Downlink Synchronization Techniques for Heterogeneous Cellular Networks

Neelakandan Rajamohan∗, Arun Pachai Kannu†

Department of Electrical Engineering
Indian Institute of Technology Madras
Chennai - 600036, India

Email: ∗ee10d011@ee.iitm.ac.in, †arunpachai@ee.iitm.ac.in

Abstract

Base stations (BS) in a cellular network send unique synchronization signals based on their physical layer cell-identity, in every frame. In order to synchronize with a BS, a mobile terminal (MT) needs to identify the following synchronization parameters - starting time of the frame, relative frequency offset and the cell-identity of the corresponding BS. In a heterogeneous cellular network, due to the increased density of low-powered base stations, the MT receives signals from multiple BS in its vicinity. We consider the downlink synchronization problem where the MT tries to recover the synchronization parameters of all the BS in its neighborhood, using which, MT can subsequently choose the suitable BS for its connection. In any given scenario, the number of BS in the neighborhood of MT is relatively small compared to the total number of BS in the entire network. Exploiting this sparseness, we present a two-stage synchronization approach where the first stage identifies the frame timing using an approximation to the maximum likelihood detector and the second stage recovers the cell identities

This work is sponsored by Dept. of Science and Technology, India.
A patent has been filed based on this work.
and the corresponding frequency offsets of all the neighborhood BS using block sparse signal recovery framework of compressive sensing. We analytically and numerically study the recovery guarantees of our proposed techniques and establish their superior performance over existing successive interference cancellation and matched filtering approaches.

Keywords— timing detection, frequency offset estimation, cell search, compressive sensing, block sparse signal recovery

I. INTRODUCTION

A cellular network comprises of base stations (BS) serving users in their respective coverage areas. Each BS in a cellular network sends its data in a sequence of frames, with each frame having a synchronization signal unique to the physical layer cell-identity of the corresponding BS. Before a mobile terminal (MT) can exchange information with a BS, MT needs to acquire the synchronization parameters of that BS which includes finding the starting time of frames from that BS, estimation of relative frequency offset and recovering the physical layer identity of that BS. This procedure of recovering the synchronization parameters by the MT in a cellular network is referred as downlink synchronization or cell search and has been studied previously in [1]–[5].

In order to address the growing demand for data communication among mobile phone users, a heterogeneous cellular network (HetNet) with many low powered BS within the coverage area of conventional BS is being envisioned. In a HetNet, there are macro cells (conventional BS with coverage radius of few kilometers and transmit power levels of 46 dBm), pico cells (operator deployed BS with coverage radius of few hundreds of meters and power levels of 20 to 23 dBm) and femto cells (user deployed BS to provide coverage inside home with open or restricted access) [6], [7].

Due to the increased density of BS in a HetNet, the MT receives superposition of signals from multiple BS in its vicinity. In a HetNet, MT may need to connect to a BS which does not
correspond to the strongest received power [8]. For instance, a MT has to connect to a (far away) macro BS while receiving the strongest signal from a nearby (restricted access) femto BS. Also, a MT may need to connect to a (low-powered) pico BS while receiving the strongest signal from a (high-powered) macro BS (in order to permit cell splitting to enhance overall throughput [9]). In view of this issue, we consider the HetNet synchronization problem, in which the MT tries to acquire the synchronization parameters of all the BS in its neighborhood. The notion of neighbor BS encompasses the base stations with received power sufficiently higher than the noise floor at the MT to enable further communication. Further, acquiring the neighbor BS parameters are essential in several advanced techniques including positioning of the MT [10], inter cell interference coordination [11] and coordinated multi-point transmission [12] .

In any given scenario, the number of BS in the neighborhood of MT is relatively small compared to the total number of BS in the entire network. Exploiting this sparseness, we present a two-stage synchronization approach where the first stage identifies the frame timing using an approximation to the maximum likelihood detector and the second stage recovers the cell identities and the corresponding frequency offsets of all the neighborhood BS using block sparse signal recovery framework of compressive sensing (CS) [13], [14]. We analytically and numerically study the recovery guarantees of our proposed techniques and establish their superior performance over existing successive interference cancellation [15] and matched filtering approaches [5], [16].

While the closely related works in [15], [17] consider perfect frame timing knowledge and absence of relative frequency offsets, we consider the initial acquisition scenario where frame timing information is not available and the residual frequency offsets are present. Due to the presence of frequency offsets, the recovery properties of sensing matrix in our corresponding CS model differs significantly from those obtained in [17]. We also study several new block sparse recovery algorithms for our CS model, in addition to the ones considered in [17].

Notation: matrices/vectors are denoted by bold uppercase/lowercase letters, $l_p$ norm by $\| \cdot \|_p$, ...
inner product by $\langle \cdot \rangle$, determinant by $\det(\cdot)$, Frobenius norm by $\| \cdot \|_F$, transpose by $(\cdot)^t$, hermitian by $(\cdot)^*$, sets by mathcal font $\mathcal{A}$, cardinality of the set by $|\mathcal{A}|$, and set minus operation by $\mathcal{A} \setminus \mathcal{B}$ and $\delta(\cdot)$ denotes Kronecker delta function.

II. SYSTEM MODEL AND PROBLEM STATEMENT

We consider the downlink scenario where the mobile terminal receives signals transmitted by the base stations in its vicinity. Let $\mathcal{I} = \{1, \cdots, C\}$ denotes the set of cell identities available in a cellular system, where $C$ denotes the total number of unique cell identities in the network. For example, in 3GPP-LTE system, the value of $C$ is 504. Let $\mathcal{B} \subset \mathcal{I}$ denotes the set of BS in the vicinity of MT from which it receives signals. We call the set $\mathcal{B}$ as the neighbor set of MT and $|\mathcal{B}| = S$ denotes the number of neighbor base stations. Without loss of generality, we assume that the cell identities of all neighbor BS are distinct. In the discrete time base band model, the received observation sequence $\{y(n)\}$ in a given MT is

$$y(n) = \sum_{q \in \mathcal{B}} \sum_{\ell=0}^{L_q-1} \sqrt{P_q} h_q(n; \ell) t_q(n - \ell) e^{j2\pi\epsilon_q n} + w(n), \quad (1)$$

where $w(n)$ denotes the additive noise, $\{t_q(n)\}$ denotes the transmit signal with power level $P_q$ from BS with identity $q$, $\epsilon_q$ denotes the normalized relative frequency offset and, $\{h_q(n; \ell)\}$ and $L_q$ denote the channel impulse response (CIR) at time $n$ between MT and BS with identity $q$ and the corresponding delay spread respectively. The frequency offset $\epsilon_q$ lies in the interval $[-\epsilon_{\text{max}}, \epsilon_{\text{max}}]$ with $\epsilon_{\text{max}}$ being the maximum possible frequency offset, which depends on the oscillator specifications in the cellular standards. The channel delay spread $L_q \leq L_{\text{max}}, \forall q \in \mathcal{B}$, where the maximum delay spread $L_{\text{max}}$ depends on the transmit power levels of the macro cell. The number of neighbor base stations in the vicinity of the mobile terminal is bounded as $S \leq S_{\text{max}}$ where $S_{\text{max}}$ depends on the density of low powered BS.

Base stations in a cellular system transmit information in a sequence of frames whose structure is specified in the cellular standards. In every frame (of length $F$), at a predetermined position
(in time), each base station transmits a training sequence (of length $N$) which is specific to its identity. This is illustrated in Fig. 1. Focusing on a specific frame, let \( \{x_q(n), n = 0, \cdots, N-1\} \) denotes the known/fixed training/synchronization signal (SS) of BS with identity $q$ and $T_q$ represents the starting location of the training signal received from BS with identity $q$. Hence, $t_q(n+T_q) = x_q(n), \, n \in \{0, \cdots, N-1\}$. We assume that the starting time of the frames from all the neighbor BS in the set $B$ are time aligned (within a degree of accuracy accounting for the propagation delays). Frame time alignment is a feasible assumption and can be accomplished if the small cells within a macro coverage area align their frames with that of their macro BS, by monitoring the macro timing using synchronization signal transmitted by the macro BS. This frame time alignment is typical in the study of HetNets [7], [18]. With the transmission of the synchronization signals from the $B$ being time aligned, we can assume that there is a time window $[T, T + L_{\text{max}}]$ such that $[T_q, T_q + L_q] \subset [T, T + L_{\text{max}}]$ for all $q \in B$ (as shown in the figure Fig. 1). This assumption is made in the neighbor cell search problem addressed in [15] as well. We consider cyclic prefixed OFDM transmission (as in LTE) and assume that the cyclic prefix duration is larger than $L_{\text{max}}$. 

Fig. 1. Alignment of SS from different BS in the received frame. (Not to scale)
We are interested in developing the downlink synchronization algorithms for the mobile terminal in our HetNet model (1). Synchronizing with a BS involves finding the starting time of its frame, its cell identity and its relative frequency offset. However, as previously discussed, synchronizing with the strongest BS may not be sufficient in HetNets. So, we are interested in recovering the identities of all the neighbors in \( B \), along with their residual frequency offsets \( \{\epsilon_q\} \) and their frame timing. Note that, once the MT recovers the synchronization parameters of all its neighbor BS, it can subsequently connect to a suitable BS for exchanging data/information.

### III. Frame Timing Detection

Since frames are sent in a continuous stream, and since each frame has a synchronization signal at a fixed location, it follows that, any contiguous collection of a frame worth \( (F) \) of received samples \( \{y(n), n = 0, \cdots, F - 1\} \) will have a synchronization signal present somewhere in that collection. By identifying the location of the training signal in the collected samples, the location of the start of the frame can be inferred. To proceed further, we assume that channel coherence time is larger than the duration of the training signal \( N \) as done in [15], which is satisfied in typical mobility environments. For convenience, let us denote \( h_q(\ell) = h_q(n; \ell) \) for \( n \in [T, \cdots, T + N - 1] \). Let us define the vectors \( y_k = [y(k), \cdots, y(k + N - 1)]^t \) for \( k \in \{0, \cdots, F - 1\} \). The observation vector \( y_T := y = [y(T), \cdots, y(T + N - 1)]^t \) is the received vector\(^1\) corresponding to the superposition of synchronization signals from the neighbor BS in \( B \). For convenience, we use the common length \( L = L_{\text{max}} \) for the CIR from all the base stations, by suitably padding zeros. Specifically, for \( \ell = 0, \cdots, L - 1 \), we define \( z_q(\ell) = \sqrt{P_q}h_q(\ell)e^{j2\pi \epsilon_q(T + \ell)}U[T - T_q + \ell] \) with \( U[n] \) being the unit step sequence (\( U[n] = 1 \) for \( n \geq 0 \) and 0 else) with the notion that \( h_q(\ell) = 0, \forall \ell \geq L_q \). We vertically concatenate \( z_q(\ell) \) to form \( L \times 1 \) vector \( z_q \), which corresponds to the CIR vector for BS with identity \( q \). Defining the matrix

\(^1\)For the value of \( T \) such that \( T_q \in [T, T + L_{\text{max}}] \forall q \in B \), the vector \( y \) captures SS from the all the BS in the set \( B \).
$X_{q,\epsilon} \in \mathbb{C}^{N \times L}$ with its $\ell^{th}$ column being $\{x_{q}(n-\ell)e^{j2\pi\epsilon(n-\ell)}, n = 0 \text{ to } N - 1, \ell = 0 \text{ to } L - 1\}$, we have

$$y = \sum_{q \in B} X_{q,\epsilon} z_{q} + w,$$

(2)

where $w = [w(T), \cdots, w(T + N - 1)]^T$ is the noise vector. The relative delays between the arrival times of the SS from different BS are taken into account in the construction of vectors $z_{q}$. Since the time window $[T, T + N - 1]$ captures the training sequences from all the BS at the MT (including all the multipaths), we can identify the common frame timing of all the BS in the set $B$ by recovering the value of $T$.

If we assume additive white Gaussian noise (AWGN) with variance $\sigma_w^2$, $z_{q}$ is zero-mean Gaussian with covariance $C_q$ and CIRs of different BS are uncorrelated, it follows that $y$ is zero-mean Gaussian with covariance

$$C_{B,\epsilon_q} = \sum_{q \in B} X_{q,\epsilon} C_q X_{q,\epsilon}^* + \sigma_w^2 I,$$

(3)

given the neighbor set $B$ and frequency offsets $\epsilon_q$. Given $k = T$, the neighbor set is $\hat{B}$ and the corresponding frequency offsets are $\hat{\epsilon}_q$, the conditional pdf of $y_k$ is zero mean Gaussian with covariance $C_{\hat{B},\hat{\epsilon}_q}$ and to find the maximum likelihood (ML) estimates, we need to maximize this conditional pdf over $k$, $\hat{B}$ and $\hat{\epsilon}_q$. Since $\epsilon_q$ is a continuous parameter, we divide the interval $[-\epsilon_{\max}, \epsilon_{\max}]$ into $M$ frequency bins (we consider $M$ to be odd for convenience) and perform a discrete search as proposed in [20]. Specifically, defining the set corresponding to the center frequencies of bins as $\mathcal{E} = \{\epsilon^1, \cdots, \epsilon^M\}$ with $\epsilon^i = \epsilon_{\max}\left(\frac{2i-1}{M} - 1\right)$, we jointly estimate the frame timing ($\hat{T}$), set of neighbor BS ($\hat{B}$) and the corresponding frequency offsets ($\hat{\epsilon}_q$) as

$$(\hat{T}, \hat{B}, \hat{\epsilon}_q)_{ML} = \arg\max_{k \in \{0, \cdots, F-1\}, \hat{B} \subset \mathcal{I} \text{ such that } |\hat{B}| \leq S_{\max} \forall q \in \hat{B} \text{ and } \hat{\epsilon}_q \in \mathcal{E}} \frac{\exp\left(-y_k^* C_{\hat{B},\hat{\epsilon}_q}^{-1} y_k\right)}{\det(C_{\hat{B},\hat{\epsilon}_q})},$$

(4)

Due to the presence of cyclic prefix, the positioning of this window has some flexibility (up to the difference between cyclic prefix duration and $L_{\max}$) [19].
where \( \exp(\cdot) \) denotes the exponential function. This optimization process is computationally expensive, since it needs to compute the determinant and inverse of \( C_{\tilde{B}q} \) for all \( k \in \{0, \cdots, F-1\} \), for all possible sets of neighbor BS (\( \tilde{B} \)) and their corresponding frequency offset hypotheses from \( E \). Also, it requires the knowledge of noise variance \( \sigma_w^2 \), transmit power levels \( P_q \) and the covariance of CIR for all BS in \( B \), which may not be typically available during the initial cell search. Due to these reasons, the above ML search procedure is not feasible.

We propose a simplified metric, referred as joint correlator detector (JCD), which identifies the timing as,

\[
(\hat{T}, \hat{\tilde{B}}, \hat{\epsilon}_q)_{\text{JCD}} = \arg \max_{k \in \{0, \cdots, F-1\}, \epsilon_i \in E} \sum_{q \in \tilde{B}} \| X_{q,\epsilon}^* y_k \|_2^2
\]

(5)

where \( \tilde{B} = \{ q : \| X_{q,\epsilon}^* y_k \|_2^2 \geq \tau_{\text{jcd}} \} \subset I \) and \( |\tilde{B}| \leq S_{\text{max}} \). Above JCD expression is equivalent to computing the correlation metric \( \| X_{q,\epsilon}^* y_k \|_2^2 \) for all \( \{k, i, q\} \), and summing the top (at-most) \( S_{\text{max}} \) metrics (provided they cross threshold \( \tau_{\text{jcd}} \)) for each timing hypothesis \( k \), and finding where the maximum sum occurs. The threshold is incorporated since the MT does not know the exact number of neighbor BS \( S \) and the choice of threshold gives a trade off between the successful detection (of identities in \( B \)) versus false alarms (wrongly detecting identities \( \notin B \)).

It can be easily verified that JCD becomes identical to ML if threshold \( \tau_{\text{jcd}} = 0 \) and all the following conditions hold: (a) There is no frequency offset, i.e., \( \epsilon_q = 0, \forall q \in \tilde{B} \). (b) Covariance matrix \( C_q = \rho I, \forall q \in \tilde{B} \) for some constant \( \rho \). (c) The training sequences have ideal correlation properties, i.e., \( \sum_{n=0}^{N-1} x_i(n) x_m^*(n+\ell) = \delta(i-m)\delta(\ell) \) for \( i, m \in I \) and \( \ell \in \{0, \cdots, L-1\} \). (d) \( \| y_k \|_2 \) is constant for all \( k \).

Though JCD is identical to ML only under some idealistic assumptions, it is computationally far less intensive (no matrix inversions) and does not require any apriori information on signal and noise powers. This makes JCD attractive for practical implementations. Since all the neighbor BS share the same frame timing (within a window of duration \( L_{\text{max}} \)), JCD works well in identifying their common frame timing. But the performance of JCD in recovering the identities of all
neighbor BS and their frequency offsets is shown to be poor (analytically and numerically) in the subsequent sections. Hence, in the following section, we develop new techniques for recovering the cell identities in the neighbor set $B$ and their corresponding frequency offsets.

IV. NEIGHBOR SET IDENTIFICATION

A. Observation Model in CS Framework

In this section, we present how the parameters of the neighbor base stations (cell identities and frequency offset values) can be recovered using the compressive sensing (CS) framework [21], [22], after their common frame timing is found. For simplicity in notations, we present the details of our approach assuming that the detected timing $\hat{T}$ from (5) is correct such that $[T_q, T_q + L_q] \subset [\hat{T}, \hat{T} + L_{\text{max}}]$, $\forall q \in B$. Based on the timing detected from JCD, we can collect the synchronization signal observations as in (2) and subsequently recover the neighbor set $B$ along with their frequency offsets. From (2), we see that only $S$ out of a total of $C$ base stations contribute to the observation vector $y$. Typically $S$ is small compared to the value of $C$ since the number of base stations in the vicinity of the MT is very small compared to the total number of base stations in the entire cellular network. This sparseness of the neighbor BS facilitate the application of compressive sensing based algorithms for the recovery of synchronization parameters of neighbor set $B$. The exact details of our approach are explained below.

Since frequency offset of each BS ($\epsilon_q$) in (2) is also unknown, we define the bin frequency in $E$ which is closest to the $\epsilon_q$ as,

$$\bar{\epsilon}_q = \min_{\epsilon \in E} |\epsilon - \epsilon_q|, \quad \forall q \in B.$$  \hspace{1cm} (6)

To proceed further, we approximate $X_{q,\epsilon_q}$ using $X_{q,\bar{\epsilon}_q}$. Specifically, we rewrite (2) as

$$y = \sum_{q \in B} X_{q,\bar{\epsilon}_q} z_q + \tilde{w},$$  \hspace{1cm} (7)

where the effective noise $\tilde{w} = w + \sum_{q \in B} (X_{q,\epsilon_q} - X_{q,\bar{\epsilon}_q})z_q$, is the sum of additive noise and the approximation error. Note that, as the number of bins $M$ increases, the difference between
$\bar{\epsilon}_q$ and $\epsilon_q$ decreases, leading to the reduction in the approximation error. Now, let us construct the super-block training matrix $X$ of size $N \times LMC$ by horizontally concatenating as $X = \begin{bmatrix} X_{1,\epsilon^1}, \cdots, X_{1,\epsilon^M}, \cdots, X_{C,\epsilon^1}, \cdots, X_{C,\epsilon^M} \end{bmatrix}$. Also, defining $Z_q^i = \begin{cases} z_q & q \in B \& \epsilon^i = \bar{\epsilon}_q \\ 0^{L \times 1} & \text{else} \end{cases}$, we define the super-block channel vector $Z$ of size $LMC \times 1$ by vertically concatenating\{ $z_1^1, \cdots, z_M^1, \cdots, z_1^C, \cdots, z_M^C$ \}. Now, (7) can be rewritten as,

$$y = X Z + \tilde{w}. \quad (8)$$

From its construction, it is clear that the only $S$ out of $CM$ blocks $z_q^i$ in $Z$ are non-zero (corresponding to $q \in B \& \epsilon^i = \bar{\epsilon}_q$). Hence $Z$ is $S$-block sparse signal [14], [17], [23]. The sensing matrix $X$ is a block matrix with concatenation of $CM$ toeplitz blocks $\{X_{q,\epsilon^i}\}$ and only $S$ of them are active in producing $y$, due to the block sparse structure of $Z$. Note that recovering the support of $Z$ (i.e., location of non-zero blocks) will lead to identification of the neighbor BS $B$ and the closest bin frequencies $\bar{\epsilon}_q$ given in (6). Hence, we focus on block-sparse signal recovery algorithms to recover $Z$ from $y$ and establish conditions for successful support recovery.

Block restricted isometry property [14], [22], [24], mutual subspace incoherence [25] and mutual incoherence [26] of the sensing matrix $X$ play a crucial role on the recovery guarantees of our algorithms. First, we characterize these properties for our sensing matrix with randomly generated training sequences $\{x_q(n)\}$, which are constant modulus complex sequences with uniform phase (CMCSUP). The matrix $X$ is said to satisfy block restricted isometry property (RIP) of order $S$ with a RIP constant $\delta \in [0, 1)$, if

$$(1 - \delta)\|Z\|_2^2 \leq \|X Z\|_2^2 \leq (1 + \delta)\|Z\|_2^2$$

for all $S$-block sparse vectors $Z$. Mutual subspace incoherence ($\mu$) of $X$ characterizes the smallest angle between column spaces of any two blocks $X_{p,\epsilon^i}$ and $X_{q,\epsilon^m}$, whereas, mutual incoherence ($\gamma$) of $X$ characterizes the smallest angle between any two columns of $X$. These can be defined
as,

\[
\mu = \max_{p,q \in I \text{ }& \text{ } \epsilon_i, \epsilon_m \in E \text{ }\& \text{ } p \neq q \text{ OR } i \neq m} \max \frac{|\langle a_p^i, a_m^m \rangle|}{\|a_p^i\|_2 \|a_m^m\|_2} \tag{9}
\]

where \( a_p^i \) and \( a_m^m \) belong to column spaces of \( X_{p,\epsilon_i} \) and \( X_{q,\epsilon_m} \) respectively, and,

\[
\gamma = \max_{i \neq m} \frac{|\langle a_i, a_m \rangle|}{\|a_i\|_2 \|a_m\|_2} \tag{10}
\]

where \( a_i, a_m \) are any two columns of \( X \). These properties for our sensing matrix \( X \) in (8) are characterized as follows.

**Theorem 1.** For the randomly generated training sequences \( x_q(n) = \frac{1}{\sqrt{N}} e^{j\theta(q,n)} \), with \( \theta(q,n) \) being i.i.d. uniformly distributed in \([-\pi, \pi]\), if the length of the sequence \( N \) satisfies

\[
N \geq \frac{12(SL)^2 \log \left( \frac{48L^2CM}{\beta} \right)}{(\delta - (M-1)\alpha)^2} \tag{11}
\]

then \( X \) satisfies block RIP of the order \( S \) with RIP constant \( \delta \in ((M-1)\alpha, 1) \) with probability at least \( 1 - \beta \), where

\[
\alpha = \max_{i,m \in \{1, \cdots, M\}, i \neq m} \frac{1}{N} \left| \sum_{n=0}^{N-1} e^{j2\pi(\epsilon_i - \epsilon_m)n} \right| \tag{12}
\]

**Proof:** See Appendix B.

**Theorem 2.** For the random training signal described in Theorem 1 with length \( N \) satisfying (11) with \( \delta \in ((M-1)\alpha, 1) \), the mutual subspace incoherence \( \mu \) of \( X \) satisfies,

\[
\mu \leq L \max \left\{ \frac{\delta - (M-1)\alpha}{SL}, \alpha \right\} \tag{13}
\]

and the mutual incoherence \( \gamma \) of \( X \) satisfies,

\[
\gamma \leq \max \left( \frac{\delta - (M-1)\alpha}{SL}, \alpha \right) \tag{14}
\]

with probability at least \( 1 - \beta \) where \( \alpha \) is given by (12).
Proof: See Appendix C.

Previously, the block RIP and $\mu$ of the sensing matrix in the absence of frequency offset ($\epsilon_q = 0$) has been characterized in [17]. Due to the presence of the frequency offsets, the structure of our sensing matrix $X$ in (8) is different from the sensing matrix for the model considered in [17]. The Rademacher distributed training sequences considered in [17] can not be used to study our sensing matrix (more details are given in Appendix A), instead, we need to use complex sequences with uniform phase. Our results in Theorem 1 and Theorem 2 illustrate the impact of the number of frequency bins $M$ on the block-RIP, $\mu$ and $\gamma$ of the sensing matrix through the bound in (11) and through the value of $\alpha$ in (12). If the bin frequencies in $E$ are not orthogonal over the duration of $N$, then $\gamma$, $\mu$ and block RIP constant $\delta$ of $X$ are bounded away from zero, unlike the results from [17]. We will see in the following section that smaller values of $\mu$, $\gamma$ and RIP constant $\delta$ give better recovery performance.

B. Detection Algorithms

In this subsection, we apply existing as well as new block sparse recovery algorithms to recover the neighbor BS parameters (cell-identities and frequency offsets) from our model in (8) and theoretically establish their detection performance.

1) Mixed Norm Minimization (MNM): The super-block channel vector $\tilde{z}$ is recovered by the solution to the following convex optimization which minimizes the $\ell_2/\ell_1$ mixed norm as [14], [17]

$$
\min_{\tilde{z}} \sum_{i=1}^{M} \sum_{q=1}^{C} \|\tilde{z}^i_q\|_2 \text{ subject to } \|X\tilde{z} - y\|_2 \leq \lambda
$$

(15)

where $\lambda$ is a threshold of our choice. Let $\hat{\tilde{z}}$ denote the reconstructed super-block channel vector (solution to (15)). From block-RIP characterization in Theorem 1 and [17, Theorem 3], we have the following recovery guarantee of MNM.

**Theorem 3.** For the model in (8), if the effective noise is bounded as $\|\tilde{w}\|_2 \leq \lambda$ and if $X$
satisfies block-RIP of order $2S$ with block-RIP constant $\delta < \frac{2\sqrt{2} - 1}{7} \approx 0.2612$, then the solution of optimization problem (15) satisfies the following inequality,

$$\|\hat{z} - \hat{\hat{z}}\|_2 \leq \frac{4\sqrt{1 + \delta}}{1 - (1 + 2\sqrt{2})\delta} \lambda. \quad (16)$$

It follows that MNM recovers $\hat{z}$ perfectly in the absence of effective noise ($\lambda = 0$) when $X$ satisfies the sufficient block RIP condition given in the above theorem. For the noisy case, we recover the neighbor BS parameters as follows. Note that the reconstructed vector $\hat{z}$ has $MC$ blocks $\{\hat{z}_q^i\}$, corresponding to $M$ frequency bins and $C$ cell identities. For each cell identity $q \in I$, we estimate the best frequency bin $\hat{\epsilon}_q$ as

$$\hat{\epsilon}_q = \epsilon_i, \quad \text{where } i = \arg\max_{m \in \{1, \ldots, M\}} \|\hat{z}_q^m\|_1, \quad (17)$$

and set the corresponding recovered channel vector for that cell identity as $\hat{z}_q = \hat{z}_q^i$. We find the set of neighbor BS $\hat{B}$ with $|\hat{B}| \leq S_{\max}$ such that $\forall q \in \hat{B}$, we have $\|\hat{z}_q\|_1 \geq \tau_{mmn}$ and $\|\hat{z}_q\|_1 \geq \|\hat{z}_p\|_1$ for any $p \notin \hat{B}$. This simply amounts to choosing the top at-most $S_{\max}$ cells in terms of the $\ell_1$ norm for their reconstructed channel (provided they are above a chosen threshold $\tau_{mmn}$). The value of threshold gives a trade off between successful detection of all neighbor BS versus falsely identifying wrong cell identities.

2) Subspace Correlation Pursuit (SCP): SCP, described below, is a greedy block sparse recovery algorithm [14], [26] and an extension of orthogonal matching pursuit to exploit the block sparse structure.

Step 1. Initialize $i = 1, \hat{\hat{B}} = \emptyset, \hat{\hat{B}}_\epsilon = \emptyset, r_0 = y, \hat{X} = \emptyset$.

Step 2. Find the index $(\hat{q}_i, \hat{\epsilon}_i)$ such that

$$(\hat{q}_i, \hat{\epsilon}_i) = \arg\max_{p \in I \setminus \hat{B}, \epsilon \in E} \|X^*_p(r_{i-1})\|_2$$

Step 3. $\hat{B} \leftarrow \hat{\hat{B}} \cup \{\hat{q}_i\}, \hat{\hat{B}}_\epsilon \leftarrow \hat{\hat{B}}_\epsilon \cup \{(\hat{q}_i, \hat{\epsilon}_i)\}$
Step 4. Concatenate: \( \hat{X} = [\hat{X} X_{q_i,\hat{\epsilon}_i}] \)

Step 5. \( \hat{z} = \arg \min_z \|y - \hat{X}z\|_2 \)

Step 6. \( r_i \leftarrow y - \hat{X}\hat{z} \) and \( i \leftarrow i + 1 \)

Step 7. If \( i \leq S_{\text{max}} \) and \( \|r_i\|_2 \geq \tau_{\text{scp}} \), go to Step 2, else set \( \hat{z} \leftarrow \hat{z} \) and stop.

When the algorithm stops, \( \hat{B}_e \) gives the set of detected BS along with their corresponding frequency offset bin index. The reconstructed CIR of detected cells are available in \( \hat{z} \). SCP chooses a block \( X_{q_i,\epsilon} \) which gives the maximum correlation with the residual in each iteration. The residual after \( i \) iterations corresponds to the projection of the observation vector onto the orthogonal complement of the column space of all the previously chosen blocks. The metric in Step 2 of SCP is similar to sufficient signal metric derived in [15] under the assumptions that the CIRs corresponding to all the neighbor BS are uncorrelated and have the same covariance matrix with high SNR. We establish the following recovery guarantees of the SCP algorithm for our synchronization problem.

**Theorem 4.** For the observation model given in (8) with \( \|\tilde{w}\|_2 \leq \lambda \) and the threshold in Step 7 set as \( \tau_{\text{scp}} = \lambda \), the parameters (cell identity and the closest frequency offset bin) recovered by SCP at the \( i \)th iteration are identical to the true values for all \( i \leq S \) (i.e., \( \hat{q}_i \subset B, \hat{\epsilon}_i = \bar{\epsilon}_{q_i} \)), if

\[
S < \frac{1}{2L} + \frac{1}{L\gamma} \left(\frac{1}{2} - \sqrt{\frac{L\lambda}{d}}\right)
\]  

and the residual after \( S \) iterations satisfies

\[
\|r_S\|_2 \leq \lambda \left(1 + \frac{1}{\sqrt{1 - \mu(S - 1)}}\right)
\]

where \( d = \min_{q \in B} \|z_q\|_2 \).

**Proof:** See Appendix D.

If the algorithm runs for exactly \( S \) iterations, then SCP perfectly identifies the set \( B \) with no false detections. It should be noted that if the algorithm is set to run for \( S_{\text{max}} \) iterations without
any threshold \( \tau_{scp} = 0 \), then the MT will perfectly detect all the BS in neighbor set \( \mathcal{B} \) if (18) is satisfied but there will be \( (S_{\text{max}} - S) \) falsely detected cell identities.

3) **Subspace Matching Pursuit (SMP)**: SMP is another greedy algorithm for block sparse signal recovery in CS framework [25], [27], [28]. With \( \Pi_{q,\epsilon}(r) \) denoting the projection of vector \( r \) onto the column space of block \( X_{q,\epsilon} \), at \( i^{th} \) iteration, SMP recovers parameters of a neighbor BS as

\[
(\hat{q}_i, \hat{\epsilon}_i) = \arg \min_{p \in \mathcal{I} \setminus \hat{B}, \epsilon \in \mathcal{E}} \| r_{i-1} - \Pi_{p,\epsilon}(r_{i-1}) \|_2
\]

which replaces the step 2 in SCP algorithm. The remaining steps are same as SCP. A block \( X_{q,\epsilon} \) is chosen by SMP, if the error corresponding to the projection of the residual onto the column space of that block has minimal norm compared to the other blocks. SMP involves more computation compared to SCP due to the projection of residual onto the column space of each block. Applying the recovery guarantee for the SMP algorithm from [17], [25] to our model, we have the following result.

**Theorem 5.** For the observation model given in (8) with \( \| \tilde{w} \|_2 \leq \lambda \), and stopping threshold set as \( \lambda \), the parameters (cell identity and the closest frequency offset bin) recovered by SMP at the \( i^{th} \) iteration are identical to the true values for all \( i \leq S \) (i.e., \( \hat{q}_i \subset \mathcal{B}, \hat{\epsilon}_i = \bar{\epsilon}_{q_i} \)), if

\[
S < \frac{1}{2} + \frac{1}{\mu} \left( \frac{1}{2} - \frac{\lambda}{c} \right)
\]

and the residual after \( S \) iterations satisfies (19), where \( c = \min_{q \in \mathcal{B}} \| X_{q,\bar{\epsilon}_q} z_q \|_2 \).

4) **Single-step Subspace Matching Pursuit (SSMP)**: In SSMP, we avoid the iterative procedure of SMP and recover the parameters of all the neighbor BS in a single step, based on the projection norm as follows.

Step 1: Compute projection norm \( G_{i,k} = \| \Pi_{i,\epsilon^k}(y) \|_2 \) for \( i \in \mathcal{I}, \epsilon^k \in \mathcal{E} \)

Step 2: For each cell identity \( q \in \mathcal{I} \), find the detected bin index corresponding to the maximum projection norm as \( i_q = \arg \max_{m \in \{1, \ldots, M\}} G_{q,m} \).
Step 3: Choose neighbor set $\hat{B}$ with $|\hat{B}| \leq S_{\text{max}}$ such that $\forall q \in \hat{B}$, we have $G_{q,i_q} > \tau_{\text{ssmp}}$ and $G_{q,i_q} > G_{p,i_p}$ for any $p \notin \hat{B}$.

Essentially, SSMP chooses the top (at-most $S_{\text{max}}$) distinct cell identities in terms of their projection norm values (provided they are above a chosen threshold).

**Theorem 6.** *For the observation model in (8) with $\|\hat{w}\|_2 \leq \lambda$, the parameters (cell identities and closest frequency offset bins) corresponding to the top $S$ largest values of $\|\Pi_{i,\epsilon^k}(y)\|_2$ are identical to the true values, if*

$$S < \frac{1}{2} \left( 1 + \frac{1}{\mu c_1} \right) - \frac{\lambda}{\mu c_1}$$

*where $c_1 = \max_{q \in B} \|X_{q,\bar{\epsilon}q}z_q\|_2, c_2 = \min_{q \in B} \|X_{q,\bar{\epsilon}q}z_q\|_2$.*

*Proof:* See Appendix E.

For the chosen threshold, if $|\hat{B}| = S$, SSMP correctly identifies the neighbor set $B$ with no false alarms and when $|\hat{B}| \neq S$, the parameters corresponding to $\min \left( S, |\hat{B}| \right)$ largest values of $G_{q,i_q}$ are the true values if (22) is satisfied. For $\lambda = 0$, SSMP recovery condition is equivalent to amplification of $\mu$ by a factor of $\frac{c_1}{c_2}$ when compared with the SMP recovery condition. Hence, complexity reduction in SSMP by avoiding the iterative process of SMP results in a poorer detection performance.

5) **Joint Correlation Detector (JCD):** When the timing window detected by JCD is correct, the performance of JCD (5) in recovering the neighbor BS parameters is characterized as follows.

**Theorem 7.** *For the observation model in (8) with $\|\hat{w}\|_2 \leq \lambda$, the parameters (cell identities and closest frequency offset bins) corresponding to the top $S$ largest values of $\|X_{i,\epsilon^k}^*y\|_2^2$ are identical to the true values, if*

$$S < \frac{1}{2} + \frac{d_2}{d_1} \left[ \frac{1}{2L} \left( 1 + \frac{1}{\gamma} \right) - \left( \frac{1}{2} + \frac{\lambda}{\sqrt{L}\gamma d_2} \right) \right]$$

*where $d_1 = \max_{q \in B} \|z_q\|_2, d_2 = \min_{q \in B} \|z_q\|_2$.\]
Given the detected timing, JCD chooses the blocks in a single step based on the largest correlation metric values $\|X_{i,c,k}^* y\|_2^2$. Above theorem can be proved along the similar principles of SSMP proof, using the bounds of the correlation metrics obtained in the proof of SCP. When compared to other recovery algorithms, JCD poses stringent condition for recovery of the parameters of all the BS in the neighbor set $\mathcal{B}$.

V. SIMULATION RESULTS

We study the synchronization performance of the algorithms using Monte-Carlo simulations. We use the following parameters; synchronization signal bandwidth of 1 MHz, carrier frequency of 2 GHz, frequency offsets are (uniformly distributed) within the range of ±7 ppm (closely aligned with the LTE specifications). We consider AWGN noise and use extended typical urban model and extended pedestrian A model [29] for the multipath channels from different BS along with Doppler corresponding to a mobility of 30 Kmph. The number of neighbor BS is set as $S = 6$, and is unknown at MT, which sets the value of $S_{\text{max}} = 9$. For the neighbor BS power levels, we consider two scenarios, 1) equal - the average received power levels of all the neighbor BS are equal and 2) unequal - the difference between strongest and weakest is set as 10 dB and the remaining power levels are chosen uniformly in between. Successful recovery of parameters of a neighbor BS is declared if its identity, frame timing and its closest frequency offset bin index are all recovered correctly. In order to fairly compare all the algorithms, we choose their thresholds to satisfy the constraint that the average number of falsely detected cells is less than 1, i.e., $E\{|\hat{\mathcal{B}} \setminus \mathcal{B}|\} < 1$. We benchmark the performance of our algorithms with that of the ideal ML detector described in (4), which in addition knows the exact value of $S$.

In Figure 2, we plot the detection performance versus signal to noise ratio (SNR), with CMCSUP training sequences described in Theorem 1. Signal power is computed by summing the received powers from all the neighbor BS. In Figure 2 (a), we plot the probability of correct timing detection of ML and JCD. We also compare with matched filter (MF) based timing
detection [5] \( \hat{T}_{MF} = \arg \max_k \max_q |\sum_{n=0}^{N-1} y(k + n)x_q^*(n)| \). In the absence of interferers, i.e., \( S = 1 \), matched filter detection is optimal in the flat Rayleigh fading channels. It can be seen that JCD performs close to ML and is significantly better than MF for both equal (Eq) and unequal (Ueq) power levels. Performance of ML in both equal and unequal power is almost identical. In Figure 2 (b), we plot the probability of successfully identifying the synchronization parameters of all the base stations in the neighbor set \( B \). Here the frame timing is obtained using JCD algorithm and \( \hat{B} \) along with closest frequency offset bin indices are obtained using corresponding detection algorithms described in Section IV-B. We also compare with the performance of the successive interference cancellation (SIC) technique presented in [15] (with suitable modifications to handle the frequency offsets). Our proposed two stage approach of timing detection using JCD and subsequent cell identity detection and frequency offset estimation using CS algorithms perform close to ML and are superior to SIC method proposed in [15] and the single stage JCD which recovers the synchronization parameters as per (5).

We also study the successful synchronization parameters identification for all the neighbor cells in \( B \) for LTE frame structure with LTE synchronization signals [1], [4] and is shown in Fig. 3. In LTE, there are 504 unique cell identities. Synchronization signals are present in two locations in each LTE frame, referred as preamble and midamble. Synchronization signal of a BS comprises of primary and secondary synchronization sequence (PSS and SSS) transmitted consecutively. PSS remains the same in both preamble and midamble while SSS differs. When PSS and SSS are concatenated, there are a total of \( 504 \times 2 \) sequences (of total length 124) which convey the cell-identity and preamble/midamble information. Our algorithms use this concatenated PSS and SSS sequences to recover the synchronization parameters. In both equal and unequal power levels of neighbor BS, our proposed methods (JCD timing detection followed by CS based recovery) perform better than the SIC method from [15]. SIC performance is worse in the equal interference case, in agreement with the common wisdom that SIC works well in the strong or weak interference regimes. On the other hand, MNM performs better in the equal
power case since it recovers the parameters of all the neighbor BS jointly.

In Fig. 4, we show the performance of our approaches in detecting the synchronization parameters for the BS with the weakest received power. For this we set the average received power levels from all the interfering neighbor BS (except the weakest one) at 0 dB. The detection performance plotted versus signal to interference plus noise ratio (SINR) of the weakest base station (treating other neighbor base stations as interference) shows that our proposed methods work well even at very low SINR levels.

In Figure 5, we show the performance of detecting synchronization parameters of all the BS present in $\mathcal{B}$ versus the number of neighbors $S$. For this study, we set the power levels of all the neighbor BS at 0 dB and the $S_{\text{max}}$ is set as $S + 3$. When $S$ is increased beyond $N/L$ (i.e., $SL > N$), the concatenated matrix $\hat{X}$ in step 5 of the algorithms (SMP and SCP) will have non trivial null space (thereby failing to find the least squares channel estimates) and hence they will not be able to recover $\mathcal{B}$. MNM recovers more than the aforementioned limit, however, the reconstructed channel $\hat{z}$ will be unreliable.

Figure 6 shows the effect of number bins $M$ on the detection performance of the algorithms.
As $M$ increases, the approximation in (6) gets better and improves the variance of the residual noise $\tilde{w}$ in (8). On the other hand, increase of $M$ makes the angle between column spaces of constituent blocks in $\mathbf{X}$ smaller, and results in the increase of $\alpha$ in (12). Since the training sequence length $N$ remains fixed, increasing $M$ beyond some value breaks the condition in (11) (due to the increase in $\alpha$) and the theoretical recovery guarantees fail. Due to these two contrasting impacts of $M$, the performances of the algorithms improves up to some value of $M$. 

Fig. 3. LTE Synchronization Performance; $M = 3$.

Fig. 4. LTE Synchronization with the weakest power BS; $M = 3$.
Fig. 5. LTE Training Sequences; $M = 3$, SNR = 12 dB.

After which degradation occurs.

Fig. 6. CMCSUP Sequences; $N = 128$, $F = 144 \times 10^2$, $C = 500$, SNR = 10 dB

VI. Conclusions

We considered the downlink synchronization problem in HetNets to identify the synchronization parameters of all the BS in the vicinity of the mobile terminal. We proposed a two
stage approach involving JCD based timing detection and CS based recovery of cell-identities and frequency offsets. We studied the performance of the proposed techniques both analytically and numerically. Our proposed methods, perform close to ML upper bound and are superior to the existing SIC and MF algorithms. Future work involves extension of the proposed ideas to utilize observations from multiple frames/antennas using the multiple measurement models of compressive sensing.

APPENDIX

A. Bound on off diagonal entries in Gram matrix of $\mathbf{X}$

In this section, we bound the absolute values of the off diagonal entries of full gram matrix $\mathbf{G} = \mathbf{X}^* \mathbf{X}$ when the training signals are randomly generated as given in Theorem 1. This bound will be subsequently used in characterizing block RIP constant $\delta$, mutual incoherence $\gamma$ and mutual subspace incoherence $\mu$ of the sensing matrix $\mathbf{X}$. The absolute values of the entries in the matrix $\mathbf{G}$, which are the inner products between the columns of $\mathbf{X}$, are represented in general form as

$$\Gamma(p,q,\ell,k,i,m) = \left| \sum_{n=0}^{N-1} x_p(n-k)x_q^*(n-\ell)e^{j2\pi[\epsilon^i(n-k) - \epsilon^m(n-\ell)]} \right|,$$

for $p,q \in \mathcal{I}$, $i,m \in \{1, \ldots, M\} = \mathcal{M}$, $k,\ell \in \{0, \ldots, L-1\} = \mathcal{L}$.

Let us collect the entries in $\mathbf{G}$ in a set as $\Phi = \{\Gamma(p,q,\ell,k,i,m) : p,q \in \mathcal{I}, \ell,k \in \mathcal{L}, i,m \in \mathcal{M}\}$. We write $\Phi = \Phi_D \cup \Phi_R$ where $\Phi_D = \{\Gamma(p,q,\ell,k,i,m) : p = q, \ell = k, i = m\} \cup \{\Gamma(p,q,\ell,k,i,m) : p = q, \ell = k, i \neq m\}$ and $\Phi_R = \Phi \setminus \Phi_D$. $\Phi_D$ contains the deterministic entries in $\mathbf{G}$ while $\Phi_R$ contains the random entries of $\mathbf{G}$. $\Phi_D$ contains diagonal entries in $\mathbf{G}$ which are equal to one. The off diagonal entries of $\mathbf{G}$ present in $\Phi_D$ can be written as,

$$\left| \sum_{n=0}^{N-1} x_p(n-\ell)x_q^*(n-\ell)e^{j2\pi[\epsilon^i(n-k) - \epsilon^m(n-\ell)]} \right|,$$

where $p \in \mathcal{I}, i,m \in \mathcal{M}, i \neq m, \ell \in \mathcal{L}$. The maximum value of the above deterministic term can be bounded by $\alpha$ as given in (12).
The random entries in $\Phi_R$ can be written as $\Phi_R = \{\Gamma(p, q, k, i, m) : p = q, k \neq \ell, i, m\} \cup \{\Gamma(p, q, k, i, m) : p \neq q, k, i, m\}$. These two disjoint cases are dealt separately. Let us first consider the case $p = q, k \neq \ell$ in (24) and these terms can be represented as,

$$\Gamma(p, p, k, \ell, i, m) = \left| \sum_{n=0}^{N-1} x_p(n - k)x_p^*(n - \ell)e^{j2\pi(\epsilon_i - \epsilon_m)n} \right|$$

for $p \in I, i, m \in M, k, \ell \in L$. We use complex Hoeffding’s inequality to bound the summation in (26) which can be stated as follows. For i.i.d. $\theta_k$ uniformly distributed in $[-\pi, \pi]$, and any vector $v \in \mathbb{C}^K$, we have

$$P\left(\left| \sum_{k=0}^{K-1} v_k e^{j\theta_k} \right| \geq t \right) \leq 4 \exp\left(-\frac{t^2}{4||v||^2_2}\right).$$

Note that multiplying $x_q(n)$ with any (deterministic) complex exponential results again in a phase uniformly distributed in $[-\pi, \pi]$ and thus complex Hoeffding’s inequality (27) can be applied to bound the sum in (26). Using Hazral-Sezemeredi theorem on equitable coloring as in [30], the terms in the summation (26) can be grouped into three sets $S_1, S_2, S_3$, in such a way that entries in different sets are independent with $S_1 \cup S_2 \cup S_3 = \{x_p(n - k)x_p^*(n - \ell)e^{j2\pi(\epsilon_i - \epsilon_m)n}\}$ for $n \in \{0, 1, \ldots, N - 1\}$ and $\lfloor \frac{N}{3} \rfloor \leq |S_w| \leq \lceil \frac{N}{3} \rceil$, for $w = 1, 2, 3$. Denoting the entries in $S_w$ by $\{s_{w,u}\}, u = 1, \ldots, |S_w|$, we can write the sum in (26) as $\Gamma(p, p, k, \ell, i, m) = \left| \sum_{w=1}^{3} \sum_{u=1}^{|S_w|} s_{w,u} \right|$. For each $S_w$, applying Hoeffding’s inequality (conveniently assuming $N$ is divisible by 3), we have for any value of $t$, $P\left(\left| \sum_{w=1}^{3} \sum_{u=1}^{|S_w|} s_{w,u} \right| \geq t \right) \leq 4 \exp\left(-\frac{t^2N}{12}\right)$ and the union bound on the sum of three such groups becomes, $P\left(\left\{| \sum_{w=1}^{3} \sum_{u=1}^{|S_w|} s_{w,u} \right| \geq t\right) \leq 12 \exp\left(-\frac{t^2N}{12}\right)$. Applying union bound over $C(L^2 - L)2M$ unique off diagonal entries of the case $p = q, k \neq \ell$,

$$P\left(\Gamma(p, p, k, \ell, i, m) \geq t\right) \leq 24ML^2C \exp\left(-\frac{t^2N}{12}\right)$$

$\forall p \in I, k, \ell \in L, i, m \in M$ and $k \neq \ell$.

$^3$This is not possible if the training sequences take values $\pm \frac{1}{\sqrt{N}}$ as in scaled Rademacher distribution.
Let us consider the other case of random sums, i.e., \( p \neq q \) in (24) and these terms can be represented as,
\[
\Gamma(p, q, k, \ell, i, m) = \left| \sum_{n=0}^{N-1} x_p(n-k)x_q^*(n-\ell)e^{j2\pi(\epsilon^i-\epsilon^m)n} \right|
\]
p, q \in \mathcal{I}, p \neq q. Since the training sequences are independent, Hoeffding inequality presented in (27) can be directly applied to bound the above term. Applying union bound over the \( 2ML^2(C^2-C) \) unique entries belong to this case, the probability that any random sum in \( \Phi_R \) belong to the case \( p \neq q \) in (24) can be bounded as,
\[
P\left( \Gamma(p, q, k, \ell, i, m) \geq t \right) \leq 8ML^2C^2 \exp \left( -\frac{t^2N}{4} \right) \quad (29)
\]
\( \forall p, q \in \mathcal{I}, \ell, k \in \mathcal{L}, i, m \in \mathcal{M} \) and \( p \neq q \). Applying union bound over (28) and (29) for both the cases, probability that absolute of any random sum \( \phi_r \in \Phi_R \), exceeding the value \( t \), \( P(|\phi_r| \geq t) \)
\[
\leq 8L^2CM \max \left\{ C \exp \left( -\frac{t^2N}{4} \right), 3 \exp \left( -\frac{t^2N}{12} \right) \right\}.
\]
With \( \beta = 48L^2CM \exp \left( -\frac{t^2N}{12} \right) \), if \( N \geq \frac{12\log \left( \frac{48L^2CM}{\beta} \right)}{t^2} \), in the above bound, second term inside the \( \max \) dominates and we have \( P(|\phi_r| \geq t) \leq \beta \). These results are summarized as follows.

**Lemma 1.** For random training signals considered in Theorem 1, the off diagonal entries of full gram matrix \( G = \mathbf{X}^*\mathbf{X} \) is bounded as follows. Any deterministic off diagonal entry (in \( \Phi_D \)) is upper bounded by \( \alpha \) given in (12). For any random entry \( \phi_r \in \Phi_R \), we have, \( P(|\phi_r| \leq t) \geq 1 - \beta \) for any \( t > 0 \), if \( N \) satisfies \( N \geq \frac{12\log \left( \frac{48L^2CM}{\beta} \right)}{t^2} \).

**B. Block RIP characterization**

We construct the matrix \( \mathbf{X}_P \) (of size \( N \times SL \)) by concatenating \( S \) blocks out of \( CM \) blocks \( \{ \mathbf{X}_{p,\epsilon^i}, p \in \mathcal{I}, \epsilon^i \in \mathcal{E} \} \) and let \( \mathcal{P} \) be a subset of the Cartesian product \( \mathcal{I} \times \mathcal{E} \) with \( |\mathcal{P}| = S \), which denotes the actual pair of indices used in the composition of \( \mathbf{X}_P \). From the Rayleigh-Ritz
theorem [31], \( \mathbf{X} \) satisfies block RIP of order \( S \) with RIP constant \( \delta \) if the eigenvalues of the gram matrix \( \mathbf{G}_P = \mathbf{X}_P^* \mathbf{X}_P \) lie in the interval \([1 - \delta, 1 + \delta]\), for all possible choices of \( P \).

Since all the diagonal entries of \( \mathbf{G}_P \) is equal to 1, from the Gershgorin disc theorem [31], it is sufficient to show that the sum of absolute values of the off diagonal entries in any column of \( \mathbf{G}_P \) is less than \( \delta \). Towards that, let us use the bound on absolute values of off diagonal entries of full gram matrix \( \mathbf{G} = \mathbf{X}^* \mathbf{X} \) given in Lemma 1 in Appendix A. If sum of any \( SL \) off diagonal entries in any column of the full gram matrix \( \mathbf{G} = \mathbf{X}^* \mathbf{X} \) is less than \( \delta \), then the eigenvalues of \( \mathbf{G}_P \) will lie in the interval \([1 - \delta, 1 + \delta]\) for any choice of \( P \).

We need to show that the worst case sum of \( SL \) off diagonal entries in any column of \( \mathbf{G} \) is less than \( \delta \). Towards that, we assume \( SL > M \) since we are interested in finding the maximum value for the sum. Note that any column of \( \mathbf{G} \) will have \( M - 1 \) deterministic off-diagonal entries whose maximum value is \( \alpha \) (from Lemma 1). Hence we have the bound on the RIP constant \( \delta \geq (M - 1) \alpha \). Now, for \( \delta > (M - 1) \alpha \), setting \( t = \frac{\delta - (M - 1) \alpha}{SL} \) in Lemma 1 in Appendix A, we are guaranteed that sum of any \( SL \) entries (including deterministic and random entries) in any column of \( \mathbf{G} \) does not exceed \( \delta \) with probability \( 1 - \beta \) if \( N \) satisfies (11).

C. Mutual Subspace Incoherence and Mutual Incoherence

The expression for mutual subspace incoherence can be alternatively written as,

\[
\mu = \max_{p,q \in \mathcal{I} \land \epsilon, \epsilon' \in \mathcal{E} \land p \neq q \lor i \neq j} \left\{ \max_{u,w \in \mathbb{C}^L} \left\| \langle \mathbf{X}_{p,\epsilon} \mathbf{u}, \mathbf{X}_{q,\epsilon'} \mathbf{w} \rangle \right\|_{2} \right\}.
\]  

(31)

Let \( \mathbf{x}_{p,i}^k \) and \( \mathbf{x}_{q,j}^k \) denote \( k^{th} \) column of \( \mathbf{X}_{p,\epsilon^i} \) and \( \mathbf{X}_{q,\epsilon^j} \) respectively. From Lemma 1 in Appendix A, we have, \( |\langle \mathbf{x}_{p,i}^k, \mathbf{x}_{q,j}^{n} \rangle| \leq \max \left\{ \frac{\delta - (M - 1) \alpha}{SL}, \alpha \right\} \) with probability at least \( 1 - \beta \) if \( N \)
satisfies (11). Now, expanding the numerator term in (31),

$$\langle X_{p,\epsilon}^* u, X_{q,\epsilon}^* w \rangle = \sum_{k=1}^{L} u_k^* \sum_{m=1}^{L} \langle x_{p,i}^k, x_{q,j}^m \rangle w_m$$

$$\leq \max \left\{ \frac{\delta - (M - 1)\alpha}{SL}, \alpha \right\} \sum_{k=1}^{L} |u_k| \sum_{m=1}^{L} |w_m|$$

$$\leq L \max \left\{ \frac{\delta - (M - 1)\alpha}{SL}, \alpha \right\} \|u\|_2 \|w\|_2$$  \hfill (32)

where the last step follows from the equivalence of $\ell_1$ and $\ell_2$ norms. Using (32) in the expression for $\mu$, we have

$$\mu \leq \max \left\{ \frac{L \max \left\{ \frac{\delta - (M - 1)\alpha}{SL}, \alpha \right\} \|u\|_2 \|w\|_2}{\|X_{p,\epsilon}^* u\|_2 \|X_{q,\epsilon}^* w\|_2} \right\}$$

$$\leq \frac{L \max \left\{ \frac{\delta - (M - 1)\alpha}{SL}, \alpha \right\}}{\min \left\{ \|X_{p,\epsilon}^* u\|_2 \|X_{q,\epsilon}^* w\|_2 \right\}}.$$

(33)

Considering the denominator term in the above equation,

$$\frac{\|X_{p,\epsilon}^* u\|_2}{\|u\|_2} = \sqrt{\frac{u^* X_{p,\epsilon}^* X_{p,\epsilon} u}{u^* u}} \geq \sqrt{\nu_{\min} \left( X_{p,\epsilon}^* X_{p,\epsilon} \right)}$$

where $\nu_{\min}$ is the minimum eigen value by Rayleigh-Ritz theorem [31]. It is evident that the off diagonal entries of $X_{p,\epsilon}^* X_{p,\epsilon}$ are the subset of random sums $\Phi_R$ explained in Appendix A and are bounded above by $\frac{\delta - (M - 1)\alpha}{SL}$ with probability at least $1 - \beta$ if $N$ satisfies (11). We can bound the sum of $L$ such entries by $\frac{\delta - (M - 1)\alpha}{SL}$ and from Gershgorin theorem, eigen values of $X_{p,\epsilon}^* X_{p,\epsilon}$ lie in the interval $[1 - \frac{\delta - (M - 1)\alpha}{S}, 1 + \frac{\delta - (M - 1)\alpha}{S}]$. Taking the minimum of eigen value in (34) and substituting in (33) we get the desired bound in (13).

The mutual incoherence defined in (10) can be written as, $\gamma = \max_{i \neq m} G_{i,m}$ where $G_{i,m}$ is the $(i,m)$th entry in the full gram matrix $G = X^* X$. From Lemma 1, the desired bound in (14) follows.
D. Recovery Guarantee for SCP

Without loss of generality, we consider the model (8) as 
\[ y = \sum_{i=1}^{S} X_{i,\epsilon^1} z_i + \tilde{w} \]
with \( \|z_1\|_2 \geq \|z_2\|_2 \geq \cdots \geq \|z_S\|_2 \), i.e., active blocks correspond to the first \( S \) cell-identities with the first bin index. SCP chooses one of the correct blocks in the first iteration if,

\[ \|X_{1,\epsilon^1}^* y\|_2 > \|X_{1,\epsilon^m}^* y\|_2 \]  \hspace{1cm} (34)

for all \( \{ t > S, \epsilon^m \in \mathcal{E} \} \) and \( \{ t \leq S, \epsilon^m \neq \epsilon^1 \} \) (which are the inactive blocks). Defining \( \tilde{y} = \sum_{i=1}^{S} X_{i,\epsilon^1} z_i \), we have,

\[ \|X_{1,\epsilon^1}^* y\|_2 \geq \|X_{1,\epsilon^1}^* \tilde{y}\|_2 - \|X_{1,\epsilon^1}^* \tilde{w}\|_2, \]  \hspace{1cm} (35)

\[ \|X_{1,\epsilon^1}^* \tilde{y}\|_2 \geq \|X_{1,\epsilon^1}^* X_{1,\epsilon^1} z_1\|_2 - \sum_{i=2}^{S} \|X_{1,\epsilon^1}^* X_{i,\epsilon^1} z_i\|_2. \]

Note that \( \|X_{1,\epsilon^1}^* X_{1,\epsilon^1} z_1\|_2 \geq \nu_{\min}(X_{1,\epsilon^1}^* X_{1,\epsilon^1}) \|z_1\|_2 \) and \( \|X_{1,\epsilon^1}^* X_{i,\epsilon^1} z_i\|_2 \leq \nu_{\max}(X_{1,\epsilon^1}^* X_{i,\epsilon^1}) \|z_i\|_2. \)

The maximum \( \nu_{\max} \) and the minimum \( \nu_{\min} \) eigen values are bounded using Lemma 1 in Appendix A and Gershgorin disc theorem as \( \nu_{\min}(X_{1,\epsilon^1}^* X_{1,\epsilon^1}) \geq (1 - (L - 1)\gamma) \) and \( \nu_{\max}(X_{1,\epsilon^1}^* X_{i,\epsilon^1}) \leq L\gamma. \) Therefore

\[ \|X_{1,\epsilon^1}^* \tilde{y}\|_2 \geq (1 - (L - 1)\gamma) \|z_1\|_2 - \sum_{i=2}^{S} L\gamma \|z_i\|_2 \]

\[ \geq (1 - SL\gamma + \gamma) \|z_1\|_2. \]

Similarly, for all \( \{ t > S, \epsilon^m \in \mathcal{E} \} \) and \( \{ t \leq S, \epsilon^m \in \mathcal{E} \setminus \epsilon^1 \} \), we have

\[ \|X_{t,\epsilon^m}^* \tilde{y}\|_2 \leq \sum_{i=1}^{S} \|X_{t,\epsilon^m}^* X_{i,\epsilon^1} z_i\|_2 \]

\[ \leq \sum_{i=1}^{S} L\gamma \|z_i\|_2 \leq SL\gamma \|z_1\|_2. \]

The correlation of noise with any block can be bounded as \( \|X_{q,\epsilon^1}^* \tilde{w}\|_2 \leq \sqrt{L}\lambda. \) Using these bounds in (34), SCP chooses the correct subspace in the first iteration if,

\[ S < \frac{1}{2L} + \frac{1}{2L\gamma} - \frac{\lambda}{\sqrt{L}\gamma \|z_1\|_2}. \]  \hspace{1cm} (36)
After the first iteration, the residual $r_1$ is orthogonal to the columns of the detected block. The above argument carries forward for each iteration and hence SCP correctly recovers the parameters of the neighbor BS in the first $S$ iterations if (18) is satisfied. Further, with this successful recovery, the residual after $S$ iterations can be written as,

$$\|r_S\|_2 = \left\| y - \sum_{q \in B} X_q \hat{z}_q \right\|_2 = \left\| \sum_{q \in B} X_q \left( z_q - \hat{z}_q \right) + \tilde{w} \right\|_2$$

$$\leq \sum_{q \in B} \|X_q e_q\|_2 + \lambda$$

(37)

and by directly applying [25, Theorem 3], the above bound (37) is equivalent to (19).

**E. Recovery Guarantee of Single-step SMP**

Without loss of generality, assume that $\mathcal{B} = \{1, 2, \cdots, S\}$ and $\bar{e}_i = \epsilon^1 \ \forall i \in \mathcal{B}$. Let us write the observation in (8) as $y = \tilde{y} + \tilde{w} = \sum_{i=1}^{S} a_i v_i + \tilde{w}$, with $|a_i| = \|X_i,\epsilon^1 z_i\|_2$, $v_i \in$ column space of $(X_{i,\epsilon^1})$ and $\|v_i\|_2 = 1$. Assume that the coefficients satisfy $|a_1| \geq \cdots \geq |a_S|$. Top $S$ largest values of the projection $\Pi_{i,\epsilon^1}(y)$ will correctly recover all the neighbor set parameters if, for $\{i \leq S\}$, we have

$$\|\Pi_{i,\epsilon^1}(y)\|_2 > \|\Pi_{t,\epsilon^m}(y)\|_2$$

(38)

for all $\{t \leq S, \epsilon^m \neq \epsilon^1\}$ and $\{t > S, \epsilon^m \in \mathcal{E}\}$. For $i \leq S$, we have $\|\Pi_{i,\epsilon^1}(y)\|_2 \geq \|\Pi_{i,\epsilon^1}(\tilde{y})\|_2 - \|\Pi_{i,\epsilon^1}(\tilde{w})\|_2$ and $\|\Pi_{i,\epsilon^1}(\tilde{y})\|_2 \geq \|\Pi_{i,\epsilon^1}(a_i v_i)\|_2 - \|\sum_{k=1, k \neq i}^{S} \Pi_{i,\epsilon^1}(a_k v_k)\|_2$. It follows that $\|\Pi_{i,\epsilon^1}(a_i v_i)\|_2 = |a_i|$ and $\|\sum_{k=1, k \neq i}^{S} \Pi_{i,\epsilon^1}(a_k v_k)\|_2 \geq |a_1| \sum_{k=1, k \neq i}^{S} \|\Pi_{i,\epsilon^1}v_k\|_2 \geq |a_1|(S-1)\mu$. The projection of effective noise on any subspace satisfy $\|\Pi_{t,\epsilon^m}\tilde{w}\|_2 \leq \lambda$. Putting these together, for all $i \leq S$,

$$\|\Pi_{i,\epsilon^1}(y)\|_2 \geq |a_S| - (S-1)\mu a_1 - \lambda.$$  

(39)

Similarly, for all $\{t > S, \epsilon^m \in \mathcal{E}\}$ and $\{t \leq S, \epsilon^m \neq \epsilon^1\}$

$$\|\Pi_{t,\epsilon^m}(y)\|_2 \leq |a_1|S\mu.$$  

(40)
Putting (39) and (40) together, recovery condition in (38) will be satisfied if $|a_S| - (S - 1)\mu|a_1| - \lambda > S\mu|a_1| + \lambda$, which is same as (22) since $|a_1| = \max_{q \in B} \|X_{q,\bar{e}_q}z_q\|_2$ and $|a_S| = \min_{q \in B} \|X_{q,\bar{e}_q}z_q\|_2$.

REFERENCES


