions will be tacitly stated in the descriptions of the systems in Section II, but the presumption throughout the paper is that the RD and the MWC are theoretically proven CS based techniques to



(a) Random demodulator

(b) Modulated wideband converter

Fig. 1. Time domain block diagrams of the random demodulator (RD) and the modulated wideband converter (MWC). The RD is characterised by T, the duration of the observation interval and M, a sampling rate parameter. The MWC is similarly characterised by W, the bandwidth of the input signal $x(t)$ and M', a sampling rate parameter. Note that the primary difference in the sampling structures is the type of filter prior to the sampling operation—the RD uses a ideal integrator and the MWC uses ideal low pass filters.

sample and reconstruct continuous-time spectrally-sparse signals.

II. Sampling mechanisms and signal models

In this section, we examine the sampling mechanisms of the RD and the MWC from a modern sampling theory perspective. We show the output samples for both systems are equal to the inner products of the input signal with a set of sampling functions that arise from the systems' designs. We observe that unlike typical sampling functions, these sampling functions involve random waveforms, a central component in many CS sampling systems. If the inner products are interpreted as analogue filtering operations, we show that the samples result from a generalised random filtering or random convolution as described by Romberg [9] and Tropp et al. [13] as a means to acquire CS samples. This analysis suggests that the RD and the MWC are two manifestations of the same sampling approach, but differ in the specific form of the sampling functions. The difference in sampling functions also reflects the difference in the assumed signal models for the RD and the MWC.

We do not introduce the notion of signal sparsity in this section because the conclusions reached do not depend on this aspect. Signal sparsity and its consequences are discussed in Section III.

A. Sampling with the random demodulator

Let $x(t)$ be a continuous-time, complex-valued signal defined on the real line. The RD acquires samples of $x(t)$ on a finite observation interval where here, we assume, without loss of generality, that the samples are collected in the interval $[0, T]$ seconds. In [10], Tropp et al. adopt a particular signal model for $x(t)$ on this interval. They assume in part that $x(t)$ has a Fourier series (FS) expansion on $[0, T]$ which has bounded harmonics, i.e. $-W \leq \frac{n}{T} < W$ Hz for $n \in \mathbb{Z}$. On this interval, $x(t)$ is

therefore modeled as

$$x(t) = \sum_{n=-N/2}^{N/2-1} X(n) e^{j\frac{2\pi}{T}nt}, \quad t \in [0, T], \tag{1}$$

where $\{X(n)\}$ denotes the FS coefficients of $x(t)$ and $N = TW$. For ease of exposition, N is assumed to be an even positive integer. This signal model is often called a *multitone* model.

To acquire the samples, a RD first multiplies $x(t)$ by a waveform $p(t)$ and then filters and samples the product $x(t)p(t)$ on $[0, T]$ (see Figure 1(a)). The signal $p(t)$ is taken to be a realisation of a continuous random process derived from a vector of Bernoulli random variables. Let $\mathbf{Z} = [Z_0, \ldots, Z_{L-1}]$ be a vector of independent and identically distributed random variables Z_l taking values ± 1 with equal probability and let $p(t; \mathbf{Z})$ denote the random process

$$p(t; \mathbf{Z}) = Z_l, \quad t \in \left[\frac{l}{W}, \frac{l+1}{W}\right), \quad l = 0, \ldots, N-1. \tag{2}$$

A realisation $\mathbf{Z}_0$ of $\mathbf{Z}$ produces a single realisation $p(t; \mathbf{Z}_0)$ of $p(t; \mathbf{Z})$. Here, we abbreviate $p(t; \mathbf{Z}_0)$ by $p(t)$ and thus consider $p(t)$ to be a deterministic quantity, although its randomness plays an important role in proving performance guarantees [10]. In this paper, we sometimes refer to $p(t)$ as a random waveform in deference to this point. We stress that when acquiring samples on $[0, T]$, the RD uses a single realisation of $p(t; \mathbf{Z})$, but different realisations may be used for other observation intervals. Note also that $p(t)$ has the FS representation,

$$p(t) = \sum_{n=-\infty}^{\infty} P(n) e^{j\frac{2\pi}{T}nt}, \quad t \in [0, T] \tag{3}$$

where $\{P(n)\}$ is the set of FS coefficients of $p(t)$.

The analogue filter in the RD design is taken to be an ideal integrator with impulse response $h(t) = \text{rect}\left(\frac{2M}{T}t - 1\right)$, where

$$\text{rect}(x) = \begin{cases} 1 & \text{for } -1 \leq x \leq 1 \\ 0 & \text{otherwise} \end{cases}, \tag{4}$$

and $M \in \mathbb{Z}^+$. The sampling period T_s is taken to be M times shorter than the observation window ($T_s = T/M$). The system therefore samples at the rate of M/T Hz. The multitone signal model and the RD sampling system are therefore parameterised by N, the parameter equal to the time-frequency product TW and M, the parameter that controls the RD's sampling rate. Here, we assume that $M < N$.

The goal of the RD is to sample $x(t)$ at low rates while retaining the ability to reconstruct it in the interval $[0, T]$. Reconstruction entails the discovery of the active frequencies (the signal's spectral support) and the amplitude of the corresponding FS coefficients. If $x(t)$ is spectrally sparse on $[0, T]$, then reconstruction is possible using CS algorithms [10]. In this case, we note that signal reconstruction only implies the recovery of the spectral content of $x(t)$ in the observation interval.

In other words, the samples $y(k), k = 0, \ldots, M-1$, do not convey information about the spectral content of $x(t)$ outside of this interval. To obtain spectral information outside of $[0, T]$, the RD must be applied to other intervals (of possibly different durations). If the RD is applied to consecutive intervals, a time-frequency decomposition of $x(t)$ similar to the short-time Fourier transform can be obtained for multitone signals.

Time domain description. By inspection of Figure 1(a), the output samples $y(k)$ can be expressed as

$$y(k) = g\big((k+1)\tfrac{T}{M}\big) = \int_0^T x(\tau)p(\tau)\mathrm{rect}(\tfrac{2M}{T}(t-\tau)-1)\ d\tau\bigg|_{t=(k+1)\frac{T}{M}},\ k = 0, \ldots, M-1. \quad (5)$$

By substituting (1) into this expression and evaluating the integral, the following equation relating the time domain samples $y(k)$ to the FS coefficients $X(n)$ results:

$$y(k) = \frac{N}{W} \sum_{n=-N/2}^{N/2-1} \sum_{l=k\frac{N}{M}}^{(k+1)\frac{N}{M}-1} p_l \, \frac{e^{j\frac{2\pi}{N}n} - 1}{j2\pi n} \, e^{j\frac{2\pi}{N}nl} \, X(n), \quad k = 0, \ldots, M-1, \quad (6)$$

where $p_l = p(l/W)$. Tropp et al. derived (6) in [10] by analysing an equivalent digital system. In Appendix VI-A, we provide an alternate derivation that explicitly shows the analogue processing inherent in sampling with the RD.

Because sampling is a linear operation with the RD, the samples $y(k)$ can be viewed as inner products of $x(t)$ with the set of sampling functions $\big\{p(\tau)\mathrm{rect}(2k+1-\frac{2M}{T}\tau)\big\}$ where

$$y(k) = g\big((k+1)\tfrac{T}{M}\big) = \big\langle x(\tau), p(\tau)\mathrm{rect}\big(2k+1-\tfrac{2M}{T}\tau\big)\big\rangle, \quad k = 0, \ldots, M-1, \quad (7)$$

and

$$\big\langle s(t), x(t)\big\rangle = \int_0^T s^*(t)x(t)\ dt,$$

for two continuous functions $x(t)$, $s(t)$ on $[0, T]$. These sampling functions have finite duration in time (T/M seconds), but because their Fourier transforms involve sinc functions, they extend infinitely in frequency. In the time-frequency plane, their support partitions the space into vertical strips of equal width (see Figure 2, left panel). We note that unlike modern sampling theory [19], the sampling functions in (7) contain the random waveform $p(t)$, and the conditions they must satisfy to ensure stable recovery is governed by CS theory and not Shannon-Nyquist based sampling theory. (Refer to [19] and [7] for details regarding the conditions that sampling functions typically must satisfy.)

From (5), it is clear the samples $y(k)$ can be thought of pointwise evaluations of the convolution between $x(t)p(t)$ and an ideal integrator. Equally valid, however, is the view that the samples are pointwise evaluations of a random, linear filtering operation involving $x(t)$ and the time-varying analogue filter $h(t,\tau) = p(\tau)\mathrm{rect}(\frac{2M}{T}(t-\tau)-1)$,

$$y(k) = \int_0^T x(\tau)h(t,\tau)\ d\tau\bigg|_{t=(k+1)\frac{T}{M}} \quad k = 0, \ldots, M-1. \quad (8)$$

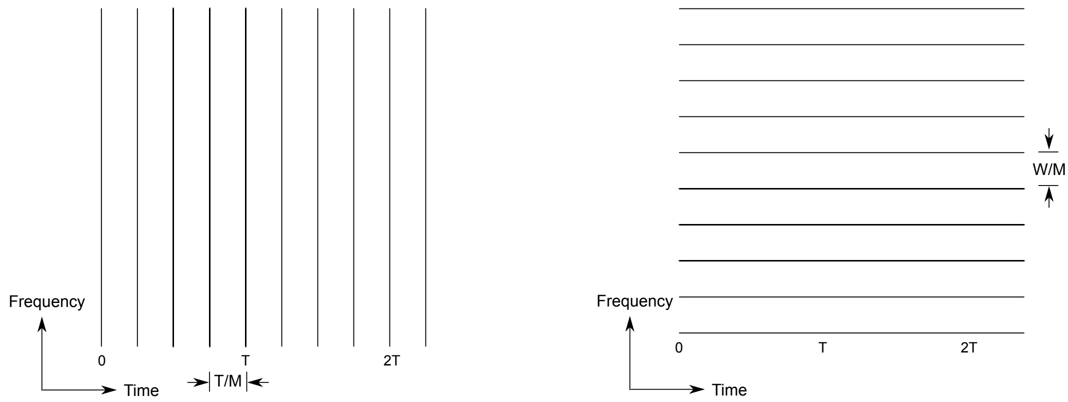


Fig. 2. The output samples of both the RD and the MWC can be described as inner products of the input signal $x(t)$ with certain sets of sampling functions. The panel on the left depicts the time-frequency support of the RD sampling functions where each vertical strip represents the support of one sampling function. Similarly, the panel on the right depicts the support of the MWC sampling functions where each horizontal strip represents the support of one sampling function. For the RD and the MWC, the support characteristics of the sampling functions directly derive from the type of analogue filters used prior to sampling. The RD and the MWC represent two extreme cases: The RD has perfectly localised support in time but completely unlocalized support in frequency. The MWC is the exact opposite.

Here, the impulse response $h(t, \tau)$ is considered random because at each time instance it is a windowed portion of a signal that randomly alternates between ± 1. The samples $y(k)$ can therefore can be thought of as the result of a random filtering operation, conceptually similar to the random filtering schemes proposed in [9] and [13]. In [13], Tropp et al. proposed a CS sampling scheme where a sparse discrete-time signal is first filtered by a digital filter whose impulse response is a realisation of a sequence of independent and identically distributed random variables, and then subsampled at a low rate. They illustrated through examples that with the use of CS recovery algorithms random filtering is a potential sampling structure to acquire CS measurements for sparse discrete time signals. In [9], Romberg proposed and examined a similar idea but considered a specific digital filter that randomly changes the phase of the input signal. Interestingly, Romberg considered the RD as a separate, follow-on processing step to his approach instead of considering it as a generalisation to his notion of random convolution. Here, (8) shows that the sampling mechanism of the RD can be viewed as a random filtering operation applied to continuous-time signals. We note, however, that the filtering operation in (8) is not a convolution because of the time-varying nature of $h(t, \tau)$. Strictly speaking then (8) is distinct from the systems proposed in [13] and [9], although random filtering remains a common thread.

Frequency domain description. An equivalent frequency domain expression to (5) can be derived (see Appendix VI-A) that relates the discrete Fourier transform (DFT) of $y(k)$, denoted by $Y(n)$, to

the FS coefficients $X(n)$,

$$Y(n) = T \sum_{m=n-\frac{N}{2}+1}^{n+\frac{N}{2}} P(m) e^{-j\frac{2\pi}{T}n} \mathrm{sinc}(\tfrac{\pi}{M}n)\, X(n-m), \quad n = 0, \ldots, M-1, \tag{9}$$

where $\mathrm{sinc}(x) = \sin(x)/x, x \in \mathbb{R}$. This equation clearly shows the frequency domain convolution caused by the multiplication with $p(t)$ and the effect of filtering with an ideal integrator, indicated by the presence of the $e^{-j\frac{2\pi}{T}n}\mathrm{sinc}(\frac{\pi}{M}n)$ term. Thus, one can also interpret $Y(n)$ as the output of a random, frequency-varying filter with impulse response $H(n,m) = P(m)e^{-j\frac{2\pi}{T}n}\mathrm{sinc}(\frac{\pi}{M}n)$. We see therefore that the RD's output in either the time or frequency domains can be viewed as the output of a random filter or convolution.

B. Sampling with the modulated wideband converter

We now let $x(t)$ be a bandlimited, continuous-time, finite energy signal. The spectral content of $x(t)$ on $\mathbb{R}$ is thus appropriately given by its Fourier transform (FT) $X(\omega)$,

$$X(\omega) = \int_{-\infty}^{\infty} x(t)\, e^{j\omega t}\, dt.$$

Here, $x(t)$ is bandlimited in the usual sense, i.e., $X(\omega)$ is assumed to be bounded: $X(\omega) = 0$ for $|\omega| \geq \pi W'$ radians per second, $W' \in \mathbb{R}^+$, where $\pi W'$ is the bandwidth of $x(t)$ and $2\pi W'$ is the Nyquist frequency in radians per second. We adopt the following definition from [20]. The class of *multiband* signals $\mathcal{B}(\mathcal{F}, W')$ is then the set of bandlimited, continuous-time, finite energy signals whose spectral support is a finite union of bounded intervals,

$$\mathcal{B}(\mathcal{F}, W') = \left\{ x(t) \in L^2(\mathbb{R}) \cap C(\mathbb{R}) : X(\omega) = 0, \omega \notin \mathcal{F} \right\} \tag{10}$$

where

$$\mathcal{F} = \bigcup_{i=1}^{K} [a_i, b_i), \quad |a_i|, |b_i| \leq \pi W' \text{ radians per second, for all } i. \tag{11}$$

In the following description of the MWC, primes are added to the parameters to distinguish them from the parameters of the RD. The same letters are, however, used for similar quantities. For example, W denotes the bound on the harmonics of multitone signals while $\pi W'$ denotes the bandwidth of the multiband signals.

Like the RD, the ith channel of the MWC multiplies $x(t)$ by a random signal $p_i(t)$, then filters and samples the product $x(t)p_i(t)$ at a sub-Nyquist rate (see Figure 1(b)). As in the original formulation, we assume each channel's filter is an ideal low pass filter, although it has been shown that the MWC can operate with non-ideal low pass filters [17]. Here, we examine its original formulation to make a clearer comparison to the RD. The signals $p_i(t), i = 0, \ldots, q' - 1$, are periodic extensions of different

realisations of a continuous random process similar to that used for the RD. Formally, let $\mathbf{Z} = \{Z_l\}$, be a sequence of independent and identically distributed random variables taking values ± 1 with equal probability and let $p(t;\mathbf{Z})$ denote the random process

$$p(t;\mathbf{Z}) = \begin{cases} Z_l, & t \in \left[\frac{l}{W'}, \frac{l+1}{W'}\right), \quad l = 0, \dots, \lfloor L' \rfloor - 1, \\ Z_{\lfloor L' \rfloor - 1}, & t \in \left[\frac{\lfloor L' \rfloor - 1}{W'}, T_p\right) \end{cases} \tag{12}$$

where $T_p = L'/W'$. The signals $p_i(t)$ are then periodic extensions of the realisations $p(t;\mathbf{Z}_i)$ of $p(t;\mathbf{Z})$:

$$p_i(t + mT_p) = p(t;\mathbf{Z}_i), \quad \text{for } t \in [0, T_p], \ m \in \mathbb{Z}, \ i = 0, \dots, q' - 1, \tag{13}$$

where $\mathbf{Z}_i$ denotes a particular realisation of $\mathbf{Z}$.

The impulse response of the ideal low pass analogue filter is $h(t) = \frac{\pi W'}{M'}\mathrm{sinc}(\frac{\pi W'}{M'}t)$, $M' \in \mathbb{R}^+$, $M' > 1$ and thus has a cut-off frequency of $\frac{2\pi W'}{M'}$ radians per second, where $\frac{1}{T_s} = \frac{W'}{M'}$ is each channel's sampling rate (in Hertz). Every channel thus samples at a rate that is M' times slower than the Nyquist rate. The system's average sampling rate is $q'W'/M'$ Hz. Mishali and Eldar [11] showed that a necessary condition for successful reconstruction is $q' \le M' \le L'$ which implies that $T_s \le T_p$. We assume this condition holds for the MWC throughout the paper.

Time domain description. By inspection of Figure 1(b), we obtain the following time-domain expression for a single channel of the MWC:

$$y_i(k) = g_i\big(k\tfrac{M'}{W'}\big) = \frac{\pi W'}{M'} \int_{-\infty}^{\infty} x(\tau) p_i(\tau) \mathrm{sinc}(\tfrac{\pi W'}{M'}(t - \tau)) \ d\tau \bigg|_{t = k\frac{M'}{W'}}, \quad \text{for all } k \in \mathbb{Z}. \tag{14}$$

This expression corresponds to the analogous time domain expression in (5) for the RD. Like the RD, the samples $y_i(k)$ can be interpreted as inner products with a set of sampling functions $\{\frac{\pi W'}{M'} p_i(\tau) \mathrm{sinc}\big(\pi(k - \frac{W'}{M'}\tau)\big)\}$ where

$$y_i(k) = g_i\big(k\tfrac{M'}{W'}\big) = \big\langle x(\tau), \tfrac{\pi W'}{M'} p_i(\tau) \mathrm{sinc}\big(\pi(k - \tfrac{W'}{M'}\tau)\big)\big\rangle, \quad \text{for all } k \in \mathbb{Z}. \tag{15}$$

In contrast to the RD, however, these sampling functions have finite frequency support and infinite support in time. In the time-frequency plane, their support partitions the space into horizontal strips of width W'/M' Hz (see Figure 2, right panel). This particular set of sampling functions represents one instance of a general theory put forth by Eldar [7] to compressively sample continuous-time signals from unions of shift-invariant spaces, of which multiband signals are members. The theory combines modern sampling theory with CS theory in such a way that samples are acquired in a typical manner by projecting the signal onto a set of sampling functions (as in (15)), but CS theory is needed to reconstruct it. We do not review the details of this theory here because it does not apply to the RD.

Interpreting (14) as a random filtering, we identify the time-vary impulse response as $h(t,\tau) = \frac{\pi W'}{M'} p_i(\tau) \mathrm{sinc}(\frac{\pi W'}{M'}(t - \tau))$. Because the MWC employs an ideal low pass filter, the impulse response

contains a sinc function instead of a rectangular function as seen in (5). Consequently, the impulse response has infinite temporal extent in this (ideal) setting. For the MWC, $h(t,\tau)$ is random in the same general sense as the RD's time-varying impulse response—the sinc function is multiplied by a realisation of a random process.

Frequency domain description. Using standard Fourier analysis techniques, Mishali and Eldar [11] derived the following frequency domain description for one channel of the MWC,

$$Y_i(e^{j\omega\frac{M'}{W'}})\mathbf{1}_{[-\frac{\pi W'}{M'},\frac{\pi W'}{M'})} = \frac{W'}{M'}\sum_{m=-\lfloor\frac{L'}{2M'}(M'+1)\rfloor+1}^{\lfloor\frac{L'}{2M'}(M'+1)\rfloor} P(m)\mathrm{rect}(\tfrac{M'}{\pi W'}\omega)\,X(\omega - m\tfrac{2\pi W'}{L'}), \qquad (16)$$

for $i = 0,\ldots,q'-1$, where $\mathbf{1}_{[\cdot)}$ denotes the indicator function and $\lfloor\cdot\rfloor$ denotes the floor rounding operation. Appendix VI-B contains a slightly different derivation of (16) than that presented in [11]. Comparing (16) to (9), we observe that the DTFT of the output sequences $y_i(k)$ can again be interpreted as the result of a random convolution, where the frequency-varying impulse response is given by $H(m,\omega) = P(m)\mathrm{rect}(\frac{M'}{\pi W'}\omega)$. The spectral content of the samples $y_i(k)$ is expressed by the DTFT, as opposed to the DFT, because $x(t)$ is defined on the entire real line for the MWC instead of on an interval. An implication of this modeling difference is explored in Section III.

Single channel MWC. There are two ways to collapse the MWC into an equivalent single channel system. One can either lengthen the observation interval by a factor of q' (keeping all other parameters fixed), or one can consider increasing the sampling rate while maintaining the same observation interval. If the observation interval is lengthened, the sequence of samples from a single channel MWC can be partitioned into q' groups of $W'T/M'$, where each group of samples is thought of as the output from an individual channel in the multi-channel configuration. Alternatively, one can set the sampling rate of a single channel MWC equal to the average rate of a multi-channel MWC, i.e., set the sampling rate to $q'W'/M'$ Hz, and accordingly adjust the low pass filter's cut-off frequency to $q'W'/M'$ Hz since it acts as an anti-aliasing filter (see Figure 3). In this case, we still maintain the requirement $M' \leq L'$, but assume specifically that $L' = q'M'$ or that T_p is q' times larger than T_s. The frequency domain description of this single channel MWC can now be obtained from (16) by substituting M'/q' for M' and $q'M'$ for L':

$$Y(e^{j\omega\frac{M'}{q'W'}})\mathbf{1}_{[-\frac{\pi q'W'}{M'},\frac{\pi q'W'}{M'})} = \frac{q'W'}{M'}\sum_{m=-\lfloor\frac{q'}{2}(M'+q')\rfloor+1}^{\lfloor\frac{q'}{2}(M'+q')\rfloor} P(m)\mathrm{rect}(\tfrac{M'}{\pi q'W'}\omega)\,X(\omega - m\tfrac{2\pi W'}{q'M'}). \qquad (17)$$

Summary. Equations (5), (9), (14), and (16) all indicate that the sampling mechanisms for the RD and the MWC are based on analogue random filtering/convolution. However, the RD's integrator and the MWC's low pass filter induce significant differences in the specific form of the random convolutions, or equivalently, in their sampling functions. In fact, the different filters induce bipolar time-frequency characterisations that make them well-suited for the signal models they target—an

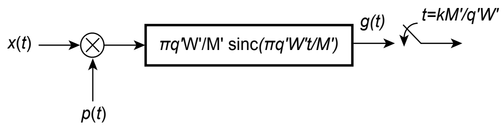


Fig. 3. Block diagram of a single channel MWC as described in [11]. To be equivalent to the multi-channel system depicted in Figure 1(b), this system samples at a rate q' times faster and has a low pass filter with a cut-off frequency q' times greater. Additional digital processing is also required to form the linear system in (19).

ideal integrator with a finite impulse response is well-suited to signals modeled on a finite interval and an ideal low pass filter with a finite frequency response is well-suited to signals modeled on a finite frequency band.

III. Model sensitivities for sparse signal models

In this section, we introduce the notion of sparsity into the multitone and multiband models and discuss two model sensitivities—one for the RD, one for the MWC. These sensitivities are a set of circumstances in which the sparsity of the assumed the models could potentially be lost, jeopardising the ability of any algorithm to correctly recover the unknown signal from a set of compressed measurements. The model sensitivity of the RD is already a familiar limitation that was first mentioned by Tropp et al. in [10] and later studied by others [21], [22]. The MWC model sensitivity presented here is new and ultimately derives from the fact that multiband signals have infinite duration. The analysis also shows that in practice, when only a finite number of samples are acquired, the MWC can at best recover an approximation of the input multiband signal instead of perfectly reconstructing it.

A. Sparse multitone signals for the RD

In the original formulation of the RD, the input signal was not only modeled as a multitone signal on the observation interval, but was also assumed to be spectrally sparse [10]. Recall the multitone signal model on the observation interval $[0, T]$ (equation (1)):

$$x(t) = \sum_{n=-N/2}^{N/2-1} X(n) e^{j\frac{2\pi}{T}nt}, \quad t \in [0, T].$$

A *spectrally sparse multitone* signal is then a multitone signal that has a small number of nonzero FS coefficients out of the $N+1$ possible. More precisely, letting K denote the number of nonzero coefficients (or equivalently the number of nonzero frequencies), a spectrally sparse multitone signal is one that satisfies $K \ll N$.

Signal reconstruction. The reconstruction of a sparse multitone signal $x(t)$ from the samples $y(k)$ hinges on the matrix form of (6),

$$\mathbf{y}(k) = \mathbf{\Sigma}\mathbf{\Psi}\,\mathbf{x}(n) \tag{18}$$

where $\mathbf{\Sigma}$ is a $M \times N$ matrix of the form

$$\mathbf{\Sigma} = \begin{bmatrix} p_0 & \cdots & p_{\frac{N}{M}-1} & 0 & \cdots & & & & \\ 0 & & & p_{\frac{N}{M}} & \cdots & p_{\frac{2N}{M}-1} & & & \\ \vdots & & & & & & \ddots & & \\ & & & & & & & p_{(M-1)\frac{N}{M}} & \cdots & p_{N-1} \end{bmatrix},$$

and

$$\begin{aligned} \mathbf{y}(k) &= [y(0), \ldots, y(M-1)]', \quad \text{(apostrophe denotes transpose)} \\ \mathbf{\Psi}_{r,l} &= e^{-j\frac{2\pi}{N}n_r l}, \\ x_r(n) &= \alpha_r X(n_r), \\ \alpha_r &= \frac{T}{j2\pi n_r}\left(e^{j\frac{2\pi}{N}n_r} - 1\right), \quad \alpha_0 = \frac{1}{W}, \end{aligned}$$

for $r = 0, \ldots, N-1$, $n_r = -N/2 + r$, and $l = 0, \ldots, N-1$. We explicitly write the arguments k and n in (18) to indicate the time domain and the frequency domain, respectively. The notational distinction becomes more important in later sections and is simply made for clarity.

By construction, (18) is an underdetermined linear system of equations ($\mathbf{\Sigma}\mathbf{\Psi}$ is $M \times N$ with $M < N$; see Section II-A) and underdetermined systems do not, in general, have unique solutions. However, CS theory has shown that because of the presumed sparsity of $\mathbf{x}(n)$, (18) can be solved by a direct application of a number of recently developed recovery algorithms, e.g., ℓ_1 minimisation [4], orthogonal matching pursuit [5], or iterative hard thresholding [23], [24]. In the CS literature, solving (18) is termed the *single measurement vector* (SMV) problem. Theoretical guarantees regarding the successful recovery of $\mathbf{x}(n)$ is provided in [10] in terms of the degree of sparsity and the number of samples (measurements) collected.

RD model sensitivity. The ability of CS recovery algorithms to recover the FS coefficients in (18) depends in part on the sparsity of $\mathbf{x}(n)$, or equivalently, on whether $x(t)$ has a sparse FS representation in the observation window $[0, T]$. A multitone signal that has a sparse FS representation on $[0, T]$ (with fundamental frequency $1/T$) will not in general have a sparse expansion if the fundamental frequency changes slightly. In other words, the FS expansion of a nonharmonic tone is generally not spectrally sparse [25, p.379-380]. The implication for the RD is that in a blind sensing scenario, where the frequencies of the tones are not known, there is no guarantee that there will not be a mismatch between the fundamental tone ($1/T$) and the observation interval, which may be slightly larger or smaller than T. Thus, in practice there is no guarantee that the FS representation of $x(t)$

on $[0, T]$ is sparse. Since successful reconstruction is conditioned on the spectral sparsity of $x(t)$ on $[0, T]$, a possible mismatch jeopardises reconstruction. This sensitivity was acknowledged by Tropp et al. in [10], highlighted in [16] and studied in [21], [22]. In particular in [22], Duarte and Baraniuk propose a heuristic solution that marries model-based CS [6], redundant DFT frames, and standard spectral estimation techniques, but it remains an open question whether this method extends to infinite dimensional signal classes such as multiband signals.

B. *Sparse multiband signals for MWC*

Recall that mulitband signals are bandlimited, continuous-time, finite energy signals whose spectral support $\mathcal{F}$ is a union of bounded intervals (see (10) and (11)). A *sparse multiband* signal is a multiband signal whose support has Lebesgue measure that is small relative to the overall signal bandwidth, i.e., $\lambda(\mathcal{F}) \ll W'$ [26]. If, for instance, all the occupied bands (intervals) have equal bandwidth B Hz and the signal is composed of K disjoint frequency bands, then a sparse multiband signal is one satisfying $KB \ll W'$. In the CS literature, signals having this type of "block" structure have been studied in various settings in which the central question is whether the additional signal structure (sparsity plus block structure) reduces the minimum number of samples required to reconstruct the original signal (see e.g., [6], [27], [28]).

Signal reconstruction and linear approximations. MWC signal reconstruction centres on the matrix form of (16),

$$\mathbf{y}(\omega) = \mathbf{\Phi\Psi s}(\omega) \tag{19}$$

where

$$\begin{aligned} y_i(\omega) &= Y_i(e^{j\omega\frac{M'}{W'}})\mathbf{1}_{[-\frac{\pi W'}{M'},\frac{\pi W'}{M'})} \\ \mathbf{\Phi}_{i,l} &= p_{il} \\ \mathbf{\Psi}_{l,r} &= \frac{M'}{W'}e^{-j\frac{2\pi}{L'}lm_r} \\ s_r(\omega) &= \beta_r \text{rect}(\tfrac{M'}{\pi W'}\omega)\, X(\omega - m_r\tfrac{W'}{L'}) \\ \beta_r &= \frac{1-e^{-j\frac{2\pi}{L'}m_r}}{j2\pi m_r}, \quad \beta_0 = 1/L', \end{aligned}$$

for $i = 0, \dots, q'-1$, $l = 0, \dots, L'-1$, $r = 0, \dots, \lfloor \frac{L'}{M'}(M'+1) \rfloor - 1$ and $m_r = -\lfloor \frac{L'}{2M'}(M'+1) \rfloor + 1 + r$. Note that the scalars β_r are the complex conjugates of α_r in (18). Like (18), this linear system of equations is underdetermined since the matrix $\mathbf{\Phi\Psi}$ has dimensions $q' \times \frac{L'}{M'}(M'+1)$ and the assumption $q' \leq L' < M'$ (see Section II-B) implies $q' < \frac{L'}{M'}(M'+1)$. If $x(t)$ is a sparse multiband signal, the vector $\mathbf{s}(\omega)$ is sparse in the sense that most of its elements (segments of $X(\omega)$) do not contain occupied frequency bands that comprise $x(t)$. This fact is important because the CS methods used

in the reconstruction process rely on a signal's sparsity to recover the support of $\mathbf{s}(\omega)$. Equation (19) can also be derived from the single channel MWC, although one has to first extract q' lower rate sample sequences from the single higher rate output sequence. We refer the reader to [11] for details regarding the extra processing steps.

We emphasise that in practice the linear system in (19) cannot, in general, be formulated because it requires an infinite amount of data. To understand this claim and its consequences, we consider the inverse DTFT of (19). From their definition, it immediately follows that the inverse DTFT of the spectra $y_i(\omega) = Y_i(e^{j\omega\frac{M'}{W'}})\mathbf{1}_{[-\frac{\pi W'}{M'},\frac{\pi W'}{M'})}$ (left hand side of (19)) are the time domain sequences $\{y_i(k), k \in \mathbb{Z}\}_i$. To take the inverse DTFT of the right hand side of (19), one interprets the spectral segments $s_r(\omega) = \beta_r\text{rect}(\frac{M'}{\pi W'}\omega)\, X(\omega - m_r\frac{W'}{L'})$ as single periods of periodic spectra. By doing so, it follows that their inverse DTFT are the sequences $\{\gamma_r(k), k \in \mathbb{Z}\}_r$ where

$$\gamma_r(k) = \frac{M'}{2\pi W'}\int_{-\pi W'/M'}^{\pi W'/M'} s_r(\omega)e^{j\frac{M'}{W'}k\omega}\; d\omega. \tag{20}$$

and

$$s_r(\omega) = \sum_{k=-\infty}^{\infty}\gamma_r(k)e^{-j\frac{M'}{W'}k\omega}, \quad \omega \in [-\pi W'/M', \pi W'/M']. \tag{21}$$

The transform pair of (19) is therefore the linear system,

$$\mathbf{Y} = \mathbf{\Phi\Psi S} \tag{22}$$

where $\mathbf{\Phi}$ and $\mathbf{\Psi}$ are as in (19), $\mathbf{Y}$ is an *infinite* column matrix whose ith row is the sequence $y_i(k), k \in \mathbb{Z}$, and $\mathbf{S}$ is an *infinite* column matrix whose rth row equals $\gamma_r(k), k \in \mathbb{Z}$. The matrix $\mathbf{S}$ is described as being *jointly sparse* because most of its rows are zero since the zero-valued elements of $\mathbf{s}(\omega)$ correspond to zero-valued sequences $\gamma_r(k)$ (rows of $\mathbf{S}$). A matrix $\mathbf{Z}$ is said to be K joint sparse if there are at most K rows in $\mathbf{Z}$ that contain nonzero elements. The recovery of $\mathbf{S}$ from the measurements $\mathbf{Y}$ in (22) is called an *infinite measurement vector* (IMV) problem [7], [11], [29] in the CS literature because each column of $\mathbf{Y}$ is viewed as a CS measurement (via the measurement matrix $\mathbf{\Phi\Psi}$) of a collection of vectors that share a common sparse support.

In practice, this IMV problem can never be formulated because we can only ever observe $x(t)$ over a finite duration window, and thus can only ever collect a finite number of samples. This practicality in effect truncates the rows of $\mathbf{Y}$ and causes (22) to become a so-called *multiple measurement vector* (MMV) problem [29]–[32], where the goal is to recover a finite number of the columns of the jointly sparse matrix $\mathbf{S}$ corresponding to the finite number of acquired samples in an underdetermined linear system. Using existing CS methods, this MMV problem can be solved exactly, or with exceedingly high probability, provided the matrix $\mathbf{\Phi\Psi}$ satisfies certain conditions and that enough samples are collected. The solution, however, only provides a linear approximation [25] to the true spectral slices $y_i(\omega) = Y_i(e^{j\omega\frac{M'}{W'}})\mathbf{1}_{[-\frac{\pi W'}{M'},\frac{\pi W'}{M'})}$ because the solution only recovers a finite number of coefficients

$\gamma_r(k)$ in (21). In general, an infinite number of coefficients would need to be recovered to perfectly reconstruct a multiband signal, or equivalently, an infinite number of samples would need to be acquired. This fact is in contrast to the sparse multitone signal model that is parameterised by a finite number of parameters and thus only requires a finite number of samples for perfect signal reconstruction. It is important to note that even though $x(t)$ can only be approximated in practice, there is an advantage in applying CS theory—it allows one to recover the same order of linear approximation obtained by sampling at the Nyquist rate, but at rates significantly below the Nyquist rate.

In [11] and [29], Mishali and Eldar proposed a two step reconstruction process termed the "continuous to finite block" that provably recovers $x(t)$ exactly given an infinite amount of data, or in other words, recovers $x(t)$ to an arbitrary precision given sufficient data. The first step recovers the joint support of $\mathbf{S}$ by solving an associated MMV problem, and the second step uses the found support to reduce the dimension of the measurement matrix $\mathbf{\Phi\Psi}$ such that its pseudoinverse can be used to find a unique solution to the MMV problem. We stress that even if this two step process perfectly solves the MMV problem derived from (22), the solution can only, in general, approximate $x(t)$ for a finite number of samples.

MWC model sensitivity. We now discuss a model sensitivity of the MWC that is similar to the RD's model sensitivity discussed above. Recall that nonharmonic signals on the observation interval $[0, T]$ the signal vector $\mathbf{x}$ in (18) may not be sparse for the RD. Similarly, $\mathbf{S}$ may become non-sparse for short duration multiband signals. To be concrete, let $x(t)$ be a sparse multiband signal with a FT $X(\omega)$ and let $z(t)$ be a windowed version of $x(t)$,

$$z(t) = x(t)w(t), \tag{23}$$

where $w(t)$ is an indicator function of some sub-interval of the observation interval. The signal $z(t)$ is thus a short duration multiband signal that has a spectrum equal to $X(\omega)$ convolved with a sinc function [33]. As is well-known, this convolution spreads the original spectrum $X(\omega)$. If it is sufficiently spread (or equivalently if $w(t)$ is sufficiently short), $\mathbf{s}(\omega)$ and $\mathbf{S}$ are no longer sparse. This situation violates one of the necessary and sufficient conditions for successful reconstruction as outlined by Mishali and Eldar in [11, Theorems 2 and 3, Condition 2]. Thus, as with the RD when sampling nonharmonic multitone signals, the MWC exhibits a limitation in recovering short duration multiband signals because the sparsity of $\mathbf{s}(\omega)$ and $\mathbf{S}$ is potentially lost. The following numerical example illustrates the point.

Example. Consider the set of sparse multiband signals with maximal frequencies of ± 1000 Hz and a MWC with ten channels each sampling at 20 Hz with each channel using a waveform $p_i(t)$ having a period of 0.05 seconds ($W' = 1000$ Hz, $q' = 10$, $L' = M' = 50$). We simulate a multiband

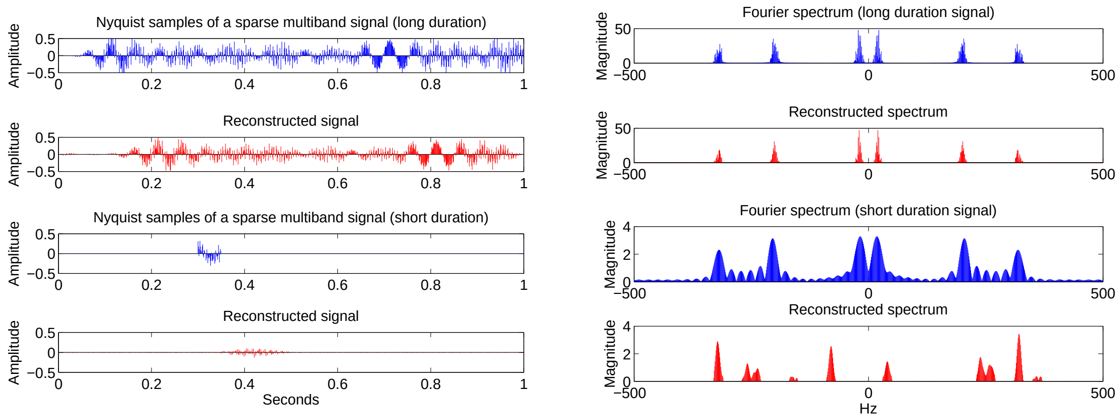


Fig. 4. The plots show the results of an experiment that demonstrates the MWC's sensitivity to short duration signals, or equivalently, short observation intervals. The top row shows a simulated multiband signal bandlimited to 500 Hz with three occupied bands (six between ± 500 Hz). The left hand column shows time domain signals, the right shows the frequency domain. The second row shows the reconstructed signal when the input is sampled by a 10 channel MWC with an average sampling rate of 200 Hz over a one second interval ($W' = 1000$ Hz, $q' = 10$, $M'L' = 50$). In this case, the MWC faithfully reconstructs the input signal, save for the time domain delay caused by the digital filter that simulates the ideal analogue low pass filter. However, if the input signal is shorten to 0.05 seconds (third row) and sampled by the same system, the MWC fails to correctly recover the multiband signal's support and thus fails to reconstruct the signal.

signal $x(t)$ by notch-filtering a discrete time white random noise process that has six active bands each having bandwidth of 20 Hz. In this example, the system average sampling rate is 200 Hz which is 66% greater than the theoretical minimum sampling rate of 120 Hz [34]. The two top panels on the right and left of Figure 4 show the reconstruction results when $x(t)$ persists over the entire 1 second observation interval. In this case, the MWC correctly recovers the support and yields an accurate approximation of $x(t)$. However, if the signal model is shortened to 0.05 seconds while keeping all other parameters fixed, the two bottom panels show that the MWC fails to correctly recover the signal. This short duration signal represents, in some sense, a signal that maximally mismatches the multiband model and the MWC's sampling functions because each have infinite support in time.

IV. Exchanging signal models and sampling continuous-time block-sparse signals

We devote this section to a simple exercise that reveals three interesting aspects of the RD and the MWC and suggests a new analogue CS sampling system. First, the exercise plainly shows that the MWC can successfully sample and recover sparse multitone signals without any modification to the recovery algorithm originally proposed by Mishali and Eldar in [11]. Second, it leads to a special case in which the MWC and the RD produce equivalent SMV problems and that underlines the fact that the RD acquires samples sequentially while the MWC acquires them in parallel. Third, it highlights a property of the MWC that differentiates it from the RD and allows it to successfully sample and recover "block" sparse signals (e.g. multiband signals). We conclude the section by proposing a

new CS sampling system that uses this property to sample and recover continuous-time block-sparse signals.

A. MWC with sparse multitone inputs

Consider the problem of using a MWC to sample and recover a sparse multitone signal instead of sparse multiband signal. Let $x(t)$ be a sparse multitone signal on the observation interval $[0, T]$ with $T = \frac{N}{W}$ and let $p_i(t)$ be as described as in Section II-B. To make the problem meaningful, we assume the observation interval is greater than or equal to the sampling period ($T \geq T_s$) and assume, for ease of exposition, that the period of $p_i(t)$ equals the sampling period ($T_p = T_s$). Equivalently, we assume $N \geq L' = M'$, where we recall from Section II-B that $L' = T_p/W'$ and $M' = T_s/W'$.

A parallel analysis to that given in Appendix VI-B, then yields an expression relating the FS coefficients of $x(t)$ to the DFT coefficients of the output samples,

$$Y_i(n) = \frac{L'}{N} \sum_{m=-\left\lfloor \frac{N}{2L'}(L'+1) \right\rfloor + 1}^{\left\lfloor \frac{N}{2L'}(L'+1) \right\rfloor} P_i(m)\, X(n - \tfrac{N}{L'}m), \tag{24}$$

$$= \frac{L'}{N} \sum_{m=-\left\lfloor \frac{N}{2L'}(L'+1) \right\rfloor + 1}^{\left\lfloor \frac{N}{2L'}(L'+1) \right\rfloor} \sum_{l=0}^{L'-1} p_{il}\, \frac{1 - e^{-j\frac{2\pi}{L'}m}}{j2\pi m}\, e^{-j\frac{2\pi}{L'}lm}\, X(n - \tfrac{N}{L'}m), \tag{25}$$

for $n = -\left\lfloor \frac{N}{2L'} \right\rfloor, \ldots, \left\lfloor \frac{N}{2L'} \right\rfloor - 1$. This expression is analogous to the frequency domain description of the RD given by (9) and is what (16) becomes assuming a sparse multitone signal model and $L' = M'$. In matrix form, (25) becomes the MMV problem,

$$\mathbf{Y} = \mathbf{\Phi\Psi S}, \tag{26}$$

where

$$\mathbf{Y}_{i,v} = Y_i(n_v)$$

$$\mathbf{\Phi}_{i,l} = p_{il}$$

$$\mathbf{\Psi}_{l,r} = \frac{L'}{N} e^{-j\frac{2\pi}{L'}lm_r}$$

$$\mathbf{S}_{r,v} = \alpha_r X(n_v - \tfrac{N}{L'}m_r)$$

$$\alpha_r = \frac{1 - e^{-j\frac{2\pi}{L'}m_r}}{j2\pi m_r}, \quad \alpha_0 = 1/L',$$

for $i = 0, \ldots, q-1$, $l = 0, \ldots, L'-1$, $v = 0, \ldots, \left\lfloor \frac{N}{L'} \right\rfloor - 1$, $n_v = -\left\lfloor \frac{N}{2L'} \right\rfloor + v$, $r = 0, \ldots, \left\lfloor \frac{N}{L'}(L' + 1) \right\rfloor - 1$ and $m_r = -\left\lfloor \frac{N}{2L'}(L'+1) \right\rfloor + 1 + r$. In contrast to sampling a sparse multiband signal, this MWC MMV problem does not result from truncation, rather its finiteness derives from the fact that multitone signals are finitely parameterised. The Fourier components of $x(t)$ can be recovered by solving (26)

using several existing CS algorithms including greedy algorithms [35], mixed norm approaches [36], MUSIC based recovery algorithms [32] and, in particular, the approach proposed by Mishali and Eldar in [11] and [29]. We conclude therefore that the MWC can successfully sample and recover sparse multitone signals without any modification to Mishali's and Eldar's recovery algorithm.

Special case. Note that when the observation interval, the period of $p_i(t)$, and the sampling period are equal, i.e. when $N = L' = M'$, (26) collapses to the SMV problem,

$$\mathbf{y}(n) = \mathbf{\Phi}\mathbf{\Psi}\mathbf{x}(n), \tag{27}$$

where $\mathbf{y}(n) = [Y_0(0), \ldots, Y_{q-1}(0)]'$ and $\mathbf{x}(n) = X(-m_r)$, for $m_r = -\left\lfloor \frac{L'+1}{2} \right\rfloor + 1 + r$, $r = 0, \ldots, L'$. The FS coefficients $\mathbf{x}(n)$ can thus be solved for using standard CS SMV recovery techniques. In this special circumstance, the MWC collects a single measurement vector in the observation interval and thus each channel samples the product $x(t)p_i(t)$ once at the end of the observation interval. Comparing (27) to (18), one sees that in this case the RD collects its samples sequentially while the MWC collects them in parallel. This special case presents an example where the MWC and the RD produce equivalent SMV problems thus provides one way to directly compare the RD and the MWC.

This example highlights a property of the MWC that differentiates it from the RD and allows it to successfully sample and recover multitone signals, as well as multiband signals. Specifically, the convolution in (25) involves shifts of $X(n)$ by integer multiplies of $\frac{N}{L'}$ that, when greater than one, yields a "block convolution". By block-convolution, we mean every DFT coefficient $Y_i(n)$ in (25) is a linear combination of finite segments of $X(n)$. Block-convolution is also seen in (16) where the FT of a multiband signal is shifted by integer multiplies of $\frac{W'}{L'}$. It is, however, in contrast to (9) where the frequency shifts describing the RD are by one.

This aspect of the MWC allows the construction of a linear system of equations like (19) and (26) that describe the original spectra in terms of linear combinations of these blocks. The blocks themselves represent a partition of the frequency axis that effectively discretised the block sparsity of multiband signals. In Section IV-C below, we incorporate this property into a multi-channel random convolution system that samples and approximately recovers *continuous-time block-sparse signals*, the time domain analogue of sparse multiband signals. In the next section, we consider the counterpart of this section and examine the case where a sparse multiband signal is sampled by a RD. Because the RD formulation does not rely on block convolution, we show that the resulting SMV problem can become computationally burdensome in certain circumstances.

B. RD with sparse multiband inputs

When a multiband signal is the assumed signal model for the RD, the system fails to produce a single measurement vector problem whose solution recovers $x(t)$. To be more concrete, let $x(t)$

be a sparse multiband signal bandlimited to W' Hertz and with a fixed spectral occupancy $\lambda(\mathcal{F})$ and consider a RD parameterised by M that samples $x(t)$ on $[0,T]$. Let $p(t)$ be as described in Section II-A. A similar analysis to that contained in Appendix VI-A then leads to the expression

$$y(k) = \sum_{m=0}^{N/M-1} p_{k\frac{N'}{M}+m} \int_{-\pi W'}^{\pi W'} X(\omega)\, \frac{e^{j\frac{\omega}{W'}}-1}{j\omega}\, e^{j\omega\left(k\frac{N'}{M}+m\right)}\, d\omega, \tag{28}$$

where here N' represents the number of Nyquist periods within the observation window ($N' = \frac{T}{W'}$). To construct a RD formulation, one could approximate the integral based on samples of the integrand:

$$y(k) \approx \sum_{m=0}^{N'/M-1} \sum_{i=0}^{D-1} p_{k\frac{N'}{M}+m}\, X(\omega_i) \frac{e^{j\frac{\omega_i}{W'}}-1}{j\omega_i}\, e^{j\omega_i\left(k\frac{N'}{M}+m\right)}\, \delta_\omega, \tag{29}$$

where $\delta_\omega = \frac{2\pi W}{D}$ for some positive integer D and $\omega_i = -\pi W + \delta_\omega(i+1/2)$. One can then see that in comparison to (18), this expression describes a single measurement vector problem $\mathbf{y}(k) = \mathbf{\Sigma\Psi}\,\mathbf{x}(\omega_i)$ where the vector $\mathbf{x}(\omega_i)$ grows infinitely long as the integral is more and more closely approximated, i.e., as $D \to \infty$. Here, the matrices $\mathbf{\Sigma}$ and $\mathbf{\Psi}$ are as in (18). Because the RD is designed to sample and recover sparse multitone signals, this approximation results if one uses a multitone signal as a model for a multiband signal, or in essence, uses a FS to approximate a FT. Clearly, in cases where one wants a fine resolution approximation, the size of the matrices in this SMV problem could grow to become computationally unwieldy.

C. *Sampling continuous-time block-sparse signals*

The class of *continuous-time block signals* $\mathcal{G}(\mathcal{T}, t_0)$ is the set of continuous-time, complex-valued, finite energy signals whose support is a finite union of bounded intervals,

$$\mathcal{G}(\mathcal{T}, t_0) = \left\{x(t) \in L^2([0,t_0]) \cap C([0,t_0]) : x(t) = 0, t \notin \mathcal{T}\right\} \tag{30}$$

where

$$\mathcal{T} = \bigcup_{i=1}^{K} [a_i, b_i), \quad 0 \le a_i, b_i \le t_0 < \infty. \tag{31}$$

A *continuous-time block-sparse signal* is a continuous-time block signal whose support has Lebesgue measure that is small relative to the signal's overall duration, i.e., $\lambda(\mathcal{T}) \ll t_0$.

The proposed system we present in the following paragraphs combines block convolution with the MWC multi-channel architecture and the random convolution ideas of Romberg to obtain a sampling system for continuous-time block-sparse signals (see Figure 5(a)). The resulting system can also be interpreted as the time domain analogue of the MWC.

Let $x(t)$ be a continuous block-sparse signal on the interval $[0,T]$ and let $\{p_i(l\frac{T}{L})\}_i$ be an ensemble of discrete time signals $l = 1, \ldots, L,\ L \in \mathbb{Z}^+$ taking values ± 1 with equal probability. The system

has multiple channels and operates in parallel like the MWC. The ith channel convolves $x(t)$ with $p_i(l\frac{T}{L})$, resulting in the continuous-time signal $g_i(t)$,

$$g_i(t) = \sum_{l=-(L-1)}^{L-1} p_i(l\tfrac{T}{L})x(t - l\tfrac{T}{L}). \tag{32}$$

By construction, this convolution is a block-convolution. The corresponding filter is a standard digital-to-analogue reconstruction filter [33], but is used here in the sampling process. Restricting the time axis to the interval $[0, T/L]$ and with a change of variables, (32) becomes

$$g_i(t)\mathbf{1}_{[0,T/L]} = \sum_{l=0}^{L-1} p_i(l\tfrac{T}{L})x(t - l\tfrac{T}{L})\mathbf{1}_{[0,T/L]}. \tag{33}$$

Representing the segments $x(t - l\frac{T}{L})\mathbf{1}_{[0,T/L]}$ in an appropriate orthogonal basis, (32) may be written as

$$g_i(t)\mathbf{1}_{[0,T/L]} = \sum_{l=0}^{L-1} \sum_{n=-\infty}^{\infty} p_i(l\tfrac{T}{L})\alpha_l(n)\psi_n(t), \tag{34}$$

where $\alpha_l(n) = \left\langle x(t - l\frac{T}{L})\mathbf{1}_{[0,T/L]}, \psi_n(t)\right\rangle$. The signal $g_i(t)$ is then sampled at a rate of $\frac{LM}{T}$ Hz over the interval $[0, \frac{T}{L}]$,

$$y_i(k) = g_i(k\tfrac{T}{LM}) = \sum_{l=0}^{L-1} \sum_{n=-\infty}^{\infty} p_i(l\tfrac{T}{L})\alpha_l(n)\psi_n(k\tfrac{T}{LM}), \quad k = 1, \dots, M, \; i = 1, \dots q. \tag{35}$$

Given the samples $\{y_i(k)\}$ one can solve for L linear approximations of the segments $x(t - l\frac{T}{L})\mathbf{1}_{[0,T/L]}$ by truncating the summation over n such that $|n| \le D < \infty$ and solving the matrix equation

$$\mathbf{Y}\mathbf{\Psi}^{\dagger} = \mathbf{\Phi}\mathbf{X} \tag{36}$$

where

$$\mathbf{Y}_{i,k} = y_i(k)$$

$$\mathbf{\Phi}_{i,l} = p_i(l\tfrac{T}{L})$$

$$\mathbf{\Psi}_{r,k} = \psi_{n_r}(k\tfrac{T}{LM})$$

$$\mathbf{X}_{l,k} = x(k\tfrac{T}{LM} - l\tfrac{T}{L})$$

for $i = 1, \dots, q$, $k = 1, \dots, M$, $r = 0, \dots, 2D$, $n_r = -D + r$ and where $\dagger$ denotes the pseudo-inverse of a matrix. The conditions on whether a unique solution can be found to (36) depends on the properties of the measurement matrix $\mathbf{\Phi}$ and the degree to which $\mathbf{X}$ is sparse. (Because $x(t)$ is assumed to be a continuous-time block sparse signal, it follows that $\mathbf{X}$ is a joint sparse matrix.) Notice that in this problem the CS measurements are not simply the samples $y_i(k)$ as they were in (18) and (22), but are the elements of the product $\mathbf{Y}\mathbf{\Psi}^{\dagger}$. Consequently, the number of CS measurements acquired is *not* related to the sampling rate LM/T, but to the number of channels q and the order

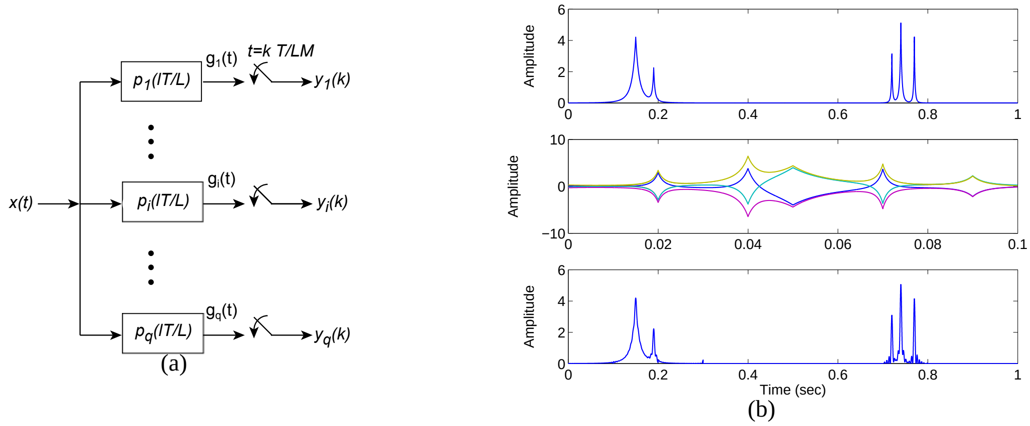


Fig. 5. (a) Schematic diagram of sampling system for continuous-time block sparse signals. The system is a generalization of random convolution as originally proposed in [9]. Each channel convolves $x(t)$ with a random sequence and then samples the result at a low rate. (b) Top panel: Simulated block sparse time domain signal on the unit interval. The signal is a modified version of the "bumps" test signal from the WaveLab toolbox [37]. Middle panel: Overlay plot of the signals $g_i(t)$ resulting from the random convolution. Here, $x(t)$ was sampled with a 6 channel system. Note that samples of $g_i(t)$ are acquired on an interval shorter than the observation interval (0.1 seconds versus 1.0 seconds). Bottom panel: Reconstructed signal approximation. Each 0.1 second duration segment is a linear approximation (Fourier basis) of $x(t)$ on that segment.

of the linear approximation D. The following example, however, shows that a good approximation of a continuous-time block sparse signal can be obtained with a relatively small number of samples.

Example. Consider the continuous-time block sparse signal depicted in the top panel of Figure 5(b). Samples of the filtered signals $g_i(t)$ (middle panel) are acquired by a 6 channel system ($q = 6$) with a sampling rate of 400 Hz ($M = 40, LM/T = 400$)where the time axis is partitioned into 10 segments ($L = 10$). Using a Fourier basis, $\psi_n(t) = e^{j\frac{2\pi}{T/L}}$, the linear system (36) was solved using the simultaneous orthogonal matching pursuit (S-OMP) algorithm [35] resulting in L linear approximations ($D = 41$) of the segments $x(t - l\frac{T}{L})\mathbf{1}_{[0,T/L]}$. The reconstructed signal is shown in the bottom panel of Figure 5(b). In this case, the reconstructed signal is a faithful representation of $x(t)$. Clearly, the quality of the reconstruction depends on how well the linear approximations approximate the segments of $x(t)$ and also on how well a particular CS algorithm solves (36). In any given application, some bases will be more appropriate than others. For example if $x(t)$ contains discontinuities, a wavelet basis would provide a better approximation for a fixed D.

Strictly speaking continuous-time block sparse signals are not bandlimited, but if one examines the spectrum of the test signal in Figure 5(b), one would discover that the signal is "essentially" bandlimited to about 1000 Hz. Thus, it could be argued according to the Shannon-Nyquist sampling theorem that a sampling rate of about 2000 Hz would be required to accurately capture this signal. Relative to 2000 Hz, 400 Hz represents a five fold savings in sampling rate.

V. Conclusion

In this paper, we showed that the sampling mechanisms of the RD and the MWC can both be thought of as being based on the underlying concept of random filtering or random convolution. The most substantial difference between the systems stems from the specific form of their sampling functions (or random filters) and from the assumed signal models. The RD has sampling functions that have finite temporal extent but infinite spectral support; the MWC employs sampling functions that have finite spectral support but infinite temporal support. The randomness in the sampling functions is a hallmark of CS theory that is fundamental in guaranteeing the invertibility of the underdetermined linear systems that characterise the RD and the MWC.

Block convolution is also an important property that differentiates the MWC from the RD because it is one approach that effectively processes infinite dimensional signals that have a block structure. The absence of this property is one primary reason the RD cannot, in general, reconstruct multiband signals. We incorporated block convolution into a new sampling system that samples continuous-time block-sparse signals.

In this paper, we also offered novel insights into the practical reconstruction of sparse multiband signals using the MWC. We showed that with a finite amount of data the MWC necessarily produces linear approximations of the spectral slices of $x(t)$. We also showed that the MWC exhibits a sensitivity to short duration signals that derives from the fact that multiband signals are modeled as extending infinitely in time.

From the perspective of this paper, one begins to consider generalisations to the RD and the MWC that target different signal classes, in particular, CS sampling systems that have different time-frequency characterisations. For example, a system that "compressively" samples radar pulses and chirps in an efficient time-frequency manner could possibly offer a means to effectively detect and classify these signals while avoiding the overhead of sampling several bands simultaneously or reconstructing the Nyquist equivalent signal. This paper takes a step towards this goal by reconciling some of the core ideas behind these sampling systems.

VI. Appendices

The following analyses yield basic time and frequency domain descriptions of the sampling/ measurement strategies. We employ standard Fourier transform properties without explicit explanation for the sake of conciseness. The notational style is that of [33]. To denote transform pairs, we use the shorthand notation,

$$x(t) \overset{\text{FT}}{\longleftrightarrow} X(\omega),$$

and use the abbrevations FT, FS, DTFT, and DFT when referring to the Fourier transform, the Fourier series, the discrete time Fourier transform, and the discrete Fourier transform, respectively. Also recall

the definitions: $\text{sinc}(x) = \sin(x)/x,\ x \in \mathbb{R}$ and

$$\text{rect}(x) = \begin{cases} 1 & \text{for } -1 \le x \le 1 \\ 0 & \text{otherwise} \end{cases} .$$

A. Random Demodulator

Time domain description. Let $x(t)$ be a sparse multitone signal on $[0, T]$ and recall the following transform pairs from Section II:

$$x(t) \xleftrightarrow{\text{FS;1/T}} X(n)$$

$$p(t) \xleftrightarrow{\text{FS;1/T}} P(n)$$

$$h(t) = \text{rect}\big(\tfrac{2M}{T}t - 1\big) \xleftrightarrow{\quad\text{FT}\quad} H(j\omega) = \tfrac{T}{M}\text{sinc}(\tfrac{T}{2M}\omega)e^{-j\omega}.$$

By inspection of Figure 1(a), the time domain description of the RD is

$$\begin{aligned} g(t) = x(t)p(t) * h(t) &= \int_{-\infty}^{\infty} x(\tau)p(\tau)h(t-\tau)\ d\tau \\ &= \int_{t-\frac{T}{M}}^{t} x(\tau)p(\tau)\ d\tau \\ &= \sum_{n=-N/2}^{N/2-1} X(n) \int_{t-\frac{T}{M}}^{t} p(\tau)e^{j\frac{2\pi}{T}n\tau}\ d\tau, \end{aligned}$$

where $*$ denotes convolution. Sampling at $t = (k+1)\frac{T}{M}$ for $k = 0, 1, \ldots$ yields

$$\begin{aligned} y(k) = g\big((k+1)\tfrac{T}{M}\big) &= \sum_{n=-N/2}^{N/2-1} X(n) \int_{k\frac{T}{M}}^{(k+1)\frac{T}{M}} p(\tau)e^{j\frac{2\pi}{T}n\tau}\ d\tau \\ &= \sum_{n=-N/2}^{N/2-1} X(n) \sum_{m=0}^{N/M-1} \int_{k\frac{T}{M}+\frac{m}{W}}^{k\frac{T}{M}+\frac{m+1}{W}} p(\tau)e^{j\frac{2\pi}{T}n\tau}\ d\tau \\ &= \sum_{n=-N/2}^{N/2-1} \sum_{m=0}^{N/M-1} p_{k\frac{N}{M}+m} X(n) \int_{k\frac{T}{M}+\frac{m}{W}}^{k\frac{T}{M}+\frac{m+1}{W}} e^{j\frac{2\pi}{T}n\tau}\ d\tau \\ &= \begin{cases} T\sum_{n=-N/2}^{N/2-1}\sum_{m=0}^{N/M-1} p_{k\frac{N}{M}+m}\, X(n)\, \frac{e^{j\frac{2\pi}{N}n}-1}{j2\pi n}\, e^{j\frac{2\pi}{N}n\left(k\frac{N}{M}+m\right)}, & n \neq 0 \\ \frac{1}{W}\sum_{n=-N/2}^{N/2-1}\sum_{m=0}^{N/M-1} p_{k\frac{N}{M}+m}\, X(n), & n = 0 \end{cases} \end{aligned} \tag{37}$$

where the first three steps follow from the additivity of the integral and the specific nature of $p(t)$. Here, $p_{k\frac{N}{M}+m} = p(k\frac{T}{M} + \frac{m}{W})$. Letting $l = k\frac{N}{M} + m$, (37) may be rewritten as

$$y(k) = \frac{N}{W} \sum_{n=-N/2}^{N/2-1} \sum_{l=k\frac{N}{M}}^{(k+1)\frac{N}{M}-1} p_l\, \frac{e^{j\frac{2\pi}{N}n} - 1}{j2\pi n}\, e^{j\frac{2\pi}{N}nl}\, X(n), \quad \text{for } k = 0, 1, \ldots, \tag{38}$$

where this expression should be understood to be consistent with (37) for $n = 0$ above. For $k = 0, \dots, M-1$, (38) may in turn be written in matrix form,

$$\mathbf{y}(k) = \mathbf{\Sigma}\mathbf{\Psi}\mathbf{x}(n) \tag{39}$$

where $\mathbf{y}(k)$, $\mathbf{\Sigma}$, $\mathbf{\Psi}$, and $\mathbf{x}(n)$ are defined in (18).

Frequency domain description. We also have the following frequency domain description of the RD. **Multiplication/Convolution:**

$$x(t)p(t) \xleftrightarrow{\text{FS;1/T}} \sum_{m=-\lfloor n-\frac{N}{2}\rfloor+1}^{\lfloor n-\frac{N}{2}\rfloor} P(m)X(n-m)$$

Convolution (filtering)/Multiplication:

$$\begin{aligned} g(t) = x(t)p(t) * h(t) \xleftrightarrow{\text{FS;1/T}} G(n) &= \sum_{m=-\lfloor n-\frac{N}{2}\rfloor+1}^{\lfloor n-\frac{N}{2}\rfloor} P(m)X(n-m)H(j\tfrac{2\pi}{T}n) \\ &= \frac{T}{M} \sum_{m=-\lfloor n-\frac{N}{2}\rfloor+1}^{\lfloor n-\frac{N}{2}\rfloor} P(m)X(n-m)e^{-j\frac{2\pi}{T}n}\text{sinc}(\tfrac{\pi}{M}n) \end{aligned}$$

Sampling/Aliasing:

$$y(k) = g\big((k+1)\tfrac{T}{M}\big) \xleftrightarrow{\text{DFT;M}} Y(n) = M \sum_{l=-\infty}^{\infty} G(n-lM)$$

Because $Y(n)$ is M periodic, we can, without loss of information, restrict it to one period. This means we need only consider one term in the summation over l. Retaining the $l = 0$ term yields

$$Y(n) = T \sum_{m=-\infty}^{\infty} P(m)e^{-j\frac{2\pi}{T}n}\text{sinc}(\tfrac{\pi}{M}n)\, X(n-m), \quad n = 0, \dots, M-1.$$

B. Modulated Wideband Converter

We have the following frequency domain description for the i^{th} channel, $i = 0, \dots, q'-1$ of the MWC.

Multiplication/Convolution:

$$x(t)p_i(t) \xleftrightarrow{\text{FT}} \sum_{m=-\infty}^{\infty} P_i(m)X(\omega - m\omega_p) \tag{40}$$

$$= \sum_{m=\lceil(\omega-\pi W')/\omega_p\rceil+1}^{\lfloor(\omega+\pi W')/\omega_p\rfloor} P_i(m)X(\omega - m\omega_p) \tag{41}$$

where $\omega_p = 2\pi W'/L'$ radians per second. The summation limits are finite for a given ω because $x(t)$ is assumed to be bandlimited.

Convolution (filtering)/Multiplication:

$$g_i(t) = x(t)p_i(t) * h(t) \overset{\text{FT}}{\longleftrightarrow} G_i(\omega) = \sum_{m=\lceil(\omega-\pi W')/\omega_p\rceil+1}^{\lfloor(\omega+\pi W')/\omega_p\rfloor} P_i(m)X(\omega - m\omega_p)H(\omega) \tag{42}$$

$$= \sum_{m=-\lfloor\frac{1}{2\omega_p}(\omega_s+2\pi W')\rfloor+1}^{\lfloor\frac{1}{2\omega_p}(\omega_s+2\pi W')\rfloor} P_i(m)X(\omega - m\omega_p)\text{rect}(2\omega/\omega_s) \tag{43}$$

where $H(\omega) = \text{rect}(2\omega/\omega_s)$ is the transfer function of an ideal low-pass filter with cut-off frequency $\omega_s/2$, $\omega_s = 2\pi W'/M'$. Note that the low pass filter windows $X(\omega)$ and its translates (i.e., restricts them to the interval $[-\omega_s/2, \omega_s/2]$) and hence removes the dependence on ω in the summation limits.

Sampling/Aliasing:

$$y_i(k) = g_i(kT_s)$$

$$\updownarrow \quad \text{DTFT}; \omega_s$$

$$Y_i(e^{j\omega\frac{M'}{W'}}) = \frac{W'}{M'} \sum_{n=-\infty}^{\infty} G_i(\omega + n\omega_s) \tag{44}$$

$$= \frac{W'}{M'} \sum_{n=-\infty}^{\infty} \sum_{m=-\lfloor\frac{L'}{2M'}(M'+1)\rfloor+1}^{\lfloor\frac{L'}{2M'}(M'+1)\rfloor} P_i(m)X(\omega - m\omega_p + n\omega_s)\text{rect}(2(\omega + n\omega_s)/\omega_s) \tag{45}$$

where (45) results from substituting (43) into (44) and the summation limits were rewritten using the definition of ω_p and ω_s. Because $Y_i(e^{j\omega\frac{M'}{W'}})$ is periodic with period $\omega_s = 2\pi W'/M'$, we can, without loss of information, restrict $Y_i(e^{j\omega\frac{M'}{W'}})$ to one period. This means we need only consider one term in the summation over n in (45). We choose to retain the $n = 0$ term and thus have the DTFT pair

$$Y_i(e^{j\omega\frac{M'}{W'}})\mathbf{1}_{[-\frac{\pi W'}{M'}, \frac{\pi W'}{M'})} = \frac{W'}{M'} \sum_{m=-\lfloor\frac{L'}{2M'}(M'+1)\rfloor+1}^{\lfloor\frac{L'}{2M'}(M'+1)\rfloor} P_i(m)X(\omega - m\omega_p)\text{rect}(2\omega/\omega_s),$$

where again, $\mathbf{1}_{[\cdot)}$ denotes the indicator function. The Fourier series coefficients of $p_i(t)$ can now be directly computed,

$$P_i(m) = \frac{1}{T_P}\int_0^{T_p} p_i(t)\ e^{-j\frac{2\pi}{T_p}mt}\ dt$$

$$= \frac{1}{T_P}\sum_{l=0}^{L'-1}\int_{l\frac{T_p}{L'}}^{(l+1)\frac{T_p}{L'}} p_{il}\ e^{-j\frac{2\pi}{T_p}mt}\ dt$$

$$= \begin{cases} \sum_{l=0}^{L'-1} \frac{p_{il}}{j2\pi m}\left(1 - e^{-j\frac{2\pi}{L'}m}\right)e^{-j\frac{2\pi}{L'}ml}, & m \neq 0 \\ \frac{1}{L'}\sum_{l=0}^{L'-1} p_{il}, & m = 0, \end{cases}$$

where $p_{il} = p_i(t)$ for $t \in [lT_p/L', (l+1)T_p/L')$, to obtain

$$Y_i(e^{j\omega\frac{M'}{W'}})\mathbf{1}_{[-\frac{\pi W'}{M'},\frac{\pi W'}{M'})} = \frac{W'}{M'} \sum_{m=-\lfloor\frac{L'}{2M'}(M'+1)\rfloor+1}^{\lfloor\frac{L'}{2M'}(M'+1)\rfloor} \sum_{l=0}^{L'-1} p_{il} \frac{1-e^{-j\frac{2\pi}{L'}m}}{j2\pi m} e^{-j\frac{2\pi}{L'}ml} \text{rect}(\tfrac{M'}{\pi W'}\omega) X(\omega - m\tfrac{2\pi W'}{L'}),$$

for $i = 0, \ldots, q'-1$, where this expression should be understood to be consistent with the expression for $P_i(0)$ given above.

We can express these q' linear equations in matrix form

$$\mathbf{y}(\omega) = \mathbf{\Phi\Psi s}(\omega) \tag{46}$$

where $\mathbf{y}(\omega), \mathbf{\Phi}, \mathbf{\Psi}$, and $\mathbf{s}(\omega)$ are as in (19).

REFERENCES

[1] D. Donoho, “Compressed sensing,” *IEEE Trans. Info. Th.*, vol. 52, no. 4, pp. 1289–1306, Apr 2006.

[2] E.J. Candes, “Compressive sampling,” in *Proc. Int. Congr. Math.*, 2006, vol. 3, pp. 1433–1452.

[3] E.J. Candes and M.B. Wakin, “An introduction to compressive sampling,” *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 21–30, Mar 2008.

[4] E. Candes, J. Romberg, and T. Tao, “Robust uncertainty principles: Exact signal reconstruction from highly incomplete frequency information,” *IEEE Trans. Info. Th.*, vol. 52, no. 2, pp. 489–509, Feb 2006.

[5] J. Tropp, “Signal recovery from random measurements via orthogonal matching pursuit,” *IEEE Trans. Info. Th.*, vol. 53, no. 12, pp. 4655–4666, Dec 2007.

[6] R.G. Baraniuk, V. Cevher, M.F. Duarte, and C. Hegde, “Model-based compressive sensing,” *Information Theory, IEEE Transactions on*, vol. 56, no. 4, pp. 1982 –2001, Apr. 2010.

[7] Y.C. Eldar, “Compressed sensing of analog signals in shift-invariant spaces,” *Signal Processing, IEEE Transactions on*, vol. 57, no. 8, pp. 2986 –2997, Aug. 2009.

[8] Y.M. Lu and M.N. Do, “A theory for sampling signals from a union of subspaces,” *Information Theory, IEEE Transactions on*, vol. 56, no. 6, pp. 2334–2345, Jun. 2008.

[9] Justin Romberg, “Compressive sensing by random convolution,” *SIAM J. Imaging Sciences*, vol. 2, no. 4, pp. 1098–1128, 2009.

[10] J.A. Tropp, J.N. Laska, M.F. Duarte, J.K. Romberg, and R.G. Baraniuk, “Beyond Nyquist: Efficient sampling of sparse bandlimited signals,” *IEEE Trans. Info. Th.*, vol. 56, no. 1, pp. 520–544, Jan 2010.

[11] M. Mishali and Y.C. Eldar, “From theory to practice: Sub-nyquist sampling of sparse wideband analog signals,” *Selected Topics in Signal Processing, IEEE Journal of*, vol. 4, no. 2, pp. 375 –391, Apr 2010.

[12] Marco Duarte, Mark Davenport, Dharmpal Takhar, Jason Laska, Ting Sun, Kevin Kelly, and Richard Baraniuk, “Single-pixel imaging via compressive sampling,” *IEEE Signal Processing Mag.*, vol. 25, no. 2, pp. 83–91, Mar 2008.

[13] J.A. Tropp, M.B. Wakin, M.F. Duarte, D. Baron, and R.G. Baranuik, “Random filters for compressive sampling and reconstruction,” *Proc. IEEE International Conf. on Acoustics, Speech, and Signal Processing*, pp. 872–875, 2006.

[14] S. Kirolos, J. Laska, , M. Wakin, M. Duarte, D. Baron, , T. Ragheb, Y. Massoud, and R. Baraniuk, “Analog to information conversion using random demodulation,” *Proc. 2006 IEEE Dallas/CAS Workshop on Design, Applications, Integration and Software*, pp. 71–74, Oct. 2006.

[15] J. Laska, S. Kirolos, M. Duarte, T. Ragheb, R. Baraniuk, and Y. Massoud, “Theory and implementation of an analog-to-information converter using random demodulation,” *Proc. IEEE Int. Symp. Circuits and Systems (ISCAS)*, pp. 1959–1962, May 2007.

[16] M. Mishali and Y.C. Eldar, "Xampling: Analog data compression," in *Data Compression Conference (DCC), 2010*, Mar. 2010, pp. 366 –375.

[17] Yilun Chen, Moshe Mishali, Yonina C. Eldar, and Alfred O. Hero, "Modulated wideband converter with non-ideal lowpass filters," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, Mar. 2010, pp. 3630 –3633.

[18] M. Mishali, A. Elron, and Y.C. Eldar, "Sub-Nyquist processing with the modulated wideband converter," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, Mar. 2010, pp. 3626 –3629.

[19] M. Unser, "Sampling-50 years after Shannon," *Proceedings of the IEEE*, vol. 88, no. 4, pp. 569 –587, Apr. 2000.

[20] Y. Bresler, "Spectrum-blind sampling and compressive sensing for continuous-index signals," *IEEE Info. Th. and Appl. Workshop*, pp. 547–554, Jan 2008.

[21] Yuejie Chi, A. Pezeshki, L. Scharf, and R. Calderbank, "Sensitivity to basis mismatch in compressed sensing," in *Acoustics Speech and Signal Processing (ICASSP), 2010 IEEE International Conference on*, Mar. 2010, pp. 3930 –3933.

[22] M. Duarte and R. Baraniuk, "Spectral compressive sensing," [Online]. Available: dsp.rice.edu/cs, Preprint 2010.

[23] T. Blumensath and M. Davies, "Interative thresholding for sparse approximations," *Jour. of Fourier Anal. and App.*, vol. 14, no. 5, pp. 629–654, 2008.

[24] T. Blumensath and M. Davies, "Interative hard thresholding for compressed sensing," *Appl. Comput. Harmon. Anal.*, vol. 27, no. 3, pp. 265–274, 2009.

[25] Stephane Mallat, *A Wavelet Tour of Signal Processing*, Academic Press, 2 edition, 1999.

[26] Ping Feng, *Universal Minimum-Rate Sampling and Spectrum-Blind Reconstruction for Multiband Signals*, Ph.D., University of Illinois at Urbana-Champaign, Urbana-Champaign, IL U.S.A., 1997.

[27] M. Stojnic, F. Parvaresh, and B. Hassibi, "On the reconstruction of block-sparse signals with an optimal number of measurements," *Signal Processing, IEEE Transactions on*, vol. 57, no. 8, pp. 3075 –3085, Aug. 2009.

[28] Y.C. Eldar, P. Kuppinger, and H. Bolcskei, "Block-sparse signals: Uncertainty relations and efficient recovery," *Signal Processing, IEEE Transactions on*, vol. 58, no. 6, pp. 3042 –3054, Jun 2010.

[29] M. Mishali and Y.C. Eldar, "Reduce and boost: Recovering arbitrary sets of jointly sparse vectors," *IEEE Trans. Signal Processing*, vol. 56, no. 10, pp. 4652–4702, Oct 2008.

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

[31] Jie Chen and Xiaoming Huo, "Theoretical results on sparse representations of multiple-measurement vectors," *Signal Processing, IEEE Transactions on*, vol. 54, no. 12, pp. 4634 –4643, Dec 2006.

[32] M. Davies and Y.C. Eldar, "Rank awareness in joint sparse recovery," [Online]. Available: arXiv:1004:4529v1 [cs.IT], Apr 2010.

[33] Richard Roberts and Clifford Mullis, *Digital Signal Processing*, Addison-Wesley Publishing Co., Inc., 1987.

[34] H.J. Landau, "Necessary density conditions for sampling and interpolation of certain entire functions," *Acta Math.*, vol. 117, no. 1, pp. 37–52, 1967.

[35] J.A. Tropp, A.C. Gilbert, and M.J. Strauss, "Algorithms for simultaneous sparse approximation. Part I: Greedy pursuit," *Signal Processing*, vol. 86, no. 3, pp. 572–588, 2006.

[36] J.A. Tropp, A.C. Gilbert, and M.J. Strauss, "Algorithms for simultaneous sparse approximation. Part II: Convex relaxation," *Signal Processing*, vol. 86, no. 3, pp. 589–602, 2006.

[37] "Wavelab 850: Wavelet analysis toolbox maintained at Stanford University," [Online]. Available: www-stat.stanford.edu/∼wavelab.