Boundary Multiple Measurement Vectors for Multi-Coset Sampler

Dong Xiao, Jian Wang and Yun Lin

***Abstract*—The rapid increase in the bandwidth of radio signals has brought great challenges to multi-coset sampler (MCS). In this paper, we propose a boundary multiple measurement vectors (bMMV) model based on i) column-partitioning and ii) row-extraction on the measurement matrix, which can achieve the theoretical minimum sampling rate of MCS. Under this model, we develop an algorithm called side-information-aided simultaneous subspace pursuit (SI-SSP) to recover MCS signals. Specifically, SI-SSP iteratively utilizes the supports of previously estimated signal sub-matrices as side information to promote the recovery performance. Theoretical analysis shows that SI-SSP can perform stable recovery of MCS signals under a mild condition on the restricted isometry property (RIP). Numerical experiments shows that our algorithm has higher recovery accuracy than existing methods, especially in low sampling rate scenarios.**

***Index Terms*—Multi-coset sampling, sub-Nyquist sampling, multiple measurement vectors, sparse signal.**

I. Introduction

The 6G technology will require a huge number of spectrum resources in the future [1]–[4]. To fully utilize the spectrum resources, dynamic spectrum access (DSA) allows cognitive radio (CR) devices to work on idle frequency bands. Since spectrum is usually occupied by wide-band and multi-frequency signals, high-speed analog-to-digital converters (ADC) are required to sample the signals [5]–[8]. This, however, is rather costly in practice, thus motivating the need of sub-Nyquist sampling that works at a rate much lower than the Nyquist sampling rate. To realize the sub-Nyquist sampling, compressed sensing (CS) based techniques have been widely used owing to the sparsity of radio signals in the frequency domain [9]–[11]. In a nutshell, they exploit the ability of CS in reconstructing high-dimensional sparse signals from a small number of measurements [12], [13].

As a classical sub-Nyquist sampling technique based on CS, multi-coset sampler (MCS) has received much attention [14]–[16]. In multiple channels, MCS realizes non-uniform sampling on time-domain coset by sampling signals with different time delays. In reconstructing MCS signals, [17] proved that accurate recovery requires the minimum sampling rate to be at least twice the Lebesgue measure of the signal frequency support. In practice, however, the system sampling rate is often much higher than theoretically predicted, since the energy in a sparse signal may be scattered across various supports.

In order to reduce the sampling rate, [18] assumed the joint sparsity in the target signal, thus transforming the recovery task into a multiple measurement vectors (MMV) problem. To date, various recovery methods exploiting the joint sparsity prior has been suggested to solve the MMV problem. Examples include simultaneous orthogonal matching pursuit (SOMP) [19], [20], CS-MUSIC [21], SA-MUSIC [22], etc. Unfortunately, they still exhibit poor accuracy when the support is scattered over multiple rows of the signal (i.e., the high joint sparsity case), or when the sampling rate falls below a certain level.

In this paper, in order to improve the recovery accuracy of MCS signals in low sampling rate scenarios, we propose a boundary MMV (bMMV) model, along with a recovery algorithm called side-information-aided simultaneous subspace pursuit (SI-SSP). Our main ideas are as follows.

- **bMMV:** The multi-coset sampling of the jointly sparse signal matrix is implemented in two steps: i) column-partition the signal matrix into multiple sub-matrices, and keep only their boundary nonzeros while setting others to zero; ii) perform multi-coset sampling on each sub-matrix in parallel. In doing so, the joint sparsity of each sub-matrix is much lower than that of the entire signal matrix. Thus, this eases the reliance on high sampling rate, and in turn promotes the recovery accuracy.
- **SI-SSP:** In bMMV, the supports of adjacent signal sub-matrices are mostly overlapping. Hence, the estimated supports of some sub-matrices can be used as side information (SI) to facilitate support estimation of other sub-matrices. This improves the recovery accuracy of the individual sub-matrices, and thus the entire signal matrix.

We provide i) a condition for bMMV to attain the theoretical lower bound of sampling rate in MCS and ii) a condition for SI-SSP to stably recover MCS signals under noise. Numerical experiments confirm effectiveness of our proposal.

II. PRELIMINARIES

We summarize notations used throughout the paper. For a complex matrix $\mathbf{X} \in \mathbb{C}^{n\times L}$ and a set $S \subseteq \{1,\cdots,n\}$, $\mathbf{X}_S$ (or $\mathbf{X}^S$) denotes the submatrix of $\mathbf{X}$ with columns (or rows) indexed by S; $\mathbf{X}_{i,j}$, $\mathbf{X}_{i,:}$ and $\mathbf{X}_{:,i}$ are the (i,j)th entry, ith row and ith column of $\mathbf{X}$, respectively; $\mathbf{X}^\dagger$, $\mathbf{X}^H$ and $\mathbf{X}^\top$ mean the Moore-Penrose pseudo-inverse, conjugate transpose and transpose of $\mathbf{X}$, respectively; $\mathrm{supp}(\mathbf{X})$ is the non-zero row indices (i.e., joint sparsity) of $\mathbf{X}$; $\|\mathbf{X}\|_F$ and $\|\mathbf{X}\|_2$ signify the Frobenius and Euclidean norm of $\mathbf{X}$, respectively. Moreover, S^c is the complement of set S; $\mathbf{I}_L$ is an $L\times L$ identity matrix.

Now, we proceed to explaining the structure of MCS. A p-channel MCS is illustrated in Fig. 1. Consider a multi-band signal $x(t)$ with N_{sig} bands, and each has the same bandwidth

D. Xiao and Y. Lin are with Harbin Engineering University, Harbin, China. J. Wang is with Fudan University, Shanghai, China. E-mail: {xiaodong1,linyun}@hrbeu.edu.cn; jian_wang@fudan.edu.cn

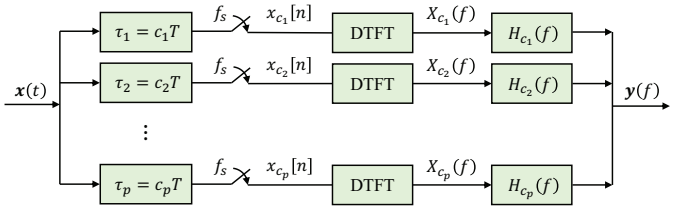

Fig. 1. An example of MCS structure with p channels

B. The frequency support $\mathcal{T}$ of $x(t)$ is defined as the union of frequency intervals: $\bigcup_{j=1}^{N_{\text{sig}}}[f_j^{\min}, f_j^{\max}]$, where $f_j^{\min}$ (or $f_j^{\max}$) is the minimum (or maximum) frequency of the jth band.

MCS uses p parallel channels that sample the signal uniformly at a decimated rate $f_s := \frac{1}{LT}$, where T is the Ny

quist sampling period and L is a decimation factor. The sampling sequence in the ith channel is then given by

$$x_{c_i}[n] = x\big(LTn + \tau_i\big), \quad n = 0, 1, \cdots \tag{1}$$

where $\tau_i := c_i T$ is the time delay of the ith channel. In particular, the delay coefficients c_i's of p channels satisfy $0 \le c_1 < \cdots < c_p \le L-1$. As a discrete signal, $x_{c_i}[n]$'s discrete time Fourier transform (DTFT) is given by

$$X_{c_i}\left(e^{\mathrm{j}2\pi fLT}\right) = f_s \sum_{m=0}^{L-1} X\left(f + mf_s\right) e^{\frac{\mathrm{j}2\pi m c_i}{L}}. \tag{2}$$

The frequency response $H_{c_i}(f)$ of $x_{c_i}[n]$ is low pass. In particular, $H_{c_i}(f) = \frac{1}{f_s}$ for $f \in [0, f_s)$ and zero otherwise. Then, the output $Y_{c_i}(f)$ of the ith channel is obtained as

$$Y_{c_i}(f) = X_{c_i}(f) H_{c_i}(f) = \sum_{m=0}^{L-1} X\left(f + mf_s\right) e^{\frac{\mathrm{j}2\pi m c_i}{L}}, \tag{3}$$

where $f \in [0, f_s)$, which is the linear combination of L non-overlapping parts of $X(f)$ in the frequency domain.

The output of p channels can be expressed as:

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

where $\mathbf{y}(f) := [Y_{c_1}^\top(f), Y_{c_2}^\top(f), \cdots, Y_{c_p}^\top(f)]^\top$, $\mathbf{x}(f) := [X(f)^\top, X(f+f_s)^\top, \ldots, X(f+(L-1)f_s)^\top]^\top$, and $\mathbf{A} \in \mathbb{C}^{p\times L}$ is the sensing matrix determined by the delay coefficients. Specifically, the (i,k)th entry of $\mathbf{A}$ is $\mathbf{A}_{i,k} = e^{\frac{\mathrm{j}2\pi c_i k}{L}}$. In practical digital systems, the analog signals $\mathbf{x}(f)$ and $\mathbf{y}(f)$ are transformed into discrete ones (i.e., $\mathbf{X} \in \mathbb{C}^{L\times N}$ and $\mathbf{Y} \in \mathbb{C}^{p\times N}$) via uniform sampling at regular intervals: $X_{i,j} = X\big((i-1+\frac{j-1}{N})f_s\big)$ and $Y_{i,j} = Y_{c_i}\big(\frac{j-1}{N}f_s\big) = \frac{1}{f_s}X_{c_i}\big(\frac{j-1}{N}f_s\big)$.

III. The bMMV Model

Our goal is to recover $\mathbf{X}$ from the MMV model:

$$\mathbf{Y} = \mathbf{A}\mathbf{X} + \mathbf{E}, \tag{4}$$

where $\mathbf{E} \in \mathbb{C}^{p\times N}$ is noise. Assume that $\mathbf{X}$ has M users, that is, $\mathbf{X} = [(\mathbf{X}^{U_1})^\top, (\mathbf{X}^{U_2})^\top, \cdots, (\mathbf{X}^{U_M})^\top]^\top$, where $\mathbf{X}^{U_j} \in \mathbb{C}^{\frac{L}{M}\times N}$ and $U_j := \big\{\frac{(j-1)L}{M} + 1, \cdots, \frac{jL}{M}\big\}$, $j = 1, \cdots, M$. Among M users, there are N_{sig} primary user (PU) signals.

We consider a typical setting of MCS where the frequency support $\mathcal{T}$ of $x(t)$ is unknown. In this case, the theoretical lower bound of sampling rate is $\min(2\lambda(\mathcal{T}), f_{\text{nyq}})$ [17], where $\lambda(\mathcal{T})$ is the Lebesgue measure of $\mathcal{T}$, and $f_{\text{nyq}} := \frac{1}{T}$ is the Nyquist sampling rate. Recalling that each PU signal's bandwidth is B, we have $\lambda(\mathcal{T}) = N_{\text{sig}}B$; thus, the theoretical lower bound of sampling rate becomes $\min(2N_{\text{sig}}B, f_{\text{nyq}})$.

For a p-channel MCS system, since the sampling rate of each channel is f_s, the actual sampling rate is pf_s. We are interested in whether pf_s can achieve the theoretical lower bound. Our answer is formally given in the following theorem. Due to page limitation, the proof is left to Supplementary [23].

Theorem 1. *The actual sampling rate of (4) is* $\min\big(pf_s, f_{nyq}\big)$*, which attains the theoretical lower bound of sampling rate in MCS when* $|supp(\mathbf{X})| \le \frac{N_{sig}B}{f_s}$.

Note that this condition may not hold when $|\text{supp}(\mathbf{X})|$ is large. To deal with this issue, we propose a new model called boundary MMV (bMMV), which consists of two operations.

- **Column-partitioning:** Decompose (4) into r sub-MMV problems and solve each problem individually:
$$\mathbf{Y}_{S_i} = \mathbf{A}\mathbf{X}_{S_i} + \mathbf{E}_{S_i}, \quad i = 1, \cdots, r, \tag{5}$$
where $S_i := \big\{(i-1)\lceil\frac{N}{r}\rceil + 1, \cdots, \min\{i\lceil\frac{N}{r}\rceil, N\}\big\}$, $i = 1, \cdots, r$, and $\mathbf{Y}_{S_i} \in \mathbb{C}^{p\times\lceil\frac{N}{r}\rceil}$, $\mathbf{X}_{S_i} \in \mathbb{C}^{L\times\lceil\frac{N}{r}\rceil}$ and $\mathbf{E}_{S_i}$ are the ith column-partitioned sub-matrices of $\mathbf{Y}$, $\mathbf{X}$ and $\mathbf{E}$, respectively. In doing so, each sub-MMV problem has lower sparsity, i.e., $|\text{supp}(\mathbf{X}_{S_i})| \le |\text{supp}(\mathbf{X})|$.
- **Row-extraction:** To further reduce $|\text{supp}(\mathbf{X}_{S_i})|$ in (5), we exploit the fact that $\text{supp}(\mathbf{X}_{S_i})$ usually occupies consecutive multiple rows. Specifically, we extract a partial matrix $\bar{\mathbf{X}} \in \mathbb{C}^{L\times N}$ from $\mathbf{X}$ by keeping only the boundary non-zeros in each PU signal $\mathbf{X}^{U_j}$, while setting other entries to zero. Clearly, once $\text{supp}(\bar{\mathbf{X}}_{S_i})$ is determined, $\text{supp}(\mathbf{X}_{S_i})$ is known immediately. Thus, identifying $\text{supp}(\mathbf{X}_{S_i})$ can be readily transferred to identifying $\text{supp}(\bar{\mathbf{X}}_{S_i})$. The latter has much lower joint sparsity.

As a result, in each sub-problem of bMMV, the actual sampling rate becomes easier to attain the theoretical lower bound compared to the original MMV problem in (4).

An illustrative example of bMMV is given in Fig. 2, in which there are 2 PU signals in $\mathbf{X}$ and each support $\text{supp}(\mathbf{X}_{S_i})$ occupies consecutive 3 or 4 rows. In this case, even if $\mathbf{X}$ is column-partitioned into 4 sub-matrices (i.e., $\mathbf{X}_{S_i}, \cdots, \mathbf{X}_{S_4}$), the joint sparsity $|\text{supp}(\mathbf{X}_{S_i})|$ of each sub-matrix does not decrease much compared to $|\text{supp}(\mathbf{X})|$. Nevertheless, row-extraction can further reduce the sparsity, effectively. Indeed, it can be seen from Fig. 2 that $|\text{supp}(\mathbf{X})| = 8$, $\max_{i\in\{1,\cdots,4\}}|\text{supp}(\mathbf{X}_{S_i})| = 8$, but $\max_{i\in\{1,\cdots,4\}}|\text{supp}(\bar{\mathbf{X}}_{S_i})| = 2$.

The column-partitioning strategy has also been studied in [16], where an approximate lower bound is obtained as $2N_{\text{sig}}f_s$ for $B \le f_s$ and $r \to \infty$. The following theorem whose proof is left to Supplementary gives a condition for actual sampling rate to attain the low bound in MCS.

Theorem 2. *When* $r \in [\lceil\frac{f_s}{\lfloor(D-1)f_s - B\rfloor}\rceil, \infty)$ *for* $D = \lfloor\frac{B}{f_s}\rfloor + 2$ *and* $B > f_s$*, we have* $\max_{i\in\{1,\cdots,r\}}\big|supp(\bar{\mathbf{X}}_{S_i})\big| \le \frac{N_{sig}B}{f_s}$.

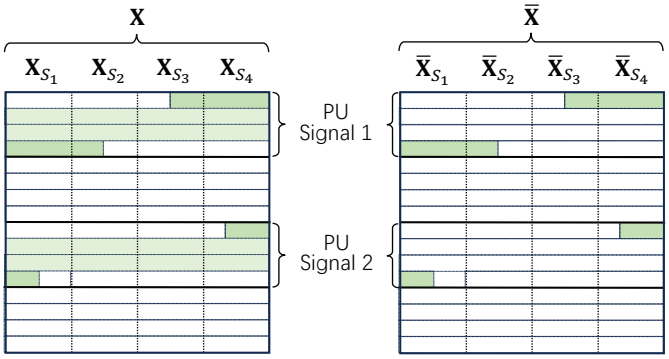

Fig. 2. An illustrative example of MCS signal $\mathbf{X}$ with 2 PU signals and $|\text{supp}(\mathbf{X})| = 8$. Via column-partitioning, $\mathbf{X}$ is decomposed into 4 sub-matrices $(\mathbf{X}_{S_1}, \cdots, \mathbf{X}_{S_4})$ with sparsity 6, 5, 5 and 6, respectively. $\bar{\mathbf{X}}$ keeps only the boundary nonzeros of PU signals in $\mathbf{X}$. The sparsity of its partitioned 4 sub-matrices $(\bar{\mathbf{X}}_{S_1}, \cdots, \bar{\mathbf{X}}_{S_4})$ are 2, 1, 1 and 2, respectively.

IV. The SI-SSP Algorithm

SI-SSP recovers the signal sub-matrices $\mathbf{X}_{S_1}, \cdots, \mathbf{X}_{S_r}$ by solving r sub-MMV problems individually.[1] The details of SI-SSP is given in Alg. 1. Initialized with a residual matrix $\mathbf{R}_{S_i}^0 = \mathbf{Y}_{S_i}$ and an estimated support $S_{S_i}^0 = \emptyset$ for $i = 1, \cdots, r$, it iteratively constructs the supports of $\mathbf{X}_{S_i}$'s through a merging-and-pruning strategy. This strategy mainly is inspired by [24]. In the kth iteration, the SI set for estimating $\text{supp}(\mathbf{X}_{S_i})$ is chosen as the union of the $r-1$ previously estimated supports of $\text{supp}(\mathbf{X}_{S_j})$'s, $j \in \{1, \cdots, r\} \setminus i$. Put formally,

$$\Lambda_{S_i}^k = \bigcup\nolimits_{j \in \{1,\cdots,r\} \setminus i} S_{S_j}^{k-1}. \tag{6}$$

Since the supports of adjacent sub-matrices are usually overlapping (see, e.g., Fig. 2), some elements in $\Lambda_{S_i}^k$ are highly likely to be contained in $\text{supp}(\mathbf{X}_{S_i})$. Hence, we wish to utilize $\Lambda_{S_i}^k$ as SI in identifying $\text{supp}(\mathbf{X}_{S_i})$.

To properly incorporate this SI, a diagonal weighting matrix $\mathbf{W}_{S_i}^k \in \mathbb{R}^{L \times L}$ is constructed with diagonal entries indexed by $\Lambda_{S_i}^k$ being $\omega \geq 0$, and zero otherwise. This matrix will be used in the subsequent selection and pruning operations, which, respectively, enhances the possibility of choosing and keeping the elements in $\Lambda_{S_i}^k$ (see Steps 4 and 7). The estimation operation in Steps 6 and 9 involves minimizing the Frobenius-norm representation distance to $\mathbf{Y}_{S_i}$, given that the signal to be estimated is supported on a merged (or pruned) set. The algorithm terminates until the residual matrix meets some tolerance, and outputs the estimated signal sub-matrices.

Next, we study the recovery performance of SI-SSP under the restricted isometry property (RIP) framework.

Definition 1. *A sensing matrix $\mathbf{A} \in \mathbb{R}^{m \times n}$ is said to satisfy the s-order RIP if for any s-sparse ($\|\mathbf{x}\|_0 \leq s$) signal $\mathbf{x} \in \mathbb{R}^n$*

$$(1-\delta)\|\mathbf{x}\|_2^2 \leq \|\mathbf{A}\mathbf{x}\|_2^2 \leq (1+\delta)\|\mathbf{x}\|_2^2, \tag{7}$$

where $0 \leq \delta < 1$. The infimum of δ, denoted by δ_s, is called the restricted isometry constant (RIC) of $\mathbf{A}$.

[1] As SI-SSP is designed to recover any signal sub-matrices from the column-partitioned MMV model, regardless of row-extraction, we use more general notations $\mathbf{X}_{S_i}$ and $\mathbf{Y}_{S_i}$ (rather than $\bar{\mathbf{X}}_{S_i}$ and $\bar{\mathbf{Y}}_{S_i}$) to describe our algorithm.

Algorithm 1: The SI-SSP Algorithm

Input : sensing matrix $\mathbf{A} \in \mathbb{C}^{L \times N}$, measurements $\mathbf{Y}_{S_i} \in \mathbb{C}^{p \times \lceil \frac{N}{r} \rceil}$, $i = 1, \cdots, r$, sparsity s and residual tolerance ϵ.

Initialize: iteration count $k = 0$, residual matrix $\mathbf{R}_{S_i}^0 = \mathbf{Y}_{S_i}$ and estimated support $S_{S_i}^0 = \emptyset$, $i = 1, \cdots, r$.

1 **while** $\|\mathbf{R}_{S_i}^k\|_F < \epsilon$ **do**
2 $k = k + 1$;
3 **Compute** SI set $\Lambda_{S_i}^k = \bigcup_{j \in \{1,\cdots,r\} \setminus i} S_{S_j}^{k-1}$ and weighting matrix $\mathbf{W}_{S_i}^k$ with diagonal entries supported on $\Lambda_{S_i}^k$ being ω and zero otherwise;
4 **Select** $\Delta S = \{s$ indices corresponding to the largest s row norms sum of $\mathbf{A}^H \mathbf{R}_{S_i}^{k-1}$ and $\mathbf{W}_{S_i}^k \mathbf{A}^H \mathbf{R}_{S_i}^{k-1}\}$;
5 **Merge** $\tilde{S}_{S_i}^k = S_{S_i}^{k-1} \cup \Delta S$;
6 **Estimate** $\tilde{\mathbf{X}}_{S_i}^k = \underset{\mathbf{\Theta}:\text{supp}(\mathbf{\Theta}) = \tilde{S}_{S_i}^k}{\arg\min} \|\mathbf{Y}_{S_i} - \mathbf{A}\mathbf{\Theta}\|_F$;
7 **Prune** $S_{S_i}^k = \{s$ indices corresponding to the largest s row norms sum of $\tilde{\mathbf{X}}_{S_i}^k$ and $\mathbf{W}_{S_i}^k \tilde{\mathbf{X}}_{S_i}^k\}$;
8 **Update** residual matrix $\mathbf{R}_{S_i}^k = \mathbf{Y}_{S_i} - \mathbf{A}\tilde{\mathbf{X}}_{S_i}^k$;
9 **Estimate** $\mathbf{X}_{S_i}^k = \underset{\mathbf{\Theta}:\text{supp}(\mathbf{\Theta}) = S_{S_i}^k}{\arg\min} \|\mathbf{Y}_{S_i} - \mathbf{A}\mathbf{\Theta}\|_F$;
10 **end**

Output : estimated signal submatrices $\mathbf{X}_{S_i}^k$'s.

The following theorem provides a condition of SI-SSP for stable recovery of sparse signals.

Theorem 3. *Consider the column-partitioned MMV model (5) with $\min_{i,j} \|(\mathbf{X}_{S_i})_{j,:}\|_2 / \|\mathbf{X}_{S_i}\|_F = \eta$ and $|\text{supp}(\mathbf{X}_{S_i})| \leq s$. Let $s_1 := \min_{i,k} |\Lambda_{S_i}^k \cap \text{supp}(\mathbf{X}_{S_i})|$, $s_2 := \min_{i,k} |\Lambda_{S_i}^k \cap \text{supp}(\mathbf{X}_{S_i}) \cap \tilde{S}_{S_i}^k|$ and $s_3 := \min_{i,k} |\Lambda_{S_i}^k \cap \text{supp}(\mathbf{X}_{S_i}) \cap \tilde{S}_{S_i}^k \setminus S_{S_i}^k|$. Then, if the sensing matrix $\mathbf{A}$ obeys the RIP with*

$$\delta_{3s} \leq \sqrt{\frac{\nu_1 \sqrt{\nu_1^2 + 4\nu_2^2} - \nu_1^2 - 1}{4\nu_1^2\nu_2^2 - 2\nu_1^2 - 1}} \tag{8}$$

where $\nu_1 := \frac{1+\omega}{1+\eta\omega\sqrt{s_2}}$ and $\nu_2 := \frac{1+\omega}{1+\eta\omega\sqrt{s_3}}$, SI-SSP produces an signal estimate $\mathbf{X}^k = [\mathbf{X}_{S_1}^k, \cdots, \mathbf{X}_{S_r}^k]$ satisfying

$$\left\|\mathbf{X} - \mathbf{X}^k\right\|_F \leq \rho^k \|\mathbf{X}\|_F + \tau \|\mathbf{E}\|_F, \tag{9}$$

where $\rho \in (0,1)$ and τ are constants depending on δ_{3s}, ν_1 and ν_2. Furthermore, after at most $k^* = \lceil \log_\rho \frac{\|\mathbf{X}\|_F}{\tau \|\mathbf{E}\|_F} \rceil$ iterations, SI-SSP estimates $\mathbf{X}$ with

$$\|\mathbf{X} - \mathbf{X}^{k^*}\|_F \leq (\tau + 1)\|\mathbf{E}\|_F. \tag{10}$$

The proof is mainly based on the analysis of Steps 4 and 7 in Alg. 1. The details are provided in Lemmas 1, 2 and 3, which are left to Supplementary [23] due to page limitation.

Sketch of proof: First, we consider one iteration and bound $\|\mathbf{X} - \mathbf{X}^k\|_F$ with $\|\mathbf{X} - \mathbf{X}^{k-1}\|_F$. To this end, let $\tau_1 = (\nu_1\sqrt{2(1-\delta_{3s})} + \sqrt{1+\delta_{3s}})(1-\delta_{3s})^{-1}$ and $\tau_2 = \sqrt{1+\delta_{3s}}$. Then, we obtain from Lemmas 2 and 3 that

$$\|\mathbf{X} - \tilde{\mathbf{X}}^k\|_F \leq \nu_1 \sqrt{\frac{2\delta_{3s}^2}{1-\delta_{3s}^2}} \|\mathbf{X} - \mathbf{X}^{k-1}\|_F + \tau_1 \|\mathbf{E}\|_F, \tag{11}$$

$$\|\mathbf{X} - \mathbf{X}^k\|_F \leq \sqrt{\frac{1}{1-\delta_{3s}^2}} \left\|\mathbf{X}_{(S^k)^c}\right\|_F^2 + \frac{\sqrt{1+\delta_{3s}}}{1-\delta_{3s}} \|\mathbf{E}\|_F. \tag{12}$$

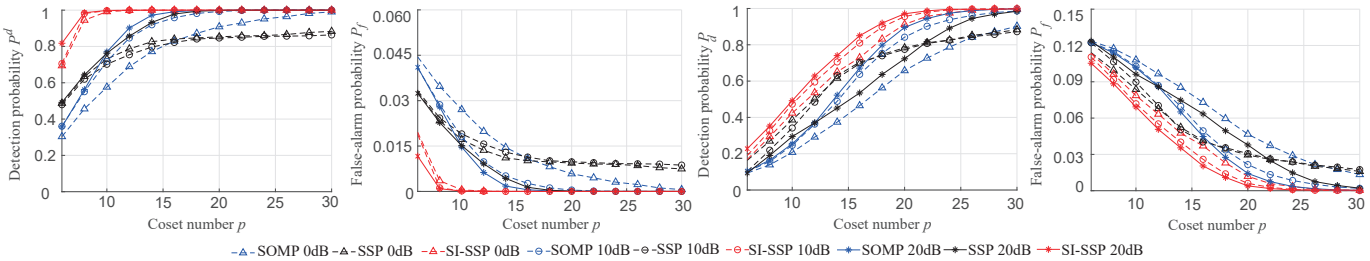

Fig. 3. P_d and P_f of SOMP, SSP and SI-SSP against p from $6 \sim 30$ with SNR = $\{0, 10, 20\}$ dB. From left to right: the signal with $\frac{L}{M} = 2$ is decomposed into $r = 6$ sub-matrices and the signal with $\frac{L}{M} = 4$ is decomposed into $r = 6$ sub-matrices.

Dividing $(S^k)^c$ into 2 disjoint subsets: $(\tilde{S}^k)^c$ and S_∇, we get

$$\begin{aligned}
\left\|\mathbf{X}_{(S^k)^c}\right\|_F^2 &= \|\mathbf{X}_{S_\nabla}\|_F^2 + \|\mathbf{X}_{(\tilde{S}^k)^c}\|_F^2 \\
&\overset{\text{Lemmas 2, 3}}{\leq} 2\left(\nu_2\delta_{3s}\|\mathbf{X}-\tilde{\mathbf{X}}^k\|_F + \nu_2\tau_2\|\mathbf{E}\|_F\right)^2 \\
&\quad + 2\left(\nu_1\delta_{3s}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_2 + \nu_1\tau_2\left\|\mathbf{E}\right\|_F\right)^2 \\
&\overset{(11)}{\leq} 2\left(\sqrt{\frac{2\nu_1^2\nu_2^2\delta_{3s}^4}{1-\delta_{3s}^2}}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F + \nu_2(\tau_1\delta_{3s}+\tau_2)\right. \\
&\quad \left.\times\|\mathbf{E}\|_F\right)^2 + 2\left(\nu_1\delta_{3s}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F + \nu_1\tau_2\left\|\mathbf{E}\right\|_F\right)^2 \\
&\overset{\text{Lemma 1}}{\leq} 2\left(\sqrt{\frac{2\nu_1^2\nu_2^2\delta_{3s}^4}{1-\delta_{3s}^2} + \nu_1^2\delta_{3s}^2}\left\|\mathbf{X}-\mathbf{X}^{k-1}\right\|_F\right. \\
&\quad \left. + \left((\nu_1+\nu_2)\tau_2 + \nu_2\delta_{3s}\tau_1\right)\left\|\mathbf{E}\right\|_F\right)^2. \qquad (13)
\end{aligned}$$

Combining (13) and (12) yields

$$\|\mathbf{X}-\mathbf{X}^k\|_F \leq \rho\|\mathbf{X}-\mathbf{X}^{k-1}\|_F + (1-\rho)\tau\|\mathbf{E}\|_F \qquad (14)$$

where $\rho := \sqrt{2}\delta_{3s}\sqrt{2\nu_1^2\nu_2^2\delta_{3s}^2 + \nu_1^2 - \nu_1^2\delta_{3s}^2}(1-\delta_{3s}^2)^{-1}$ and $\tau := \sqrt{2}\delta_{3s}\nu_2(\nu_1\sqrt{2(1-\delta_{3s})} + \sqrt{1+\delta_{3s}})(1-\delta_{3s}^2)^{-1/2}(1-\delta_{3s})^{-1}(1-\rho)^{-1} + (\nu_1\nu_2\sqrt{2(1-\delta_{3s})} + \sqrt{1+\delta_{3s}})(1-\delta_{3s})^{-1}$.

Second, we recursively apply (14) to obtain

$$\|\mathbf{X}-\mathbf{X}^k\|_F \leq \rho^k\|\mathbf{X}\|_F + \tau\|\mathbf{E}\|_F \qquad (15)$$

where $\rho < 1$ under (8). When $k^* = \lceil \log_\rho \frac{\|\mathbf{X}\|_F}{\tau\|\mathbf{E}\|_F} \rceil$, we have $\rho^k\|\mathbf{X}\|_F \leq \tau\|\mathbf{E}\|_F$, and thus the stability result (10).

Remark 1. *The RIP condition of SI-SSP depends on constants ν_1 and ν_2. When the weighting parameter $\omega = 0$ (i.e., SI is not used), we have $\nu_1 = \nu_2 = 1$. In this case, SI-SSP reduces to SSP, whose RIP condition becomes $\delta_{3s} < \sqrt{\sqrt{5}-2} \approx 0.4859$. When $\omega > 0$ and s_2, s_3 increase (i.e., more SI is available), ν_1 and ν_2 can generally be less than 1, for example, when $\nu_1 = 0.7$ and $\nu_2 = 0.8$, we have $\delta_{3s} < 0.6$, which is less restrictive than $\delta_{3s} < 0.4859$, indicating that SI-SSP performs better due to utilization of the SI.*

V. NUMERICAL EXPERIMENTS

We perform numerical experiments to test the performance of SI-SSP. In our experiments, we consider a multi-band signal with frequency range $[-5, 5]$GHz and $N_{\text{sig}} = 3$ PU's signals, whose bandwidths are $B = 100$ and 300MHz corresponding to $\frac{L}{M} = 2$ and 4, respectively. We divide the frequency range into $L = 100$ cosets and use $p \in [6, 30]$ channels to get the MCS signal, with delay coefficients randomly generated. Moreover, we add Gaussian white noise to the signal with signal-to-noise ratio (SNR) of $\{0, 10, 20\}$dB. In bMMV, we consider $r = 6$ partitioning sub-matrices and set $\omega = 0.5$ to the weighting matrix $\mathbf{W}$. To evaluate the recovery accuracy of SI-SSP, we use the detection probability P_d and false alarm probability P_f as metrics. For comparative purpose, we consider SOMP [19] and simultaneous subspace pursuit (SSP, i.e., SP for MMV).

In Fig. 3, we plot the performance curves of SI-SSP at $\frac{L}{M} = 2$ and $\frac{L}{M} = 4$. For the case of $\frac{L}{M} = 2$, the left two figures show that when $p < 2|\text{supp}(\mathbf{X})| = 12$ (i.e., lower bound of sampling rate), P_d of SI-SSP reaches 1 and is higher than SSP and SOMP, while the P_f reaches 0. For SSP and SOMP, the convergence speed is slow at 0db, indicating that its noise resistance ability is inferior to SI-SSP. In the case of $\frac{L}{M} = 4$, SI-SSP achieves a P_d close to 1 and a P_f close to 0 for $p = 2|\text{supp}(\mathbf{X})| = 24$, which significantly outperforms SSP and SOMP. These experimental results clearly demonstrate superiority of our algorithm.

VI. CONCLUSION

In this paper, we have proposed a bMMV model by exploiting the structure prior of MCS signals. Our model reduces the joint sparsity significantly and attains theoretical lower bound of sampling rate. We have also proposed an algorithm called SI-SSP for recovering MCS signals from the bMMV model, which enhances the recovery accuracy by utilizing properly designed side information. Experimental results show that SI-SSP has improved detection and false-alarm probability compared to existing methods.

The present work opens up some promising future directions. Firstly, we are interested in recovering more general multi-band signals with PU signals of different bandwidths. Secondly, we wish to refine the side information in SI-SSP by, e.g., focusing on fewer signal neighbor sub-matrices, rather than all other sub-matrices. Thirdly, it would be interesting to apply bMMV and SI-SSP to more sub-Nyquist sampling structures, such as the modulated wideband converter and random demodulater.

REFERENCES

[1] M. Marcus, J. Burtle, B. Franca, A. Lahjouji, and N. McNeil, "Federal communications commission spectrum policy task force," *Report of the unlicensed devices and experimental licenses working group*, no. 02-135, 2002.

[2] B. Wang and K. R. Liu, "Advances in cognitive radio networks: A survey," *IEEE Journal of selected topics in signal processing*, vol. 5, no. 1, pp. 5–23, 2010.

[3] M. Jia, Z. Gao, Q. Guo, Y. Lin, and X. Gu, "Sparse feature learning for correlation filter tracking toward 5g-enabled tactile internet," *IEEE Transactions on Industrial Informatics*, vol. 16, no. 3, pp. 1904–1913, 2019.

[4] M. Wang, Y. Lin, Q. Tian, and G. Si, "Transfer learning promotes 6g wireless communications: Recent advances and future challenges," *IEEE Transactions on Reliability*, vol. 70, no. 2, pp. 790–807, 2021.

[5] H. Sun, W.-Y. Chiu, J. Jiang, A. Nallanathan, and H. V. Poor, "Wideband spectrum sensing with sub-nyquist sampling in cognitive radios," *IEEE Transactions on Signal Processing*, vol. 60, no. 11, pp. 6068–6073, 2012.

[6] S. Ren, Z. Zeng, C. Guo, and X. Sun, "A low complexity sensing algorithm for wideband sparse spectra," *IEEE Communications Letters*, vol. 21, no. 1, pp. 92–95, 2017.

[7] B. Khalfi, B. Hamdaoui, M. Guizani, and N. Zorba, "Efficient spectrum availability information recovery for wideband dsa networks: A weighted compressive sampling approach," *IEEE Transactions on Wireless Communications*, vol. 17, no. 4, pp. 2162–2172, 2018.

[8] L. Yang, J. Fang, H. Duan, and H. Li, "Fast compressed power spectrum estimation: Toward a practical solution for wideband spectrum sensing," *IEEE Transactions on Wireless Communications*, vol. 19, no. 1, pp. 520–532, 2020.

[9] D. D. Ariananda and G. Leus, "Compressive wideband power spectrum estimation," *IEEE Transactions on Signal Processing*, vol. 60, no. 9, pp. 4775–4789, 2012.

[10] Mishali, Moshe and Eldar, Yonina C, "From theory to practice: Sub-nyquist sampling of sparse wideband analog signals," *IEEE Journal of selected topics in signal processing*, vol. 4, no. 2, pp. 375–391, 2010.

[11] A. Lopez-Parrado and J. Velasco-Medina, "Cooperative wideband spectrum sensing based on sub-nyquist sparse fast fourier transform," *IEEE Transactions on Circuits and Systems II: Express Briefs*, vol. 63, no. 1, pp. 39–43, 2015.

[12] D. L. Donoho, "Compressed sensing," *IEEE Transactions on information theory*, vol. 52, no. 4, pp. 1289–1306, 2006.

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

[14] R. Venkataramani and Y. Bresler, "Optimal sub-nyquist nonuniform sampling and reconstruction for multiband signals," *IEEE Transactions on Signal Processing*, vol. 49, no. 10, pp. 2301–2313, 2001.

[15] T. Moon, H. W. Choi, N. Tzou, and A. Chatterjee, "Wideband sparse signal acquisition with dual-rate time-interleaved undersampling hardware and multicoset signal reconstruction algorithms," *IEEE Transactions on Signal Processing*, vol. 63, no. 24, pp. 6486–6497, 2015.

[16] Z. Song, J. Yang, H. Zhang, and Y. Gao, "Approaching sub-nyquist boundary: Optimized compressed spectrum sensing based on multicoset sampler for multiband signal," *IEEE Transactions on Signal Processing*, vol. 70, pp. 4225–4238, 2022.

[17] M. Mishali and Y. C. Eldar, "Blind multiband signal reconstruction: Compressed sensing for analog signals," *IEEE Transactions on signal processing*, vol. 57, no. 3, pp. 993–1009, 2009.

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

[19] 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.

[20] J. A. Tropp, "Algorithms for simultaneous sparse approximation. part ii: Convex relaxation," *Signal Processing*, vol. 86, no. 3, pp. 589–602, 2006.

[21] J. M. Kim, O. K. Lee, and J. C. Ye, "Compressive music: Revisiting the link between compressive sensing and array signal processing," *IEEE Transactions on Information Theory*, vol. 58, no. 1, pp. 278–301, 2012.

[22] K. Lee, Y. Bresler, and M. Junge, "Subspace methods for joint sparse recovery," *IEEE Transactions on Information Theory*, vol. 58, no. 6, pp. 3613–3641, 2012.

[23] J. W. D. Xiao and Y. Lin, "Boundary multiple measurement vectors for multi-coset sampler–supplementary," Online avaliable: https://github.com/cherry5-lang/Supplementary, 2023.

[24] W. Dai and O. Milenkovic, "Subspace pursuit for compressive sensing signal reconstruction," *IEEE Transactions on Information Theory*, vol. 55, no. 5, 2009.