Real-time Multi-Gigahertz Sub-Nyquist Spectrum Sensing System for mmWave

Zihang Song, Haoran Qi and Yue Gao

z.song.h.qi,yue.gao@qmul.ac.uk
Queen Mary University of London
London

ABSTRACT
A real-time sub-Nyquist wideband spectrum sensing system for millimeter wave (mmWave) implemented on National Instruments mmWave software-defined radio system is presented. Based on compressed sensing theory and multicoset sampling architecture, the system is capable of achieving real-time spectrum sensing of 3.072 GHz-bandwidth signal at the centre frequency of 28.5 GHz. Bayesian sparsity estimation and data decimation are applied to realize robust performance of spectrum reconstruction under dynamic spectrum scenarios and enable real-time processing, respectively. This paper presents and comments on the impact of noise corruption, spectrum sparsity on the recovery performance and evaluates two low-complexity sparse recovery greedy algorithms of interest.

CCS CONCEPTS
• Hardware → Digital signal processing.

KEYWORDS
sub-Nyquist sampling; wideband spectrum sensing; millimetre wave

ACM Reference Format:

1 INTRODUCTION
The cost, complexity and energy consumption of Nyquist-rate analog-to-digital converters (ADCs) are the practical challenges for real-time wide-band spectrum sensing implementations in millimetre wave (mmWave) frequencies [11]. Compressed sensing (CS) has been proposed in recent years [8] aiming at alleviating the pressure on ADCs in wide-band spectrum sensing (WSS). CS utilizes the sparsity structure of wideband signals in the frequency domain reconstructs the spectrum from compressed sub-Nyquist measurements[2]. Specially designed sub-Nyquist baseband sampling schemes have been designed to construct the compressed measurements, among which the multicoset sampler has advantages of relatively simple implementation and applies multiple measurement vectors (MMV) model to achieve signal reconstruction with lower computational cost [4][6].

We hereby demonstrate a real-time sub-Nyquist mmWave spectrum sensing system based on National Instruments (NI) mmWave software-defined radio (SDR) system [1]. Multiband signals with 3.072 GHz bandwidth downconverted from 28.5 GHz are sampled by multicoset sampler with each ADC sampling at a few tens of megahertz, then the sub-Nyquist measurements are processed in parallel and the real-time spectrum is recovered based on an MMV model via two particular greedy algorithms respectively. Bayesian sparsity estimation is used to estimate the sparsity as the input for these greedy algorithms. Data decimation is applied for reducing computation cost and enabling real-time processing.

The following parts of the paper is organized as follows. The mathematical model of multicoset sampler is shown in Section 2 in detail. In Section 3, analyses are made to showcase the influence of multicoset sampler’s design parameters on the system performance. The system implementations on the NI SDR system are demonstrated in Section 4.

2 MATHEMATICAL MODEL
2.1 Multicoset sampling
A band-limited complex signal \( x(t) \) with frequency range \( [-B_2, B_2] \) is equally divided into C channels, where C is a even integer. The input signal is sampled by a multicoset sampler with \( P \) cosets (\( P < C \)). In each coset the signal is sampled at Landau rate \( B/C \) and a unique delay \( c_p/B \), where \( c_p \) is a positive integer satisfying

\[
0 \leq c_p < C, \quad p = 1, 2, ..., P.
\]
Thus the sampled data in channel $p$ can be expressed as

$$x_p[n] = x[nC/B + p/B], \quad n = 0, 1, 2, \ldots$$  \hspace{1cm} (1)

The discrete-time Fourier transform (DTFT) of the digital signal in channel $p$ can be regarded as the Fourier transform of the product of the original continuous input with offset $c_p/B$ and a periodic impulse sequence. It is worth noting that $x(t)$ is a band-limited signal with bandwidth $[-\frac{B}{2}, \frac{B}{2}]$, for $f \in [0, \frac{B}{2})$, the result is reduce to $C$ terms as

$$X_p(\exp(j2\pi fC/B)) = \mathcal{F}\{x(t + c_p/B) \cdot \sum_{n=-\infty}^{+\infty} \delta(t - nC/B)\}$$

$$= \frac{B}{C} \exp(j2\pi c_p f/B) \sum_{n=-[\frac{C}{2}]}^{+[\frac{C}{2}]-1} X(f - nB/C) \exp(-j2\pi ncp/C)$$  \hspace{1cm} (2)

Denote $Y_p(f) = \frac{C}{B} \exp(-j2\pi ncp/C)X_p(\exp(j2\pi fC/B))$, then eq.(2) can be rewritten as

$$Y_p(f) = \sum_{n=-[\frac{C}{2}]}^{+[\frac{C}{2}]-1} X(f + nB/C) \exp(j2\pi ncp/C).$$  \hspace{1cm} (3)

Take even $C$ as an example, eq. (3) in all $P$ cosets can be connected as a matrix manipulation

$$
\begin{bmatrix}
Y_c(f)
Y_p(f)
\vdots
Y_p(f)
\end{bmatrix} = A \cdot
\begin{bmatrix}
X[f + \frac{B}{2}(-\frac{C}{2})] \\
X[f + \frac{B}{2}(-\frac{C}{2} + 1)] \\
\vdots \\
X[f + \frac{B}{2}(\frac{C}{2} - 1)]
\end{bmatrix}
$$

$$y(f) = A \cdot x(f),$$  \hspace{1cm} (4)

where $A$ is a $P \times C$ matrix with $A_{p,q} = \exp[j2\pi \frac{pc}{C}(-\frac{C}{2} + q - 1)]$ for $1 \leq p \leq P$ and $1 \leq q \leq C$.

Considering the dynamic frequency-domain characteristics of actual signal, only finite-length time-domain data per frame can be obtained for spectrum reconstruction. Thus, the discrete Fourier transform (DFT) is applied in practice. Comparing the definition equations of DTFT and DFT, we can obtain $f = \frac{B}{2C}$. Then (4) can be rewritten as

$$
\begin{bmatrix}
Y_c[k]
Y_p[k]
\vdots
Y_p[k]
\end{bmatrix} = A \cdot
\begin{bmatrix}
\hat{X}[kB/C + \frac{B}{2}(\frac{C}{2})] \\
\hat{X}[kB/C + \frac{B}{2}(\frac{C}{2} + 1)] \\
\vdots \\
\hat{X}[kB/C + \frac{B}{2}(\frac{C}{2} - 1)]
\end{bmatrix}
$$

$$y[k] = A \cdot \hat{x}[k].$$  \hspace{1cm} (5)

Equation (5) is a single measurement vector (SMV) model, where $N$ is the sample number acquired per frame and $k = 0, 1, 2, \cdots, N - 1$ is the sample index. If we stack all the $N \hat{y}[k]$ and $\hat{x}[k]$ together by column respectively, a multiple measurement vector (MMV) model

$$\hat{Y} = A \hat{X}$$  \hspace{1cm} (6)

can be obtained, where $\hat{Y}_{p\times N} = [\hat{y}[0] \; \hat{y}[1] \cdots \hat{y}[N - 1]]$ and $\hat{X}_{C\times N} = [\hat{x}[0] \; \hat{x}[1] \cdots \hat{x}[N - 1]]$.

**Figure 1: Basic multicest undersampling scheme with data decimation and FFT, the detailed demonstration of data decimation is shown in subsection 2.4.**

### 2.2 Compressed sensing reconstruction

In this spectrum sensing scheme, equation (6) is a set of underdetermined equations and has infinite number of solutions. To apply CS recovery, the input signal $x(t)$ needs to be sparse, or more strictly speaking, compressible in frequency domain, which means only a few channels are occupied in all the $C$ channels. That is to say $\hat{x}[k]$ has only a few nonzero elements and $\hat{x}$ is a row-sparse matrix. In practice, such a noise-free signal is unattainable, so $\hat{Y}$ is actually a measurement with additive noise $A\mathbf{X} + \mathbf{N}$, where $\mathbf{N}$ denotes the Gaussian noise whose columns are independent and identically distributed as $\mathcal{N}(0, \sigma^2I)$. On the other hand, the delays of multisets sampler in different cosets should stay distinct ($c_{p1} \neq c_{p2}$ for $p1 \neq p2$), which makes $A$ a partial Fourier basis with natural mutual coherence.

Under the circumstances above, eq.(6) turns into a basic CS problem aiming at solving a row-sparse representation $\hat{X}$ of the original signal from the compressed measurements $\hat{Y}$ under dictionary $A$, i.e.

$$\arg\min_{\hat{X} \in \mathbb{C}^{C\times N}} ||\hat{X}||_{l_2,1} \text{ s.t. } ||\hat{Y} - A\hat{X}||_{l_2} < \epsilon,$$  \hspace{1cm} (7)

where $|| \cdot ||_{l_2,1}$ denotes the $l_{2,1}$ norm of the matrix and $\epsilon$ defines the tolerance related to the noise level. Solving such a problem requires optimization algorithms which could bring
in huge computational cost. In contrast, greedy algorithms divide the problem into several parts and seek for local optimal solutions part by part. Greedy algorithms cost far less computation burden than optimization algorithms and yield acceptable solution in high-SNR scenarios. Thus, greedy algorithms, instead of optimization algorithms are adopted in the proposed scheme for better real-time performance. A basic greedy algorithm for MMV problem is Simultaneous Orthogonal Matching Pursuit (SOMP) [9]. Given an initial condition $\mathbf{X}_0 = O_{CN}$ empty initial support $S_0$, starting at $k = 1$, it iterates

$$
\mathbf{R}_k = \mathbf{Y} - \mathbf{A}\mathbf{X}_{k-1}, \quad (8a)
$$

$$
S_k = S_{k-1} \cup \arg \max_i \|\{\mathbf{A}\}_{i,i} \cdot \mathbf{R}_k\|_F, \; i \notin S_{k-1}, \quad (8b)
$$

$$
\{\mathbf{X}_k\}_{S_k} = \{\mathbf{A}\}_{S_k}^\dagger \mathbf{Y}, \quad (8c)
$$

where $\{\cdot\}_{i,i}$ denotes the pseudo-inverse of the matrix. The SOMP algorithm solves one non-zero row of the sparse matrix in each iteration. Given the sparsity $K$ of the original spectrum, the SOMP algorithm converges after the $K$th iteration.

Joint-block-sparse Hard-thresholding pursuit (J-BHTP) algorithm, comparing to the SOMP, seeks for the $\hat{K}$ largest rows of the matrix in Frobenius-norm (F-norm) for a given $\hat{K}$, the largest possible sparsity of the original spectrum in each iteration. It is proved that J-BHTP converges in finite iterations given normalized sensing matrix $\mathbf{A}$. With an initial condition $\mathbf{X}_0 = O_{CN}$ empty initial support $S_0$, starting at $k = 1$, it iterates

$$
\mathbf{R}_k = \mathbf{Y} - \mathcal{A}\mathbf{X}_{k-1}, \quad \mathbf{X}_k = O_{CN} \quad (9a)
$$

$$
\mathcal{T}_k \leftarrow \{\hat{K}\} \text{ largest row entries in F-norm of } \mathbf{X}_{k-1} - \mathcal{A}\mathbf{R}_k, \quad (9b)
$$

$$
\{\mathbf{X}_k\}_{\mathcal{T}_k} = \{\mathcal{A}\}_{\mathcal{T}_k}^\dagger \mathbf{Y}, \quad (9c)
$$

where $\mathcal{A}$ denotes the normalized sensing matrix proportional to the original matrix $\mathbf{A}$ to meet $\|\mathcal{A}\|_2 \leq 1$ and $\mathbf{Y}$ is the normalized version of the measurement matrix $\hat{Y}$ with the same proportion. The J-BHTP iteration ends when $\mathcal{T}_k = \mathcal{T}_{k-1}$. In high-SNR scenario and appropriate estimation of sparsity $\hat{K}$, it can be theoretically proved that J-BHTP converges with small number of iterations.

### 2.3 Bayesian Estimation of Spectrum Sparsity

Due to the uncertainty of the number of channel occupied, the spectrum sparsity is always unknown in practice. If inaccurate $k$ is given to greedy algorithms, the recovery will be erroneous. Thus, obtaining the prior knowledge of the spectrum sparsity $k$ is of great importance for fast and satisfactory reconstruction.

Recall the MMV model with noise

$$
\hat{\mathbf{Y}} = \mathbf{A}\mathbf{X} + \mathbf{N}, \quad (10)
$$

Calculate the auto-correlation matrix of $\hat{\mathbf{Y}}$ and $\hat{\mathbf{X}}$

$$
\mathbf{R}_\mathbf{Y} = \mathbf{A}^{-1}\mathbf{R}_\mathbf{X}\mathbf{A}^H + \sigma^2 \mathbf{I}, \quad (11)
$$

where $\mathbf{R}_\mathbf{Y} = E[\mathbf{Y}\mathbf{Y}^H] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbf{y}[n]\mathbf{y}[n]^H$ and $\mathbf{R}_\mathbf{X} = E[\mathbf{x}\mathbf{x}^H] = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{n=1}^{N} \mathbf{x}[n]\mathbf{x}[n]^H$. In practice, only finite sample length $N$ can be achieved, eq.(11) is written as

$$
\hat{\mathbf{R}}_\mathbf{Y} = \mathbf{A}^{-1}\hat{\mathbf{R}}_\mathbf{X}\mathbf{A}^H + \frac{1}{N}\mathbf{N}\mathbf{N}^H, \quad (12)
$$

where $\hat{\mathbf{R}}_\mathbf{Y}$ and $\hat{\mathbf{R}}_\mathbf{X}$ are the approximate values of $\mathbf{R}_\mathbf{Y}$ and $\mathbf{R}_\mathbf{X}$ calculated by finite length of samples.

Perform eigendecomposition to $\hat{\mathbf{R}}_\mathbf{Y}$,

$$
\hat{\mathbf{R}}_\mathbf{Y} = \sum_{i=1}^{P} \hat{\lambda}_i \hat{\mathbf{v}}_i \hat{\mathbf{v}}_i^H, \quad (13)
$$

where $\hat{\lambda}_1 \geq \hat{\lambda}_2 \geq \cdots \geq \hat{\lambda}_P$ and $\hat{\mathbf{v}}_1, \hat{\mathbf{v}}_2, \cdots, \hat{\mathbf{v}}_P$ denote the descending-order eigenvalues of $\mathbf{R}_\mathbf{Y}$ and corresponding eigenvectors. Both signal and noise subspace are included in different eigenvalues of $\hat{\mathbf{R}}_\mathbf{Y}$. By estimating the amplitudes of $\hat{\lambda}_1$, the eigenvalues and subspace corresponding to signal components can be differentiated from those only representing for noise subspace, and the sparsity order is obtained by the dimensionality of signal subspace [7].

The problem of spectrum sparsity estimation is now how to find the integer $1 < k < P$ that best split two groups of eigenvalues of signal and noise subspace, given an estimate of autocorrelation matrix $\hat{\mathbf{R}}_\mathbf{Y}$ and its eigendecomposition. It is proposed in [7] that some information theoretical criteria, such as enhanced Bayesian information criterion (BICe) can be used where the the following is defined as the objective function of finding the best $k$

$$
\text{BICe}(k) = -2 \log f(\hat{\mathbf{Y}}|\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_P, k) - 2 \log f(\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_P | k) + C_k \log N, \quad (14)
$$

where $\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_P$ is the maximum-likelihood (ML) estimation for the theoretical eigenvalues $\lambda_1, \lambda_2, \cdots, \lambda_P$ and $f(\cdot)$ denotes the probability density function and $C_k = k(2P - K)$ is the number of free parameters related to $k$. Here, the range of integer $k$ is bounded to $[1, P]$. The estimation of sparsity $k$ can be obtained by solving the problem

$$
k = \arg \min_{k} \text{BICe}(k), \quad (15)
$$

The posterior probability of $N$ independent observations $\hat{\mathbf{Y}}$, i.e. $f(\hat{\mathbf{Y}}|\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_P, k)$ is expressed by a multivariate Gaussian model [3]. The joint probability density term, i.e. $\log f(\hat{\lambda}_1, \hat{\lambda}_2, \cdots, \hat{\lambda}_P)$ is proposed in [5] to be approximated by the product of the probability density of signal subspace and
the noise subspace eigenvalues, with the ML estimates to approximate true values as conditions

\[ f(\hat{\lambda}_1, \ldots, \hat{\lambda}_P; k) = f(\hat{\lambda}_1, \ldots, \hat{\lambda}_k | \lambda_1 = \hat{\lambda}_1, \ldots, \lambda_k = \hat{\lambda}_k) \cdot f(\hat{\lambda}_{k+1}, \ldots, \hat{\lambda}_P | \sigma^2 = \hat{\sigma}^2). \]  

(16)

where \( \hat{\sigma}^2 := \sum_{i=k+1}^{P} \hat{\lambda}_i / (P-k) \) denotes the maximum-likelihood estimation of noise subspace eigenvalues. Probability densities on the right hand side, i.e. \( f(\hat{\lambda}_1, \ldots, \hat{\lambda}_k | \lambda_1, \ldots, \lambda_k) \) and \( f(\hat{\lambda}_{k+1}, \ldots, \hat{\lambda}_M | \sigma^2) \) are derived in [3] and [10] respectively. Finally, the estimation can be solved by

\[ \hat{k} = \arg \min_k \text{BICe}(k) \]  

(17)

2.4 Data Decimation

The spectrum sensing model can be concluded as follows. Firstly, the baseband analog signal is sampled in \( P \) ways by the multicoset sampler with different random delays with each other. Then the \( P \)-way digital signal was windowed into \( N \)-point frames and transformed by parallel FFT operation. The FFT results are concatenated by rows and weighted to \( \tilde{Y} \). The Bayesian estimator gives the signal sparsity \( \hat{k} \) based on information in \( \tilde{Y} \). At last, greedy algorithms SOMP as well as JB-HTP are applied on \( \tilde{Y} \) according to \( \hat{k} \) to obtain the spectrum recovery \( \tilde{X} \).

In the above process, the number of samples in each coset \( N \) will significantly influence the computational cost in the recovery process. The time complexity of each operation mentioned above is shown in Table 1. 

<table>
<thead>
<tr>
<th>Operation</th>
<th>Time complexity</th>
</tr>
</thead>
<tbody>
<tr>
<td>N-point FFT</td>
<td>( O(N \log N) )</td>
</tr>
<tr>
<td>SOMP</td>
<td>( O(k^2NP) ) (eq.(8a)~(8c))</td>
</tr>
<tr>
<td>JB-HTP</td>
<td>( O(kNP) ) (eq.(9a)~(9c))</td>
</tr>
<tr>
<td>Sparsity estimation</td>
<td>( O(P^3) ) (eq.(13))</td>
</tr>
</tbody>
</table>

The computational cost will be notably reduced if less samples are processed in one frame of data. However, the FFT spectrum suffers from leakage when the window narrows down. Thus, “data decimation” is proposed to reduce the computational cost without lost of accuracy. Define the decimation factor \( d \), the N-point time-domain data \( x_{c_0}[n], n = 0, 1, \ldots, N-1 \) can be evenly divided into \( d \) equi-long fractions (\( N \) is chosen to be capable of being exactly divided by \( d \)). By making all the fractions overlapped and added with each other, the original N-point data can be reduced to a length of \( \frac{N}{d} \)-point data \( \sum_{i=0}^{d-1} x[n + iN/d], n = 0, 1, \ldots, N/d - 1 \) without discarding any symbol. Denoting \( k_d \) as the symbol index of \( \frac{N}{d} \)-point DFT, the DFT of the decimated signal is

\[ X_d[k_d] = \sum_{n=0}^{N/d} x[n] \exp(-j2\pi dk_d \frac{n}{N}) \]  

(18)

Comparing Equation 18 with the original N-point DFT \( X[k] = \sum_{n=0}^{N-1} x[n] \exp(-j2\pi nk/N) \), it can be founded that \( X_d[k_d] = X[dk] \), which means the DFT of decimated signal is a down-sampling of the DFT of original signal, and no additional aliasing or accuracy loss is introduced because all the amplitude information is retained during decimation. With only acceptable sacrifice of resolution, the computational cost can be largely reduced and the real-time performance of the system can be improved.

3 NUMERICAL EVALUATION

3.1 Noise performance

By extracting elements from the reconstructed matrix \( \tilde{X} \) by rows and tent them together as a vector, the reconstructed spectrum \( \tilde{X}_f \) is obtained. Consider a complex [-1.536GHz, 1.536GHz] base-band signal with 3.072GHz bandwidth evenly divided into 50 channels where 2 channels with 61.44MHz bandwidth are active. With channel SNR=-10dB, both greedy algorithms successfully detect the two activate channels (Figure 2 (b), Figure 2 (c)).

![Figure 2](image-url)

(a) Spectrum of received signal with noise \( x_f \), (b) spectrum reconstructed by SOMP \( \tilde{x}_f \), (c) spectrum reconstructed by JB-HTP \( \tilde{x}_f \) with window length \( l = 10000 \), SNR = -10dB, bandwidth \( B = 3072MHz \), channel number \( C = 50 \) with 2 active channels -1075.2MHz, 768MHz, coset number \( P = 8 \) and spectrum sparsity \( k = 2 \).
Comparing to the spectrum $\mathbf{x}_f$ of the received signal (Figure 2(a)) obtained by FFT, both $\mathbf{\tilde{x}}_f$ contains non-zero values only in the $k$ active channels while has all zero values in the unoccupied channels, which is because both greedy algorithms only reconstruct $k$ channels while regard other channels as zero. With accurate estimation $\mathbf{\tilde{k}}$ of spectrum sparsity $k$, the capability of eliminating noise in unoccupied channels of CS recovery makes it a noise-reduction process.

Define the normalized mean square error (NMSE) as

\[
NMSE = \frac{\|\mathbf{\tilde{x}}_f - \mathbf{x}_0\|_2^2}{\|\mathbf{x}_0\|_2^2}.
\]

By varying the SNR from $-20$dB to $50$dB, the NMSE of SOMP and JB-HTP are shown in Figure 3 (a). When $SNR > 0$, both algorithms show lower NMSE than the direct FFT, which is because of the elimination impact on inactive channels. When $-20$dB $< SNR < 0$dB, SOMP performance remains lower, while the NMSE of JB-HTP occasionally reaches higher values than SOMP. This is because as noise level increasing, it begins more difficult for JB-HTP to recognize all the active channels accurately in a single iteration. By contrast, SOMP searches for the maximum-possible active channel in each iteration, which helps reduce the false-detection probability in low-SNR scenario.

Define detection probability

\[
P_d = \frac{\text{total # correctly detected channels}}{\text{total # active channels in all recovery trials}}.
\]

The $P_d$ test result is shown in Figure 3 (b) under 1000 times of Monte-Carlo test for each integer SNR value in $[-20 dB, 50 dB]$. 2 of 50 channels are occupied and Gaussian white noise are added. When $SNR \geq -1$dB, both algorithms can achieve 100% reliable reconstruction. Then the detection probability of SOMP drops later but steeper than that of JB-HTP as SNR decreasing. In very-low-SNR situation (in this case $SNR < -10$dB), JB-HTP can achieve higher $P_d$, while in situation with higher SNR, SOMP offers larger $P_d$.

### 3.2 Sparsity performance

Sparsity of the signal in frequency domain is a basic prerequisite for the proposed CS reconstruction scheme. Taking a 50-channel 8/16-coset multicoset samplers as example, the relationship between detection probability and sparsity under different SNR is shown in Figure 4 (a), (b). SOMP shows acceptable detection probability for small sparsity values and loses accuracy on active channel detection with decreasing SNR or coset number. JB-HTP shows similar patterns with SNR and $P$ but performs better in detecting the active channels than SOMP, especially for high-SNR scenario (it keeps $100\%$ $P_b$ for all $k \in \{1, \ldots, 20\}$). For JB-HTP, despite more accurate $P_d$, the recovery performance of the signal amplitude is worse, which is reflected in the NMSE-sparsity curves in Figure 4 (c), (d). It is worth noting that a peak on NMSE curve of JB-HTP occurs at sparsity values about half the channel numbers as SNR dropping down. This is mainly caused by the synergistic effect of false support set detection and unreliable amplitude recovery.

### 4 EXPERIMENTAL IMPLEMENTATION

The hardware implementation for spectrum sensing is shown in Figure 5. Two NI mmWave SDR systems are used as the transmitter and receiver respectively in our implementation. Both the transmitter and receiver have modular configurable
We hereby demonstrated a real-time multi-gigahertz sub-Nyquist spectrum sensing system for mmWave bands implemented on NI mmWave SDR platform. Low-complexity sparse spectrum recovery algorithms has been designed with a practical sparsity estimation scheme. Data decimation was developed to largely reduce the time-complexity of recovery algorithms and facilitate real-time processing without loss of amplitude accuracy.

Acknowledgments
Support by the Engineering and Physical Sciences Research Council of United Kingdom under the Grant EP/R00711X/1.

References