

SUB-NYQUIST PROCESSING WITH THE MODULATED WIDEBAND CONVERTER

Moshe Mishali, Asaf Elron and Yonina C. Eldar

Department of Electrical Engineering
Technion — Israel Institute of Technology, Haifa, Israel
{moshiko@tx, elron@tx, yonina@ee}.technion.ac.il

ABSTRACT

Sub-Nyquist systems capture the signal information in a different fashion than uniform high-rate samples. Consequently, digital processing, which is the prime reason for leaving the analog domain, becomes challenging. We propose a digital algorithm that translates samples obtained by the modulated wideband converter, a recent sub-Nyquist system, to the standard format of existing software packages. Our algorithm works in baseband, that is without the need to interpolate the samples to the Nyquist grid. Related methods such as nonuniform sampling or the random demodulator are shown to lack the baseband processing option.

Index Terms— Analog to digital conversion, balanced quadrature correlator, compressed sensing, digital signal processing, modulated wideband converter.

1. INTRODUCTION

Analog to digital conversion (ADC) lies at the heart of modern signal processing, isolating the delicate interaction with the continuous world, so that sophisticated algorithms can be developed in a flexible software environment. In today's technology, analog signals appear around carrier frequencies which are over tens of GHz, for which sampling according to the highest possible frequency exceeds by far the capabilities of commercial ADC devices. When the spectral support is a priori known, carrier demodulation provides the information of interest at a low rate.

In this paper, we focus on sub-Nyquist systems which do not assume the carrier locations [1–3], and more specifically on the ability to perform digital signal processing (DSP) with these systems – the prime reason for shifting to the digital domain. The body of this work investigates the modulated wideband converter (MWC) system [1] in that DSP aspect. After describing the system in Section 2, we define the term baseband processing and elaborate on the digital input type that standard DSP packages are designed to deal with. As we show, the MWC generates outputs that are incompatible with this format. In Section 3, we present a digital algorithm that translates the MWC outputs to the desired standard. In turn, any existing DSP algorithm of interest can smoothly interface with the MWC. Our method consists of three stages: refining the frequency support estimate, isolating the transmissions, and finally a digital carriers recovery stage. The last step utilizes the balanced quadrature correlator, a known and reliable frequency detector which is suitable for many data transmission techniques, both analog and digital [4].

Combining the proposed algorithm with the MWC leads to a powerful sub-Nyquist system, which is capable of sampling and processing, both at a low rate, of signals with unknown spectral support. In Section 4, we demonstrate the proposed algorithm and its accuracy in the presence of noise. The software implementation is avail-

able online at [5]. Section 5 addresses related sub-Nyquist strategies, such as periodic nonuniform sampling [2] and the random demodulator [3], which are both shown to lack the capability of baseband processing at sub-Nyquist rates.

2. THE MODULATED WIDEBAND CONVERTER

2.1. System description

The MWC system treats analog signals $x(t)$, referred to as multi-band, whose Fourier transform $X(f)$ is supported on no more than N frequency intervals (bands), each with width no greater than B . The band locations are unknown.

In order to reduce the sampling rate below Nyquist, the MWC utilizes an analog front-end with m channels, as depicted in Fig. 1. In the i th channel, the input signal $x(t)$ is multiplied by a periodic waveform $p_i(t)$, lowpass filtered, and then sampled every T seconds, resulting in the sequence $y_i[n]$. In the basic configuration, we choose $m \geq 4N$, filter cutoff $1/2T$, rate $1/T = B$, and T -periodic waveforms $p_i(t)$. An example of sign alternating periodic waveforms was studied in [1]. The system samples at rate $m/T \geq 4NB$, where the information rate NB is assumed to be significantly smaller than the Nyquist rate f_{NYQ} of $x(t)$. This basic configuration suffices for our purposes; other configurations with practical advantages are detailed in [1].

In the digital domain, the sample sequences $y_1[n], \dots, y_m[n]$ are processed by a continuous-to-finite (CTF) block. The CTF determines the support of $X(f)$ at resolution B , meaning an index set S , such that $l \in S$ indicates the presence of signal energy in the length B spectrum slice, centered around lB . For example: in the illustration of Fig. 2, $S = \{\pm 1, \pm 2, \pm 3, \pm 4\}$. The support estimate S is then used to (pseudo-)invert a certain matrix \mathbf{C}_S , which depends solely on the choice of periodic waveforms $p_i(t)$ and on the set S [1]. Finally, the sample sequences $y_i[n]$ are translated, via the linear mapping \mathbf{C}_S^\dagger , to another set of lowrate sequences $z_{-L}[n], \dots, z_L[n]$, with L being the smallest integer satisfying $M = 2L + 1 \geq f_{\text{NYQ}}/B$. The rate B complex-valued sequence

$$z_l[n] = \left(x(t) e^{-j2\pi l B t} \right) \star h(t) \Big|_{t=\frac{n}{B}}, \quad -L \leq l \leq L, \quad (1)$$

with \star denoting convolution, contains the information about the l th spectrum slice of $x(t)$; all together, the sequences $z_l[n]$ determine $x(t)$.

2.2. Baseband processing

Shifting from analog to digital is motivated by storage on reliable digital media, and more importantly by the ability to process the data in a flexible software environment. By sub-Nyquist baseband

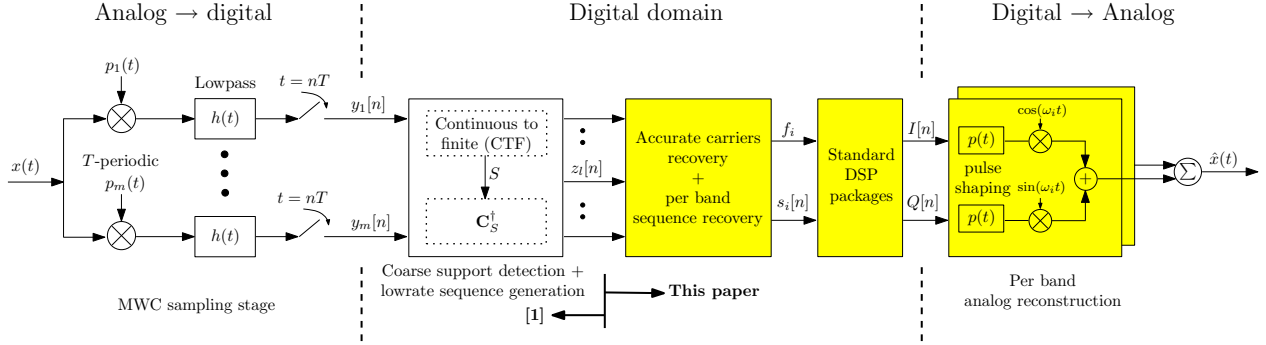


Fig. 1: The modulated wideband converter samples analog signals according to their information rate. The digital processing of [1] generates the lowrate sequences $z_l[n]$, from which point the present contribution starts.

processing we mean the ability to extract the information bits from each band of interest, directly from the samples, namely without performing any kind of interpolation to the Nyquist grid. Baseband processing at the low rate is perhaps the most practical property of sub-Nyquist systems to consider, since the theoretical interest in perfect reconstruction of the original high-rate analog $x(t)$ is often less useful. To achieve this functionality with standard DSP packages, we need

- a sequence $s[n]$, preferably at a low rate, which contains the whole spectral contents of a single band of interest, and
- a fairly accurate estimate of the carrier frequency f_i .

The term carrier frequency usually refers to the conventional quadrature representation of communication signals [4]

$$s(t) = I(t) \cos(2\pi f_c t) + Q(t) \sin(2\pi f_c t), \quad (2)$$

where $I(t), Q(t)$ are real-valued narrowband information signals, and f_c is a relatively high carrier frequency. For example: analog amplitude modulation (AM) obeys the representation (2) with $Q(t) = 0$, whereas frequency modulation has both $I(t), Q(t)$ nonzero [6]. Various digital techniques, such as frequency- or phase-shift keying (FSK/PSK) also conform with (2). When f_c is known prior to sampling, analog demodulation of the carrier followed by filtering is used to provide the DSP module with $I(t), Q(t)$, or their sampled versions according to their actual bandwidths. The software then invokes various synchronization techniques to compensate for slight carrier offsets in the analog demodulation.

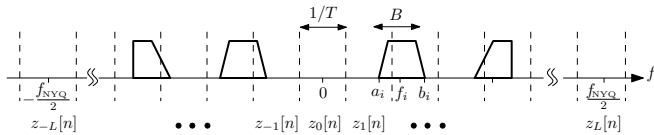


Fig. 2: The Fourier transform of a multiband signal.

The MWC does not utilize carrier knowledge prior to sampling and consequently the sequences $z_l[n]$ do not relate directly to the information signals $I(t), Q(t)$ of a band of interest. For example: in Fig. 2, the energy of the i th band split between the first and the second spectrum slices, and in general $z_l[n]$ may contain simultaneous contributions of several bands. One possible solution for this shortcoming is to reconstruct the Nyquist rate samples $x(n/f_{\text{NYQ}})$ from $z_l[n], l \in S$ and then manipulate the high rate sequence. However, since f_{NYQ} is prohibitively large this approach may be computationally infeasible, which motivates the baseband processing definition.

Our goal is to use only lowrate computations (i.e., proportional to NB rather than to f_{NYQ}) and to obtain the band edges $[a_i, b_i]$, the carrier f_i and a lowrate sequence $s_i[n]$, per band.

In the next section, we propose an algorithm that provides the MWC with the baseband processing functionality, with computation complexity proportional to $NB \ll f_{\text{NYQ}}$.

3. ALGORITHM

We propose a three-steps algorithm, which begins in refining the support estimate to the actual band edges $[a_i, b_i]$. We rely on two additional model parameters: the minimal width of a single band b_{\min} and the smallest spacing between bands Δ_{\min} . These quantities are often implied by the application specification. The second step processes $z_l[n]$ and incorporates the edges $[a_i, b_i]$ to provide a lowrate sequence $s_i[n]$ per band $1 \leq i \leq N/2$. Finally, an accurate carrier detection technique finds f_i assuming $s_i[n]$ corresponds to a signal of the form (2). In the sequel, we mention the relevant MATLAB commands (in verbatim style) that are used in our implementation.

Step 1. For convenience, we start with converting the complex-valued $z_l[n]$ to real-valued counterparts. Recall that the input $x(t)$ is real-valued with conjugate symmetric $X(f)$. Therefore, $l \in S$ implies $-l \in S$ and $z_{-l}[n] = z_l^*[n]$. In step (1.1) of Fig. 3, a real-valued sequence $x_l[n]$ at rate $2B$ for each $l \in S, l > 0$ is obtained by re-positioning $z_l[n], z_{-l}[n]$ from both sides of the origin. Mathematically, $x_l[n] = I_{2,0.5B}\{z_{\pm l}[n]\}$, where the operator

$$I_{r,F}\{z_{\pm l}[n]\} \triangleq (z_l[n] \uparrow r) e^{-j2\pi F n} + (z_{-l}[n] \uparrow r) e^{j2\pi F n}, \quad (3)$$

and $\uparrow r$ denotes rate conversion by a factor of r , with the appropriate post-filtering. By abuse of notation, here and in the sequel the same index n is used before and after the rate conversion, where the context resolves the ambiguity. The case $l = 0 \in S$, has $x_0[n] = z_0[n]$. We used `interpft` to carry out the interpolations in (3). The information rate is not changed; $z_l[n]$ is complex-valued at rate B , while $x_l[n]$ is real-valued at rate $2B$.

Power spectral density (PSD) estimation of $x_l[n]$ is invoked in (1.2) in order to locate the energy concentration within each spectrum slice. We used the Welch PSD estimation method, implemented by `pwelch`, which divides the input to overlapping sections with overlap ratio 50%, filters each section by a Hamming window, performs a discrete Fourier transform (DFT) on each section, and finally averages the results. The frequency resolution and the window size are determined by:

$$f_{\text{res}} = \min(b_{\min}, \Delta_{\min}), \quad W \geq \frac{2B}{f_{\text{res}}}. \quad (4)$$

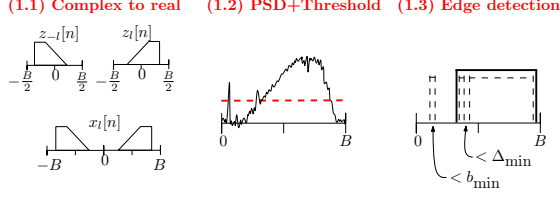


Fig. 3: Fine support (band edges) detection (Step 1).

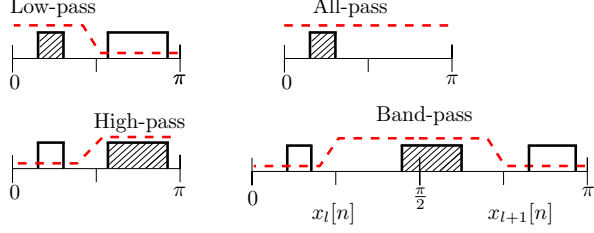


Fig. 4: Band isolation (Step 2). Merging occurred in the right-bottom drawing.

The PSD estimation produces $P_{xx}^{(l)}[k]$ for $1 \leq k \leq K \approx W/2$, where the accuracy of the estimation increases with the number of samples in $x_l[n]$. The window size introduces an inherent trade-off, where a short one gives better averaging of additive noise, while a longer one allows higher DFT orders and thus improves the number of frequency bins K . In (4), the shortest possible window is used. A logarithmic threshold

$$\log_{10}(\text{Threshold}) = \frac{1}{K} \sum_{k=1}^K \log_{10} P_{xx}^{(l)}[k], \quad (5)$$

translates $P_{xx}^{(l)}[k]$ to a binary decision on the energy concentration.

Finally, in step (1.3) we mitigate undesired noise effects that were encountered in simulations; Support regions that are closer than Δ_{\min} are united, and isolated regions with width smaller than b_{\min} are pruned. The operations (1.1)-(1.3) are carried out for each $l \in S, l \geq 0$. To conclude this step, only the $N/2$ most powerful bands, according to the PSD value, are retained to further mitigate noise effects. The output of Step 1 consists of $N/2$ pairs $[a_i, b_i]$ roughly indicating the start and the stop edges of the information bands. To this end, the pairs are ordered such that $a_i < b_i < a_{i+1}$, and by convention $a_0 = b_0 = 0, a_{(N/2)+1} = b_{(N/2)+1} = f_{\text{NYQ}}/2$.

Step 2. The purpose of this step is to isolate a sequence $s_i[n]$ for each $1 \leq i \leq N/2$, such that $s_i[n]$ contains the entire contribution of exactly one band of information. Using the edges $[a_i, b_i]$ we identify cases in which the information resides in adjacent spectrum slices $x_l[n], x_{l+1}[n]$ for some $0 \leq l \in S$; see Fig. 2 for an illustration. In such cases, merging occurs via

$$\tilde{s}_i[n] = I_{4,0.5B}\{z_{\pm l}[n]\} + I_{4,B}\{z_{\pm(l+1)}[n]\}, \quad (6)$$

whereas $\tilde{s}_i[n] = x_l[n]$ when both a_i, b_i lie within the same spectrum slice. By now, $\tilde{s}_i[n]$ contains the entire energy of the i th band with possible contributions from other bands. The information $[a_i, b_i]$ from Step 1 is utilized again to decide on the next actions.

Consider the i th band, and for brevity assumes no merging step was required, such that $[a_i, b_i] \subseteq [lB - B/2, lB + B/2]$ for some $0 \leq l \in S$. Let $[\omega_{p,L}, \omega_{p,H}]$ be the normalized angular frequencies of $x_l[n]$ corresponding to $[a_i, b_i]$, and set $\omega_{s,L} = 0, \omega_{s,H} = \pi/2$. If either b_{i-1}, a_{i+1} resides within the same spectrum slices $z_l[n]$,

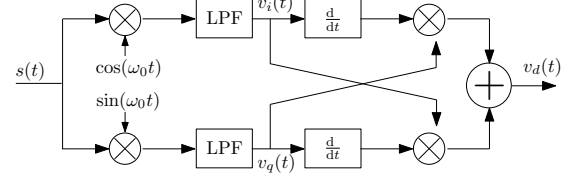


Fig. 5: The analog balanced-quadricorrelator.

update the normalized angular frequencies $\omega_{s,L}, \omega_{s,H}$, respectively. Next, design a filter with passband $\omega_{p,L} \leq \omega \leq \omega_{p,H}$ and stopping bands $[0, \omega_{s,L}]$ and $[\omega_{s,H}, \pi/2]$. As Fig. 4 shows, the resulting filter may be low-, high-, band- or all-pass, depending on the specific values of $\omega_{s,L}, \omega_{s,H}$. We used `firpmord` and `firpm` to determine the filter order and for the actual realization. We allowed passband ripple of 10^{-6} and stopband ripple of 10^{-2} . The filter order is often small, since the actual spacing between the bands relaxes the cutoff constraints. At last, $s_i[n]$ is obtained by filtering $\tilde{s}_i[n]$ with the designed finite impulse response (FIR). Similar filter design holds with merging occurred, eq. (6).

At this point, we have a sequence $s_i[n]$ for each band $1 \leq i \leq N/2$ at a uniform rate of either $2B$ or $4B$, depending on whether the merging (6) was required. In either case, the middle frequency $\tilde{f}_i = (a_i + b_i)/2$ can serve as a rough estimate of the unknown carrier of (2). In fact, in simulations we observed that \tilde{f}_i is not far from the true carrier, as long as the PSD is sufficiently accurate, a situation which occurs for a high signal to noise ratio (SNR) and many samples from $x_l[n]$. The next step incorporates an accurate carrier frequency detector which relies on (2) to better predict the carriers even in cases that the PSD curve is noisy and inaccurate.

Step 3. We start with describing the balanced quadricorrelator (BQ), that was analyzed in [4] and whose circuit appears in Fig. 5. The BQ receives an input $s(t)$ of the form (2), assumes a certain carrier frequency $f_0 = \omega_0/2\pi$, and outputs $v_d(t)$ whose expected value is proportional to the carrier error

$$\mathbb{E}[v_d(t)] = -K_G(f_c - f_0)(\mathbb{E}[I^2(t)] + \mathbb{E}[Q^2(t)]). \quad (7)$$

In practice, time averaging replaces the expectation. The signals $I(t), Q(t)$ that build $s(t)$ are assumed random with zero-cross correlation, $\mathbb{E}[I(t_1)Q(t_2)] = 0$ for all t_1, t_2 . The constant K_G in (7) captures the analog gains along the way: the mixers, the filters, and the differentiators. Note that zero-cross correlation holds for AM, and also for FSK/PSK with a preceding source coding stage [4].

In the proposed algorithm, we implement a digital version of the BQ. A fundamental requirement for the BQ operation, either in analog or digital, is that the first mixing yields non-overlapping copies of $s(t)$. To ensure this property, each $s_i[n]$ is interpolated by a factor of three, and the positive and negative frequencies are re-positioned in $[\pi/3, 2\pi/3], [-2\pi/3, -\pi/3]$, respectively. For example, when no merging occurs this operation boils down to $I_{6,1.5B}\{z_{\pm l}[n]\}$ with the relevant l . The digital BQ is applied on the outcome.

Our digital implementation consists of FIR lowpass arm filters, and the continuous derivatives are approximated by the finite difference – a filter with the discrete impulse response $[1, -1]$. The BQ is initialized with a normalized angular frequency $\omega_0 = \pi/2$ and is repeatedly applied. At each iteration, the time-averaged $v_d[n]$ is multiplied by a loop gain $G = 5 \cdot 10^6$, and ω_0 accumulates $G v_d[n] / \|s_i[n]\|^2$, while monitoring $\omega_0 \in [\pi/3, 2\pi/3]$. The procedure terminates upon convergence or if a pre-defined number of iteration is reached. Note that a wide family of filters can substitute the true differentiators [4].

Properties. Upon completion, the carrier f_i is computed from the normalized angular frequency ω_0 that the BQ converged to by

$$f_i = B \left(c \frac{\omega_0 - \pi/2}{\pi/3} + l + \frac{c-1}{2} \right), \quad (8)$$

where $c = 1$ when merging was not required, and $c = 2$ otherwise. In addition, we supply $s_i[n]$ for the DSP module of Fig. 1. For each band, the digital BQ operates at rate either $6B$ or $12B$, depending on the rate of $s_i[n]$. The recovered carrier f_i and the detected band edges $[a_i, b_i]$ allow to reduce the rate of $s_i[n]$ to the minimal rate $2(b_i - a_i)$. The DSP can extract the information $I[n]$ by no more than multiplying by $\cos[2\pi f_i n]$ and lowpass filtering; $Q[n]$ is obtained similarly with $\sin[2\pi f_i n]$.

For applications in which b_{\min}, Δ_{\max} are unknown, Steps (1.1)-(1.2) may yield many possible support regions. However, since only the $N/2$ powerful regions are selected, and since the BQ iterates on the carriers f_i , the missing b_{\min}, Δ_{\max} has little effect in practice.

A final comment regarding the conversation to analog in Fig. 1. In [1], $x(t)$ is reconstructed directly from $z_l[n]$ by interpolation to $z_l(t)$ and properly positioning of the spectrum slices. Since the scenario of band splitting is common, it can be verified that this procedure requires $2N$ mixers and filters at the most. A nice feature of the present approach is that only N mixers and filters are required. Furthermore, the reconstruction reduces to standard modulation of the narrowband information signals $I[n], Q[n]$.

4. SIMULATIONS

To evaluate the proposed algorithm, we considered an example of a multiband model $N = 3, B = 50$ MHz. Quadrature phase-shift keying (QPSK) modulation was used to generate $x(t) = \sum_{i=1}^3 x_i(t)$ via

$$x_i(t) = \sqrt{\frac{2E_i}{T_{\text{sym}}}} \left(\sum_n I_i[n] p(t - nT_{\text{sym}}) \right) \cos(2\pi f_i t) + \left(\sum_n Q_i[n] p(t - nT_{\text{sym}}) \right) \sin(2\pi f_i t) + n(t), \quad (9)$$

where $E_i = \{1, 2, 3\}$, $1/T_{\text{sym}} = 30$ MHz, $p(t) = \text{rcos}(t/T_{\text{sym}})$ are the symbol energy, rate and raised-cosine pulse shape with 30% rolloff. The carriers $f_i \in [0, 5]$ GHz, the bit streams $I_i[n] = \pm 1, Q_i[n] = \pm 1$, and the additive white Gaussian noise $n(t)$ were all drawn independently at random.

An MWC with the basic configuration was used with $m = 30$ channels and sign alternating waveforms $p_i(t)$, $M = 195$ alternations points per period T . To verify the performance of the proposed algorithm, we assume the spectrum slices $z_l[n]$ were obtained successfully by the preceding stages, namely the CTF and matrix inversion steps of Fig. 1. For each one of 40 test signals, we measure the carrier frequency offset (CFO) of each f_i . Fig. 6 reports the distribution of the CFOs encountered in our simulations. Evidently, in most cases our algorithm approaches the true carriers as close as 150 kHz. For comparison, the 40ppm CFO specifications of IEEE 802.11 standards tolerates this error for transmissions located around 3.75 GHz. The offset is to be further reduced in the DSP, based on QPSK-specific synchronization techniques.

5. RELATED WORK

Two leading alternative sub-Nyquist approaches appear in the literature. We considered periodic nonuniform sampling in [2], where

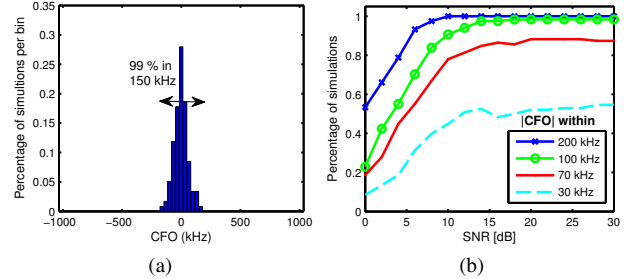


Fig. 6: The distribution of CFO vs. SNR (a). In (b), SNR=10 dB and the curves represent the percentage of simulations in which the CFO magnitude is within the specified range.

the CTF and the matrix inversion were introduced. This approach however requires Nyquist-rate ADCs, and in addition necessitates interpolating the lowrate nonuniform sequences to the high Nyquist grid before the CTF or any other processing can occur [1].

The random demodulator (RD) of [3] is a special case of the MWC with $m = 1$, a pseudo-random (aperiodic) sign alternating $p_1(t)$, and $h(t)$ which is restricted to an ideal integration. The RD is suited to signals consisting of a finite set of pure sinusoids, rather than analog multiband signals, such as QPSK transmissions. In the examples considered earlier, $N = 6, B = 50$ MHz, $f_{\text{NYQ}} = 10$ GHz, the recovery problem boils down to a sparse inversion problem with a matrix dimension $\approx 10^9 \times 10^{10}$ [1]. For comparison, in the same setting, the CTF of the MWC involves a sparse inversion of a matrix size of 35×195 . More inherently, the reconstruction from the RD samples aims at recovering the Nyquist rate signal directly, with no intermediate lowrate sequences such as $z_l[n]$ of the MWC. Consequently, the prominent advantage of sub-Nyquist baseband processing is not accomplished at present by neither the RD nor nonuniform sampling.

6. REFERENCES

- [1] M. Mishali and Y. C. Eldar, "From theory to practice: Sub-Nyquist sampling of sparse wideband analog signals," *arXiv.org 0902.4291*; to appear *IEEE J. Sel. Topics Signal Process.*
- [2] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Trans. Signal Process.*, vol. 57, no. 3, pp. 993–1009, Mar. 2009.
- [3] J. A. Tropp, J. N. Laska, M. F. Duarte, J. K. Romberg, and R.G. Baraniuk, "Beyond nyquist: Efficient sampling of sparse bandlimited signals," *arXiv.org 0902.0026*; to appear *IEEE Trans. Inf. Theory*.
- [4] F. Gardner, "Properties of frequency difference detectors," *IEEE Trans. Commun.*, vol. 33, no. 2, pp. 131–138, Feb. 1985.
- [5] A. Elron, M. Mishali, and Y. C. Eldar, "Carrier frequency recovery from sub-Nyquist samples: Online documentation and simulations," Available: <http://webee.technion.ac.il/Sites/People/YoninaEldar/Info/software/FR/FR.htm>.
- [6] N. Boutin and H. Kallel, "An arctangent type wideband PM/FM demodulator with improved performances," in *Circuits and Systems, 1990., Proceedings of the 33rd Midwest Symposium on*, 1990, pp. 460–463.