Supplementary Document for Paper: Numerical Evaluation on Sub-Nyquist Spectrum Reconstruction Methods

1 Signal and System Models

1.1 Multiband Signal

The class of multiband signals is essential in various communication scenarios. In this context, we adopt the widely-known multiband signal model. This model describes N_{sig} complex transmissions situated in the baseband of interest $\mathcal{F} = [-f_{\text{Nyq}}/2, f_{\text{Nyq}}/2]$, where f_{Nyq} is the baseband Nyquist frequency. Each transmission, without loss of generality, is assumed to possess a bandwidth no larger than B and does not overlap with other transmissions.

A wideband cognitive receiver, which passively senses the spectrum over $\mathcal{F}$, receives the baseband signal represented in the temporal domain by:

$$x(t) = \sum_{i \in N_{\text{sig}}} s_i(t) + n(t), \tag{1}$$

where $s_i(t)$ denotes the i^{th} transmission signal with incorporated channel effects, and $n(t)$ signifies additive Gaussian noise. We further assume that the baseband signal exhibits sparsity in the frequency domain, which means $N_{\text{sig}}B \ll f_{\text{Nyq}}$.

1.2 Multicoset Sampling

The cognitive receiver is assumed to be equipped with a well-known periodic non-uniform sampling architecture, known as the multicoset sampler [7], to acquire the baseband signal samples at a sub-Nyquist rate. The basic structure of a multicoset sampler is shown in Fig. 1. A power splitter divide the analog signal into p ways (referred as "cosets" afterwards) and impose different analog delays $\tau_i = c_i/f_{\text{Nyq}}$ specified by the delay pattern $C = \{c_i \in \mathbb{Z} \mid 0 \leq c_i < L, i = 1, 2, ..., p\}$ to each coset of the signal. Then each coset of the signal is digitized by a low-rate ADC working at a sampling rate $f_{\text{s}} = f_{\text{Nyq}}/L$, where integer L satisfies $L > p$ to ensure compressed sampling. The sample sequence acquired in the i^{th} coset is expressed as

$$x_i[n] = x\left(\frac{nL}{f_{\text{Nyq}}} + \frac{c_i}{f_{\text{Nyq}}}\right), \quad n \in \mathbb{Z}. \tag{2}$$

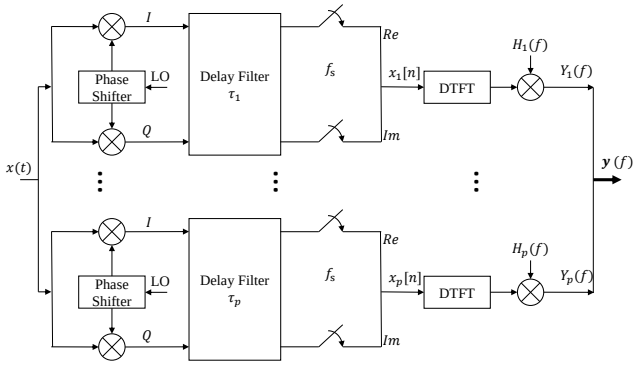


Figure 1: The diagram of the basic implementation of multicoset sampler.

It makes sense that the samples we obtain from each coset are a downsampled version of the Nyquist samples with a downsampling coefficient L, and the overall average sampling rate of the multicoset sampler is pf_{Nyq}/L. By denoting the i^{th} measurement as

$$Y_i(f) = \frac{L}{f_{\mathrm{Nyq}}} e^{-j2\pi c_i T_{\mathrm{Nyq}} f} X_i(e^{j2\pi fL/f_{\mathrm{Nyq}}}) \tag{3}$$

where $X_i\left(e^{j2\pi fL/f_{\mathrm{Nyq}}}\right) = \sum_{n=-\infty}^{\infty} x_i[n] e^{-j2\pi fLn/f_{\mathrm{Nyq}}}$ is the discrete-time Fourier transform (DTFT), we can establish a linkage between the compressed measurement and the original spectrum $X(f)$, given as

$$Y_i(f) = \sum_{n=-\lfloor \frac{L}{2} \rfloor}^{+\lfloor \frac{L}{2} \rfloor - 1} X\left(f + \frac{nf_{\mathrm{Nyq}}}{L}\right) \exp\left(j2\pi \frac{nc_i}{L}\right). \tag{4}$$

Equation (4) indicates that the measurement $Y_i(f)$ in the i^{th} coset can be represented by a linear combination of L separate samples on the signal's original spectrum $X(f)$. By stacking the p single-coset measurements columnwise we get the matrix form the linkage

$$\mathbf{y}(f) = \mathbf{A}\mathbf{x}(f) \tag{5}$$

where $\mathbf{y}(f) = [Y_1(f) \quad Y_2(f) \quad ... \quad Y_p(f)]^{\mathrm{T}}$ and $\mathbf{x}(f)$ is a column vector of length L composed of $\{X(f + nf_{\mathrm{Nyq}}/L) | n = 0, 1, ..., L-1\}$.

The coefficient matrix $\mathbf{A}$ is a $p \times L$ matrix whose ilth element is given by

$$A_{i,l} = e^{j2\pi c_i(-2/L + l - 1)/L}. \tag{6}$$

In practice, a finite window is usually employed and discrete Fourier transform (DFT) is used instead of DTFT. Denoting the window length as N per coset, a multiple measurement vectors (MMV) model

$$\mathbf{Y} = \mathbf{A}\mathbf{X} \tag{7}$$

can be derived, where $\mathbf{Y}$ is a $p \times N$ measurement matrix whose $(i,n)^{\text{th}}$ element is

$$Y_{i,n} = Y_i\left((n-1)f_{\text{Nyq}}/NL\right), \tag{8}$$

and $\mathbf{X}$ is a $L \times N$ matrix whose $(l,n)^{\text{th}}$ element is

$$X_{l,n} = X\left(\frac{-L/2 + l - 1 + (n-1)/N}{L/f_{\text{Nyq}}}\right). \tag{9}$$

1.3 Problem Formulation

Equation (7) can be viewed as a decomposition problem where $\mathbf{A}$ is a $p \times L$ dictionary matrix, $\mathbf{Y}$ is a $p \times N$ measurement matrix consisting of N individual p-dimensional measurement vectors, and the $L \times N$ matrix $\mathbf{X}$ houses the coefficients of columns in $\mathbf{A}$ that need estimation.

For $P < L$, $\mathbf{A}$ becomes an overcomplete dictionary, making (7) indeterminate with countless $\mathbf{X}$ solutions. Given the block-sparse nature of the spectrum, $\mathbf{X}$ is typically seen as a joint-sparse coefficient matrix with each column sharing an identical support. Consequently, resolving (7) transforms into a ℓ_0-norm minimization challenge:

$$\min_{\mathbf{X} \in \mathbb{C}^{L \times N}} \quad \|\mathbf{X}\|_0 \quad \text{s.t.} \quad \mathbf{Y} = \mathbf{A}\mathbf{X}.$$

In real scenarios, the measurement $\mathbf{Y}$ invariably incurs noise disturbances. Thus, a relaxation on the aforementioned equation introduces a ℓ_F-norm-bounded error tolerance ϵ:

$$\arg\min_{\mathbf{X} \in \mathbb{C}^{L \times N}} \quad \|\mathbf{X}\|_0 \quad \text{s.t.} \quad \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_{\mathsf{F}} < \epsilon.$$

Although addressing both of these equations directly is NP-hard, a plethora of computationally feasible algorithms have been proposed to approximate these sparse approximation problems. Notably, several of them have been implemented in practice, yielding reliable results.

In subsequent sections, we offer a cursory overview of three primary types of sparse approximation methods. Furthermore, we present one type of non-sparse approximation method derived from sub-Nyquist samples, with a detailed examination of symbolic algorithms contributed by GBSense Challenge participants.

2 Algorithms

2.1 Greedy Algorithms

A detailed analysis of greedy iterative algorithms for combinatorial optimization problems can be found here. These algorithms use the connection between the signal and the atomic dictionary to gauge an atom's effectiveness. A few notable greedy algorithms are matching pursuit (MP), orthogonal matching

Algorithm 1 SA-SOMP algorithm [2]

Input: $\mathbf{Y}$; $\mathbf{A} \in \mathbb{C}^{p \times L}$; $\hat{\kappa}$
Output: $\hat{\mathcal{S}}$, $\hat{\mathbf{\Theta}}$

$\mathbf{R} \leftarrow \mathbb{E}[\mathbf{Y}(f)\mathbf{Y}^{\mathrm{H}}(f)]$
$[\mathbf{U}_s, \mathbf{\Lambda}_s] \leftarrow \mathrm{RREVD}(\mathbf{R}, \hat{\kappa}), \chi_s \leftarrow \mathbf{U}_s\sqrt{\mathbf{\Lambda}_s}$
$t \leftarrow 1$; $\mathbf{R}_0 \leftarrow \chi_s$; $\hat{\mathcal{S}}_0 \leftarrow \varnothing$;$\mathbf{A}_0 \leftarrow \varnothing$;
while $t \leq \hat{\kappa}$ **do**
    $\hat{\alpha}_t \leftarrow \arg\max_i \left\|\mathbf{A}_{:,i}^{\mathrm{H}}\mathbf{R}_{t-1}\right\|_2$, $i = 0, 1, \cdots, L-1$
    $\hat{\mathcal{S}}_t \leftarrow \hat{\mathcal{S}}_{t-1} \cup \hat{\alpha}_t$
    $\mathbf{A}_t \leftarrow \mathbf{A}_{\hat{\mathcal{S}}_t}$
    $\hat{\mathbf{\Theta}}_t \leftarrow \arg\min_\nu \left\|\mathbf{Y} - \mathbf{A}_t\hat{\mathbf{\Theta}}_t\right\|_2$
    $\mathbf{R}_t \leftarrow \mathbf{Y} - \mathbf{A}_t\hat{\mathbf{\Theta}}_t$
    $t \leftarrow t + 1$
end while
$\hat{\mathcal{S}} \leftarrow \hat{\mathcal{S}}_t$
fill the rows of $\hat{\mathbf{\Theta}}$ indexed by $\hat{\mathcal{S}}$ with $\hat{\mathbf{\Theta}}_t$

pursuit (OMP), stagewise orthogonal matching pursuit (StOMP), compressive sampling matching pursuit (CoSaMP), iterative hard thresholding (IHT), and gradient descent with sparsification (GraDeS) [3–6]. The algorithm's complexity mainly depends on the iterations required to determine the correct support set.

The JB-HTP algorithm [4] and SA-SOMP [2] are also elaborated upon, with SA-SOMP's full procedure being:

2.2 Optimization Algorithms

The l_0 optimization problem, its relaxation, and various strategies such as the basic pursuit (BP) strategy, FOCUSS algorithm, GPSR, and SpaRSA are all detailed. The "slow kill" (SK) algorithm by She et al. considers a specific optimization problem, represented by:

$$\arg\min_{\mathbf{X} \in \mathbb{C}^{L \times N}} \quad \|\mathbf{Y} - \mathbf{A}\mathbf{X}\|_F^2 + \frac{\eta_0}{2}\|\mathbf{X}\|_F^2 \quad \text{s.t.} \quad \|\mathbf{X}\|_{2,0} \leq \hat{\kappa}. \tag{10}$$

2.3 Bayesian Framework

The Bayesian framework's reconstruction algorithm considers the time correlation of signals. Key algorithms under this category include the Expectation-Maximization (EM) algorithm, Bayesian compressive sensing (BCS) algorithm, sparse Bayesian learning (SBL) algorithm, and MSBL [1, 8, 9, 11].

Algorithm 2 MSBL algorithm [11]

Input: $\mathbf{Y}$; $\mathbf{A} \in \mathbb{C}^{p\times L}$; σ^2; θ, $maxIter$
Output: $\hat{\mathcal{S}}$, $\hat{\mathbf{\Theta}}$

$\boldsymbol{\gamma} \leftarrow \mathbf{1}$; $\mathbf{\Gamma} \leftarrow \mathrm{diag}(\boldsymbol{\gamma})$; σ^2
while $t \leq maxIter$ **do**
  $\boldsymbol{\mu} \leftarrow \mathbf{\Gamma}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{A}\mathbf{\Gamma}\boldsymbol{A}^{\mathrm{H}} + \sigma^2\mathbf{I})^{-1}\mathbf{Y}$
  $\mathbf{\Sigma} \leftarrow \mathbf{\Gamma} - \mathbf{\Gamma}\boldsymbol{A}^{\mathrm{H}}(\boldsymbol{A}\mathbf{\Gamma}\boldsymbol{A}^{\mathrm{H}} + \sigma^2\mathbf{I})^{-1}\boldsymbol{A}\mathbf{\Gamma}$
  $\gamma_i \leftarrow \frac{\frac{L}{N}\|\boldsymbol{\mu}(i,:)\|^2}{1-\mathbf{\Sigma}(i,i)/\gamma_i}$, for $i = 1, 2, \cdots, L$
  $\sigma^2 \leftarrow \frac{\frac{L}{N}\|\boldsymbol{Y}-\boldsymbol{A}\boldsymbol{\mu}\|_F^2}{p-L+\mathrm{Tr}(\mathbf{\Gamma}^{-1}\mathbf{\Sigma})}$
  $\mathbf{\Gamma} \leftarrow \mathrm{diag}(\boldsymbol{\gamma})$
end while
Let $\hat{\mathcal{S}}$ be the set of all index i such that $\gamma_i \geq \frac{\theta}{L}\sum_{i=1}^{L}\gamma_i$
$\hat{\mathbf{\Theta}}(\hat{\mathcal{S}},:) \leftarrow \boldsymbol{\mu}(\hat{\mathcal{S}},:)$, $\hat{\mathbf{\Theta}}(\hat{\mathcal{S}}^c,:) \leftarrow \mathbf{0}$

Algorithm 3 Fast compressed power spectrum estimation [10]

Input: $\mathbf{Y}$; C;
Output: $\hat{\boldsymbol{\theta}}$

$h[n]\begin{cases} Y(i,j), & n = jN + c_i \\ 0, & \text{otherwise} \end{cases}$, $I[n]\begin{cases} 1, & n = jN + c_i \\ 0, & \text{otherwise} \end{cases}$
$\overline{h}[n]\begin{cases} h[n], & 0 \leq n \leq LN-1 \\ 0, & -LN+1 \leq N < 0 \end{cases}$, $\overline{I}[n]\begin{cases} I[n], & 0 \leq n \leq LN-1 \\ 0, & -LN+1 \leq N < 0 \end{cases}$
$\boldsymbol{r}_h = \mathbb{F}^{-1}(|\mathbb{F}(\overline{\boldsymbol{h}})|^2)$, $\boldsymbol{q} = \mathbb{F}^{-1}(|\mathbb{F}(\overline{\boldsymbol{I}})|^2)$
$r_x[k] = r_h[k]/q[k]$, $k = -LN+1, -LN+2, \cdots, LN-1$
$\hat{\boldsymbol{\theta}} = \mathbb{F}(\boldsymbol{r}_x)$

2.4 Non-sparse reconstruction

The fast compressed power spectrum estimation (FCPSE) provides an alternative for power spectrum estimation:

3 Numerical Results

The algorithms shortlisted for the GBSense Challenge 2021 have undergone various numerical tests for a thorough evaluation based on the competition's criteria. In this section, typical detection performance and time complexity data are shown.

Using a simulated 16-channel multicoset sampler($p = 16$), each piece of data in the aforementioned datasets is sampled. Each channel's sampling rate is 1/40 of the Nyquist rate ($L = 40$). The codes of the shortlisted entries are embedded to reconstruct the spectrum and of each piece of data. To avoid the reconstruction errors imposed by the delay pattern ,which are not the subject of this evaluation work, a large number of randomly generated delay patterns

are utilized in the simulation, and their correspondintg results are averaged. Specifically, the test result on each piece of data is averaged using 100 random sampling delay patterns (except for FCPSE, which has a uniquely developed delay pattern creation module).

For each dataset, the reconstruction accuracies of the shortlisted algorithms are measured by the average Area Under Curve (AUC) of the receiver operating characteristic curve (ROC).

To reveal comprehensive features of the shortlisted algorithms, additional simulations are conducted. Typical ROC curves for the shortlisted algorithms under the aforementioned settings and 10dB SNR are shown in Figure 2. For the $k = 6$ case in Figure 2(a), the SBL and SK methods generally outperform the other greedy algorithms and non-sparse FCPSE algorithm. That is because the strong time correlation of the signal is utilized by SBL, and the global minimum is better approached by optimization problems. For the $k = 8$ case in Figure 2(b), a significant performance drop occurs for the SK algorithm. Note that $k = 8$ is the theoretical upper bound for ensuring a unique solution with $p = 16$, which means this direct l_0-minimization strategy is vulnerable to sparsity change. It is also worth noting that JB-HTP tends to have better detection performance than SOMP when k increases because the misselected atoms can be fixed as the iteration of JB-HTP goes on, while a mismatch in SOMP cannot be excluded in the following iterations and will lead to slow convergence. SA-SOMP achieves exactly the same results as SOMP does because they share the same reconstruction strategy, and their only difference is the choice of the measurement matrix.

The shortlisted algorithms are tested for their detection probabilities defined by

$$P_d = \frac{\text{total \# correctly detected channels}}{\text{\# total active channels in all recovery trials}}. \tag{11}$$

under multiple SNR levels with $L = 40$, $p = 16$ and signal sparsity of $k = 6$, and the results are shown in Figure 3. Note that because FCPSE provides a non-sparse recovery, an additional screening of the $\hat{\kappa}$ channels with the highest energy are kept as the recovered support. Among the curves in Figure 3, SK is the most sensitive method to noise which only achieve detection probability over 0.9 when $SNR \geq 10dB$. On the contrary, FCPSE is more resilient to noise level changes because Gaussian noise can be effectively damped by its statistical nature. However, detection probability above 0.9 is difficult to be achieved by FCPSE because the unbiased estimation is not always a good approximation, especially with sub-Nyquist samples and narrow sampling windows. SBL algorithm can achieve nearly 0.95 detection probability when $SNR \geq 5dB$ and 100% when $SNR \geq 10dB$, which outperforms the other algorithms within the simulation data.

For signal with different sparsity k from 1 to 16, simulations are conducted with $L = 40$, $p = 16$ and $SNR = 20dB$, and the results are shown in Figure 4. Taking advantage of the strong time correlation of the testing data, SBL outperforms the other algorithms on detection probability. The SK algorithm

Table 1: Average AUC & running time on the GBSense test sets

Algorithm	**AUC**	**Time (ms)**
SBL	0.9771	3.5103
SA-SOMP	0.8190	1.3999
SK	0.6277	23.9034
FCPSE	0.8631	1.3734

performs well under $k \leq 6$ but quickly deteriorates when k gets larger. SOMP and SA-SOMP provide better results than JB-HTP when $k < 8$ but get worse when $k > 8$ because the atomwise matching is more accurate when fewer atoms are contributing to the signal but unreliable when a large number of weighted atoms are mixed. FCPSE gives a relatively detection probability for $k > 8$ than the greedy algorithms and SK.

As spectrum sensing is a task of high real-time requirement, the time complexity of a reconstruction algorithm is of great significance to be considered taking into practice. The time efficiency of the algorithms is measured by the average running time of the codes in MATLAB simulation. The test results for the four shortlisted algorithms are provided in Table 1.

By varying the length N of the sub-Nyquist sampling window, the real running time of each shortlisted algorithm on the CPU is tested by MATLAB, and the results are shown in Figure 5 in logarithmic coordinates, which in general agrees with Table 1. SK, as a l_0-optimization method, has naturally computational complexity orders higher than the other candidates. The SA-SOMP method achieves exactly the same detection results as SOMP but only costs running time comparable with JB-HTP, thus can be regarded as a promising method to be applied in very sparse conditions to achieve both high detection rates and high real-time performance. SBL costs moderate calculation time among the tested algorithms. The running time of three greedy algorithms and SBL is not sensitive to the increase of sampling window length N because they process the columns in the coefficient matrix $\mathbf{X}$ as joint-sparse. By contrast, the running time of FCPSE increases nearly exponentially with N because of the large amount of convolution included.

References

[1] Shihao Ji, Ya Xue, and Lawrence Carin. Bayesian compressive sensing. *IEEE Transactions on signal processing*, 56(6):2346–2356, 2008.

[2] Yuan Ma, Xingjian Zhang, and Yue Gao. Joint sub-nyquist spectrum sensing scheme with geolocation database over tv white space. *IEEE Transactions on Vehicular Technology*, 67(5):3998–4007, 2017.

[3] Deanna Needell and Joel A Tropp. Cosamp: Iterative signal recovery from incomplete and inaccurate samples. *Applied and computational harmonic analysis*, 26(3):301–321, 2009.

[4] Haoran Qi, Xingjian Zhang, and Yue Gao. Low-complexity subspace-aided compressive spectrum sensing over wideband whitespace. *IEEE Transactions on Vehicular Technology*, 68(12):11762–11777, 2019.

[5] Joel A Tropp and Anna C Gilbert. Signal recovery from random measurements via orthogonal matching pursuit. *IEEE Transactions on information theory*, 53(12):4655–4666, 2007.

[6] Joel A Tropp, Anna C Gilbert, and Martin J Strauss. Simultaneous sparse approximation via greedy pursuit. In *Proceedings.(ICASSP'05). IEEE International Conference on Acoustics, Speech, and Signal Processing, 2005.*, volume 5, pages v–721. IEEE, 2005.

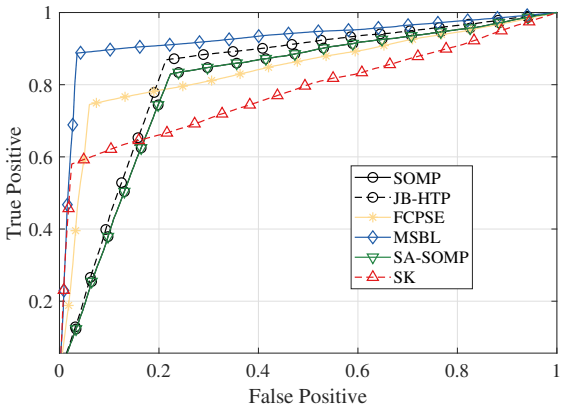
[7] Raman Venkataramani and Yoram Bresler. Perfect reconstruction formulas and bounds on aliasing error in sub-nyquist nonuniform sampling of multiband signals. *IEEE Transactions on Information Theory*, 46(6):2173–2183, 2000.

[8] David P Wipf and Bhaskar D Rao. Sparse bayesian learning for basis selection. *IEEE Transactions on Signal processing*, 52(8):2153–2164, 2004.

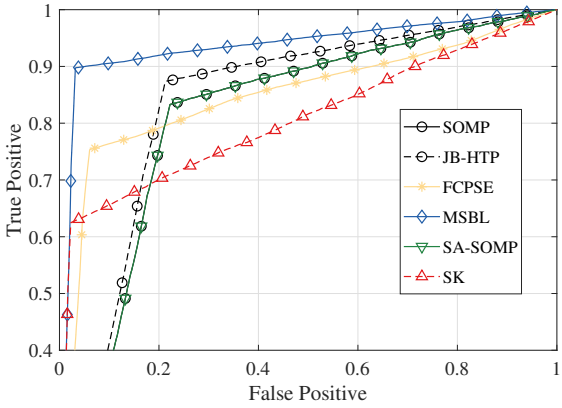
[9] David P Wipf and Bhaskar D Rao. An empirical bayesian strategy for solving the simultaneous sparse approximation problem. *IEEE Transactions on Signal Processing*, 55(7):3704–3716, 2007.

[10] Linxiao Yang, Jun Fang, Huiping Duan, and Hongbin Li. Fast compressed power spectrum estimation: toward a practical solution for wideband spectrum sensing. *IEEE Transactions on Wireless Communications*, 19(1):520–532, 2019.

[11] Hadi Zayyani, Massoud Babaie-Zadeh, and Christian Jutten. Decoding real-field codes by an iterative expectation-maximization (em) algorithm. In *2008 IEEE International Conference on Acoustics, Speech and Signal Processing*, pages 3169–3172. IEEE, 2008.



(a)



(b)

Figure 2: Typical ROCs for the GBSense shortlisted algorithms compared with traditional SOMP and JB-HTP ($L = 40$, $p = 16$) on signal with sparsity (a) $k = 6$ and (b) $k = 8$.

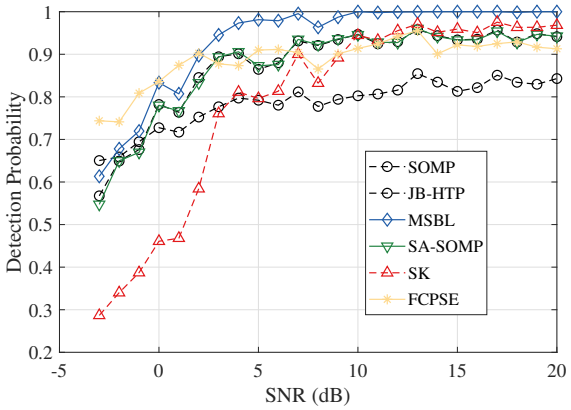


Figure 3: Detection probability versus signal SNR for the four shortlisted algorithms compared with traditional SOMP and JB-HTP

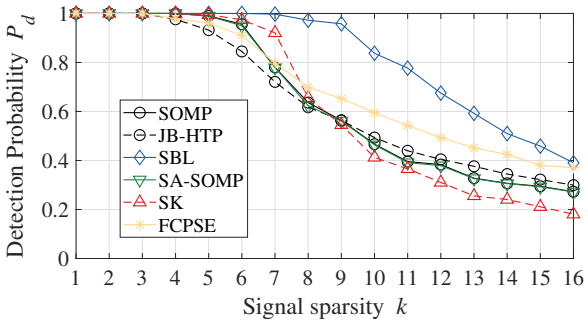


Figure 4: Detection probability versus coset number p for the four shortlisted algorithms compared with traditional SOMP and JB-HTP

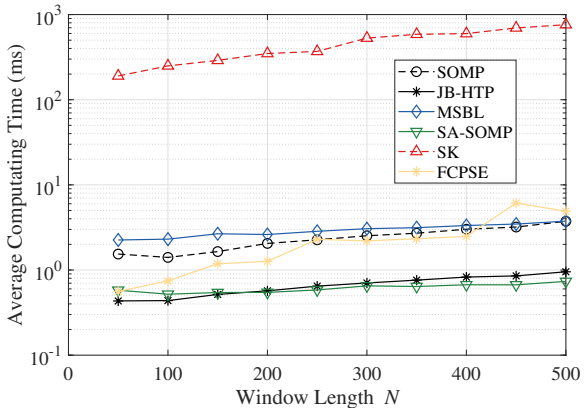


Figure 5: Average running time versus the length N of the sampling window.