

“Cocktail Party in the Cloud”: Blind Source Separation for Co-operative Cellular Communication in Cloud RAN

Abolfazl Hajisami, Hariharasudhan Viswanathan, and Dario Pompili
 Department of Electrical and Computer Engineering
 Rutgers University–New Brunswick, NJ
 e-mail: {hajisamik, hari_viswanathan, pompili}@cac.rutgers.edu

Abstract—Due to the rapid growing popularity of mobile Internet, broadband cellular wireless systems are expected to offer higher and higher data rates even in high-mobility environments. Cloud Radio Access Network (C-RAN) is a new centralized paradigm for broadband wireless access that addresses efficiently the fluctuation in capacity demand through real-time inter-Base Station (BS) cooperation. An innovative Blind Source Separation (BSS)-based cellular communication solution for C-RANs, *Cloud-BSS*, which leverages the inter-BS cooperation, is proposed. *Cloud-BSS* groups contiguous cells into *clusters* – sets of neighboring cells inside which mobile stations do not need to perform handovers – and allows them to use all of the frequency channels. The proposed solution is studied under different network topologies, and a novel strategy, called *Channel-Select*, to improve the Signal-to-Noise Ratio (SNR) is introduced. *Cloud-BSS* enhances the cluster spectral efficiency, decreases handovers, eliminates the need for bandwidth-consuming channel estimation techniques, and mitigates interference. Simulation results, which are discussed along with concepts, confirm these expectations.

Index Terms—Cloud Radio Access Network; Cellular Systems; Virtual Base Station; Blind Source Separation.

I. INTRODUCTION

Over the last few years, proliferation of personal mobile computing devices like tablets and smartphones along with a plethora of data-intensive mobile applications has resulted in a tremendous increase in demand for ubiquitous and high data-rate wireless communications. To meet this demand, the fourth generation (4G) cellular communication system with peak downlink data rate of 1 Gbps has been envisioned. Long Term Evolution (LTE) systems based on Orthogonal Frequency Division Multiple Access (OFDMA) represent a major breakthrough in terms of achieving downlink peak data rates of 300 Mbps [1]. However, because LTE systems do not match yet the International Mobile Telecommunications Advanced (IMT-Advanced) “True 4G” requirements, a significant effort is being made towards the development of LTE-Advanced.

The current practice to enhance data rates is to increase the number of Base Stations (BSs) and go for smaller cells so to increase the band reuse factor. However, additional deployment and maintenance of a large number of cellular BSs are highly inefficient due to excessive capital and operating expenditures. Moreover, with small cells the Inter-Cell Interference (ICI) problem becomes more challenging. Smaller cells also lead to a higher number of *handovers* between cells, which is the process of transferring an ongoing call (or data session) from one cell to another to avoid call termination

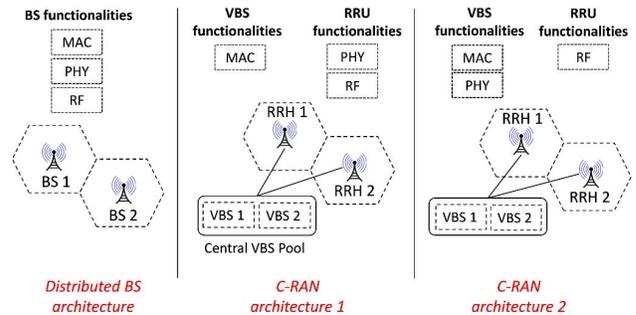


Fig. 1: Comparison of the traditional Distributed Base Station (BS) architecture against *two C-RAN architectures* differing on how the communication functionalities are split among the local Remote Radio Unit (RRH) and the Virtual Base Station (VBS) in a remote datacenter. *We focus on architecture 2.*

when the mobile user gets outside the range of a cell. Enabling frequent near-seamless and smooth handovers, with minimal service disruption, calls for very high-speed communication and co-operation among BSs.

A. Cloud Radio Access Network

Cloud Radio Access Network (C-RAN) [2] is a new paradigm for broadband wireless access that addresses the fluctuation in capacity demand efficiently while keeping the cost of delivering services to the users low. It also enables a higher degree of co-operation and communication among BSs. *C-RAN represents a clean-slate design and allows for dynamic reconfiguration of computing and spectrum resources.* Characteristics of C-RAN are: i) centralized management of computing resources, ii) reconfigurability of spectrum resources, iii) collaborative communications, and iv) real-time cloud computing on generic platforms. C-RAN is composed of Remote Radio Heads (RRHs) distributed over a wide geographic region controlled by remote Virtual Base Stations (VBSs) housed in centralized BS pools (Fig. 1). VBSs and their corresponding RRHs should be connected by high-bandwidth low-latency media (e.g., use of optic fibers allows for a maximum distance of separation of 40 Km between the RRH and its VBS) [2].

Packet-level processing, Medium Access Control (MAC), physical-layer (PHY) baseband processing, and Radio Frequency (RF) functionalities may be split between the VBSs

and the RRHs depending on the specific C-RAN architecture (Fig. 1). *In this paper, we focus on the second architecture as it allows exploitation of the full potential of this paradigm.* The communication functionalities of the VBSs are implemented (in software) on Virtual Machines (VMs) hosted over general-purpose computing platforms (servers), which are housed in one or more racks of a small cloud datacenter. In a centralized VBS pool, as all the information from the BSs is resident in a datacenter, BSs can exchange control data at Gbps speeds.

Cooperation and communication among BSs can improve user Quality of Service (QoS) and system performance by exploiting the extra degrees of freedom to make optimized decisions. For instance, efficient schemes for handover and interference cancelation, which need a high volume of data to be exchanged among cooperating BSs, can be employed in a C-RAN architecture because of the low-latency inter-BS communication. However, even in C-RAN, where we have Gbps connections among the VBSs, the aforementioned schemes need a considerable time (almost 13% of the frame time) to process the received data.

B. Contributions

We propose an uplink *cooperative joint PHY and MAC* solution for next-generation cellular communications that exploits synergistically the advantages of C-RAN and Blind Source Separation (BSS). BSS is a well-known technique in signal processing to recover the underlying source signals from a set of mixtures, where the mixing system is unknown (popularly known as the “cocktail party problem”). C-RAN’s characteristics are well suited for BSS-based cooperative cellular communications as the source separation problem relies on inter-BS cooperation. In our solution, named *Cloud-BSS*, we divide a set of neighboring cells into *clusters* and allow them to use all of the frequency channels in the system band. In each cluster, the RRHs receive a mixture of the Mobile Station (MS) signals. Then, the MS signals are separated from the mixtures through BSS. *Cloud-BSS* provides the following benefits: i) enhancement of the cluster spectral efficiency, ii) decrease in the number of handovers, iii) elimination of the need for bandwidth-consuming channel estimation, and iv) interference mitigation. We study our solution under different network topologies and introduce a strategy, named *Channel-Select*, to increase the Signal-to-Noise Ratio (SNR) of the estimated signals. In other words, *Cloud-BSS* separates the intra-cluster mixtures and *Channel-Select* mitigates the defective impact of background noise (including inter-cluster interference) during the separation process.

C. State of The Art

Centralized management of computing resources, i.e., BS pooling, renders BS information global and, hence, enables cooperative communication techniques at the MAC and PHY layers that were previously not implementable due to strict inter-BS coordination requirements (in terms of throughput and latency). Examples of MAC- and PHY-layer enhancements include joint flow scheduling and load balancing [3],

macro diversity, interference alignment and cancelation [4], and advanced mobility management [5]. Even though work has been done on the aforementioned cooperative communication techniques that can benefit from C-RAN, research on enabling technologies for C-RAN itself is at a nascent stage and, hence, there are only a few works in this area.

In [6] and [2], the authors introduce the centralized-BS idea and study its advantages, challenges, and requirements; they argue that the behavior of PHY and MAC layers are quite different and, hence, are better suited for two different processor architectures. In [7], the authors assume that only the MAC-layer functionality is centralized at the BS pool, propose a MAC-layer consolidation solution, and compare the multiplexing gain over today’s distributed BS architecture, depicted in Fig. 1(left). In [5], the authors address the problem of inter-BS communication, study the latencies involved in information exchange among distributed BSs, and consider BS pooling as a potential solution for higher degree of BS co-operation in broadband cellular networks. The authors also review a few well-known co-operative Multiple Input Multiple Output (MIMO) techniques w.r.t. their inter-BS communication needs and their challenges in a distributed architecture, and discuss how these are mitigated in a pooled BS model.

In summary, prior work on C-RANs focused on the overall system architecture, on the feasibility of virtual software BS stacks as well as on the performance gains. In contrast to existing works, which do not explore the potential of C-RAN to enhance the system spectral efficiency, we propose a novel solution based on BSS to improve the cluster (and, hence, the system) spectral efficiency and system performance using the centralized advantage of C-RAN.

The rest of the paper is organized as follows. In Sect. II, we present our solution, which aims at enhancing the cluster spectral efficiency and system performance using the combined advantages of C-RAN and BSS; first, we formalize our solution and explain why for different network topologies the system performance varies; then, we introduce our strategy to maximize the performance. In Sect. III, we discuss the other benefits of *Cloud-BSS*. Finally, in Sect. IV, we draw the main conclusions and wrap up the paper discussing future work.

II. PROPOSED SOLUTION: CLOUD-BSS

Generally, in cellular networks, neighboring cells avoid to reuse the same set of frequencies (or channels) so to keep the interference below a certain threshold and to ensure user QoS. As a drawback, however, the whole cellular band cannot be used by the cells, which leads to low data rates. Here, we introduce *Cloud-BSS*, our BSS-based solution, which increases the user capacity – thus achieving high data rates – by exploiting the characteristics of the C-RAN architecture; specifically, *Cloud-BSS* is able 1) to grant multiple users access to the same OFDMA channels simultaneously and 2) to assign multiple channels to the same user. These features, together, lead to a higher per-user capacity and hence, data rate, given a fixed cellular band. We also explain how the proposed solution mitigates the interference problem and decreases the number

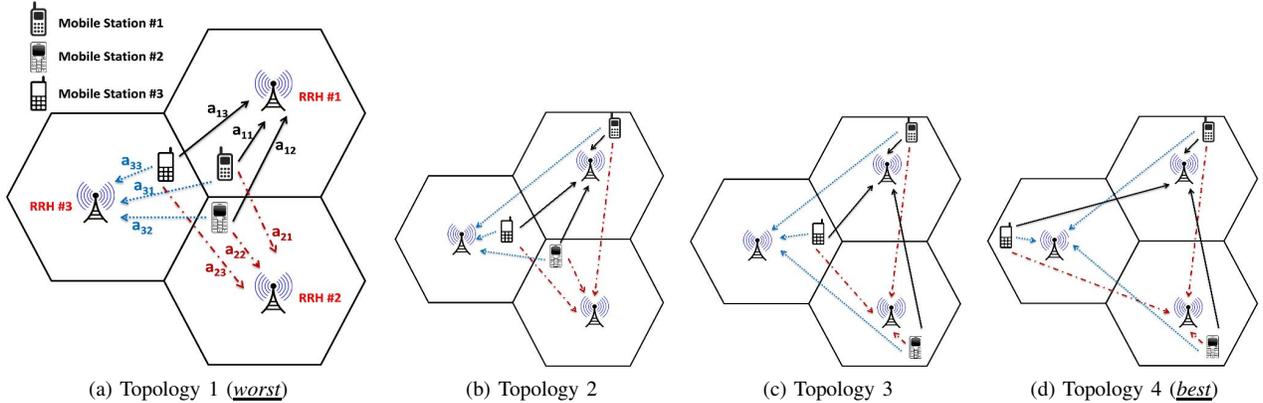


Fig. 2: **Four network topologies** in which the RRHs receive different combinations (‘mixtures’) of the MS signals. From Topology 1 to 4 (worst to best), the distance between a MS in a cell and the RRHs in the other cells becomes higher.

of handovers, regardless of the handover procedure used by the system. In fact, a handover will only be needed when a MS moves from one cluster to another (as opposed to from one cell to another). In other words, as long as a MS remains in a certain cluster, no handover is needed, which provides the following advantages: an increase in user QoS and a decrease in overall system computation and communication overhead.

In *Cloud-BSS*, we divide a set of contiguous cells into a *cluster* and allow them to use all of the frequency channels in the system band, *thus achieving a frequency-reuse factor of 1*. Hence, in each cluster, the RRHs receive a mixture of the MS signals. Figure 2(a) shows our clustering idea: here the cluster size is 3 and a_{ij} is the channel coefficient between MS # j and RRH # i . The relationship between the *received* and the *MS signals* at different time instants can be expressed through the following linear noisy model (for clarity time is omitted),

$$\mathbf{x} = \sum_{i=1}^N s_i \mathbf{a}_i + \mathbf{n} = \mathbf{A}\mathbf{s} + \mathbf{n}. \quad (1)$$

Here, $\mathbf{s} = [s_1, \dots, s_N]^T$ is the $N \times 1$ vector of complex-valued MS signals (sources), $\mathbf{x} = [x_1, \dots, x_M]^T$ is the $M \times 1$ vector of signals (mixtures) received by the RRHs, \mathbf{A} is the $M \times N$ complex-valued channel coefficient (mixing) matrix with linearly independent columns (\mathbf{a}_i being its i^{th} column), and $\mathbf{n} = [n_1, \dots, n_M]^T$ is the $M \times 1$ Gaussian noise vector with independent and identically-distributed (i.i.d.) components. Note that, in (1), the inter-cluster interference is part of the background noise, and that the MS signals are assumed to be *statistically independent*; such assumption is almost always met in practice for physically-separated transmitters. Now, to extract the MS data in the VBS we need to separate the MS signals (sources) from the received RRH signals (mixtures); *in a cluster, this is in fact a BSS problem*.

First, we provide some preliminary background on BSS and Independent Component Analysis (ICA). Then, we argue that the topology configuration of the MSs affects the system performance and show that *diagonal dominant topologies* lead to better performance, i.e., to a lower Bit Error Rate (BER).

We introduce a metric to measure how much a $M \times N$ mixing matrix is diagonal dominant and use it in the simulation analysis to show the performance associated with different topologies. Finally, we introduce a strategy, named *Channel-Select*, to ‘group’ the best set of active MSs (i.e., assign them to the same OFDMA channel) based on their locations so to induce diagonal dominance in the mixing matrices.

A. Blind Source Separation (BSS)

In BSS, a set of mixtures of different source signals is available and the goal is to separate the source signals when we have no information about the mixing system or the source signals (hence the name *blind*) [8]. The mixing and separating systems can be represented mathematically as,

$$\mathbf{x}(t) = \mathbf{A}\mathbf{s}(t), \quad \mathbf{y}(t) = \mathbf{B}\mathbf{x}(t), \quad (2)$$

where $\mathbf{s}(t) = [s_1(t), \dots, s_N(t)]^T$ is the vector of sources that are mixed by the mixing matrix \mathbf{A} and $\mathbf{x}(t) = [x_1(t), \dots, x_M(t)]^T$ is the vector of available observations. Let \mathbf{A} be a $M \times N$ matrix of full-column rank, which means that the observations are linearly independent; the goal is to design a separating matrix \mathbf{B} such that $\mathbf{y}(t) = [y_1(t), \dots, y_N(t)]^T$ is an estimate of the sources. A method to solve BSS is ICA, which exploits the assumption of *source independence* and estimates \mathbf{B} such that the outputs $y_i(t)$ s are statistically independent. For this assumption to hold, however, the number of observations must be equal or greater than the number of sources (i.e., $M \geq N$). The essence of ICA can be understood better by considering the “*cocktail party problem*,” in which many people are talking simultaneously: if several microphones at different positions are available, then different mixtures of the voices can be recorded. Given such mixtures and the assumption that the original voice signals are independent from each other, ICA can recover the original voices from the mixtures. However, most of the ICA algorithms are only applicable to real signals, whereas in digital communication systems we deal with complex-valued signals. To solve this problem, some ICA algorithms have been proposed (such as the ones in [9]–[11]) to deal with complex-valued signals.

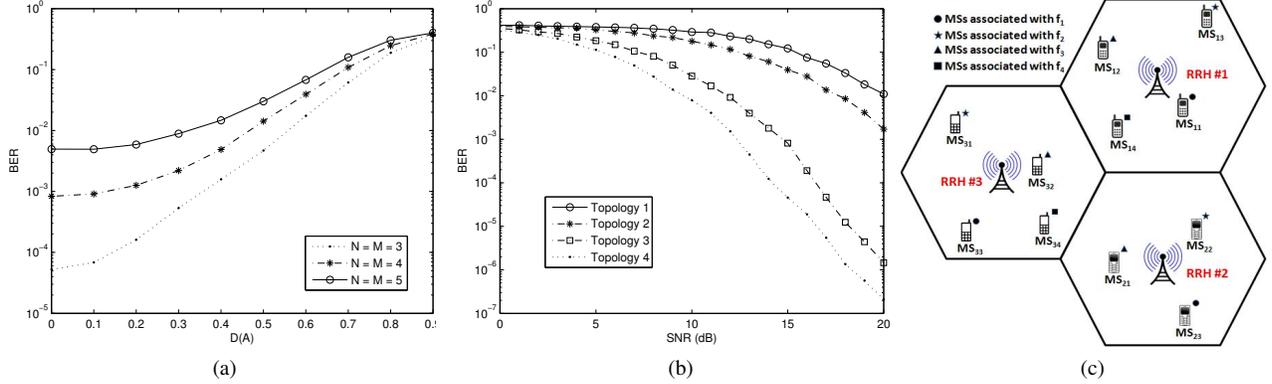


Fig. 3: (a) Increase in Bit Error Rate (BER) with increase in non-diagonal dominance of mixing matrix \mathbf{A} (the closer $D(\mathbf{A})$ to 1, the more \mathbf{A} is non-diagonal dominant); note that in this simulation we considered $N = M$; (b) Decrease in BER with Signal-to-Noise Ratio (SNR) and topology order for each of the four different topologies considered in Fig. 2; (c) Example for allocating frequency channels to the MS's.

B. System Performance for Different Topologies

In cellular networks the interfering signals from other cells decrease the performance of the system. To overcome this problem, BSs avoid reuse of the same set of frequencies. However, in the case of C-RAN, as we have access to all of the BSs' received signals in a cluster, the MS signals can be separated by the use of ICA algorithms. Hence, with reference to the model in (1), since we have access to all the x_i s ($1 \leq i \leq M$), we can separate the MS signals (sources) from the RRH signals (mixtures) through ICA. Note that it was studied in [9] that the case of multiple paths, where several coherent wireless signals from a single transmitter are mixed in the received signal, does not affect the ICA problem.

Without considering the noise – and as long as the sources are independent and the mixing channel coefficient matrix is full rank – ICA methods can extract the source signals simply by estimating the inverse of the mixing matrix. However, in the presence of noise, when the ICA algorithms estimate \mathbf{B} as $\hat{\mathbf{A}}^{-1}$ and multiply it by the observation so to extract the source signals, from (1) we obtain,

$$\mathbf{B}\mathbf{x}(t) = \hat{\mathbf{s}}(t) + \mathbf{B}\mathbf{n}(t), \quad (3)$$

where each estimated source signal is associated with a combination of the additive noises at all the receivers. If we assume that the noises at all the receivers have the same variance σ_n^2 , then the noise in the i^{th} estimated source has a variance of $(b_{i1}^2 + \dots + b_{iM}^2) \cdot \sigma_n^2$, where b_{ij} is the $(i, j)^{\text{th}}$ component of the separation matrix \mathbf{B} . As the b_{ij} s are dependent on the determinant of the mixing matrix \mathbf{A} , the noise level of the estimated sources is highly dependent on the mixing matrix. The determinant of a matrix is the volume of the parallelepiped composed of its rows/columns. It is straightforward to prove

that if a diagonal dominant¹ and a non-diagonal dominant matrix have the same row/column norm, then the former has the greater determinant and, hence, lower component values in its inverse matrix [12].

Theorem 1. *Let us assume that a $N \times N$ matrix \mathbf{C} is diagonally dominant by rows, and let us set $\beta = \min_i (|c_{ii}| - \sum_{j \neq i} |c_{ij}|)$. It follows that $\|\mathbf{C}^{-1}\|_{\infty} < 1/\beta$. Proof in Appendix.*

From Theorem 1, we infer that when β is low the upper bound of $\|\mathbf{C}^{-1}\|_{\infty}$, which is the maximum absolute row sum of the matrix \mathbf{C}^{-1} , becomes high [13]. Hence, considering \mathbf{A} to be a $N \times N$, when the non-diagonal components of the mixing channel coefficient matrix \mathbf{A} in each row are close to the diagonal component, β becomes lower and, as a drawback, the maximum absolute row sum of the separation matrix \mathbf{B} becomes higher, leading to low SNR in the estimated sources in (3). Hence, in order to have a high SNR in the estimated source signals, the absolute value of the diagonal dominant component of \mathbf{A} should be as high as possible, and the absolute value of non-diagonal components of \mathbf{A} should be as low as possible. *This translates into the following observation: in a certain frequency channel, the i^{th} MS needs to be as close as possible and the j^{th} MS needs to be as far as possible to the i^{th} RRH (with $j \neq i$).*

In our solution, where the mixing channel coefficient matrix depends on the topology of the network, we expect that for different network topologies the performance would vary. Depending on the topology, the mixing channel coefficient matrix \mathbf{A} and thus the variance of the noise in the estimated sources are different. As we discussed earlier, we expect that for topologies with a diagonally dominant mixing (channel

¹A matrix is said to be *row/column diagonally dominant* if, for every row/column, the magnitude of the diagonal entry in a row/column is larger than or equal to the sum of the magnitudes of all the other entries in that row/column; i.e., $\mathbf{A} = (a_{ij})_{ij}$ is *row diagonally dominant* if $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|, \forall i$; and *column diagonal dominant* if $|a_{ii}| \geq \sum_{i \neq j} |a_{ij}|, \forall j$.

coefficient) matrix the performance would be better than for topologies with a non-diagonally dominant mixing matrix; in the diagonally dominant case, in fact, the variance of the noise associated with each estimated transmitted signal is lower than in the non-diagonally dominant case. Hence, *the more diagonal dominant the mixing matrix, the better the performance*. Based on the formulation we described for Fig. 2(a), we introduce metric $D(\mathbf{A})$ to define how much a $M \times N$ mixing matrix \mathbf{A} is row diagonally dominant as,

$$D(\mathbf{A}) = \frac{1}{N(N-1)} \sum_{j=1}^N \left(\frac{\sum_{k=1}^N |a_{jk}|}{\max_i |a_{ij}|} - 1 \right), \quad (4)$$

where a_{ij} is the $(i, j)^{th}$ component of matrix \mathbf{A} and $\max_i |a_{ij}|$ is the maximum absolute value in the j^{th} column of the matrix. In (4), we find the maximum components in each column of matrix \mathbf{A} , then perform a normalized sum only over those rows where the maximum components exist. In fact, in the mixing matrix \mathbf{A} , the components of each column/row correspond to a certain MS/RRH, respectively. In the case where $N < M$, (4) eliminates $(M - N)$ rows corresponding to the RRHs for which there is no MS in their cells. With this definition, $D(\mathbf{A})$ always ranges in $[0, 1]$, being equal to 0 when \mathbf{A} is diagonal. For the topologies in which $D(\mathbf{A})$ is high, the maximum absolute row sum of the separation matrix \mathbf{B} becomes higher than for the topologies with lower $D(\mathbf{A})$. An increase in $D(\mathbf{A})$ causes a decrease in SNR and consequently an increase in the BER. Figure 3(a) shows the increase of the BER as $D(\mathbf{A})$ approaches 1, for the cases of $N = M = 3, 4, 5$ and SNR = 15 dB.

To verify that the system performance depends on the topology – as inferred from the above mathematical analysis – we analyzed four topologies, as shown in Fig. 2, and considered a cluster of three cells. We implemented the JADE algorithm [9] for separating the MS signals from their mixtures. JADE uses the whole fourth-order statistics of the received data, from which very good separation results can be achieved. In Topology 1 (worst), the channel coefficient matrix \mathbf{A} is not diagonally dominant at all, whereas in Topology 4 (best) all of the columns of the matrix are diagonally dominant. In between, we have Topology 2 and 3, in which only one and two columns, respectively, are diagonally dominant. If we consider $D_l(\mathbf{A})$ as the diagonal dominance metric of the l^{th} topology, then we have $D_4(\mathbf{A}) < D_3(\mathbf{A}) < D_2(\mathbf{A}) < D_1(\mathbf{A})$. Consequently, moving *orderly* from Topology 1 to 4, we expect progressively better performance due to the decrease in noise level of the estimated MS signals. The decreasing BER curves with the SNR (*not surprising*) and with topology order (*corroborating our analysis*) are depicted in Fig. 3(b).

C. Proposed Strategy for Operation under Non Idealities

We propose the *Channel-Select* strategy, which changes the transmitting frequency (channel) of each user so to make the mixing matrices “as diagonally dominant as possible”. Depending on the ‘instantaneous’ topology, we optimize the

frequency allocation for each MS so that the channel coefficient matrix in each channel is as diagonal dominant as possible. To achieve this goal, *Channel-Select* uses the following objective function,

$$\mathcal{F}(\mathbf{A}_1, \dots, \mathbf{A}_L) = \min \max [D(\mathbf{A}_1), \dots, D(\mathbf{A}_L)], \quad (5)$$

where \mathbf{A}_l is the mixing matrix in the l^{th} frequency channel f_l and L is the total number of channels available in the OFDMA system. The complexity of this combinatorial optimization problem depends on the number of MSs and cluster size, and grows exponentially as these parameters increase. Assuming K MSs are uniformly distributed in the cluster, the problem complexity is $\mathcal{O}((K/M)^M)$: consequently, as we are dealing with a large number of MSs in a cluster, a solution under real-world timing constraints is infeasible. Therefore, we introduce a simple heuristic algorithm that makes the diagonal and non-diagonal component of matrix \mathbf{A} higher and lower, respectively, causing $D(\mathbf{A}_l)$'s in (5) to be low.

Without any loss in generality, we assume that: **(i)** We have L frequency channels (f_1, \dots, f_L) and K MSs; **(ii)** We have enough frequency channels to admit all the users in each cell; **(iii)** All cells have the same size and in each cluster all the RRHs can receive all the MSs’ signals; **(iv)** All MSs use the same output power (i.e., no power control is performed); **(v)** We approximately know where the MSs are, but we do not know their trajectory (i.e., no horizon).

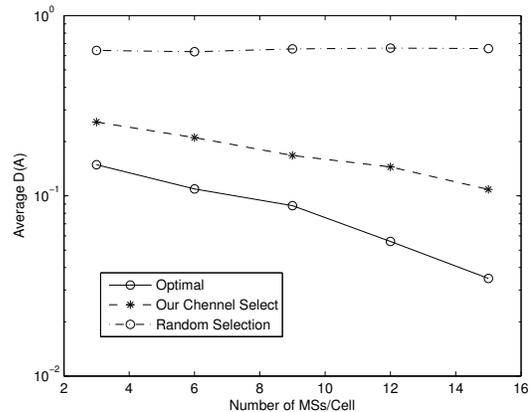


Fig. 4: *Channel-Select* vs. optimal solution (by exhaustive search) for small problem sizes (cluster size $M = 3$).

With these realistic working assumptions, we propose *Channel-Select*, an algorithm to allocate the frequency channels to the MSs in such a way that $D(\mathbf{A}_l)$'s become as low as possible. Figure 3(c) clarifies our explanations for the case when the cluster size M is 3 and the number of channels L is 4. Firstly, based on the location of the MSs, we calculate the distances between the MSs and RRHs. Then, we define how many MSs should be allocated to each frequency channel as $N = \lceil K/L \rceil$ (as in Fig. 3(c), where $K = 11$, $L = 4$, and $N = 3$). Furthermore, we find the nearest MS in cell #1 (MS_{11}) and $N - 1$ farthest MSs in the other cells (MS_{23} and MS_{33}) to the RRH #1 and allocate channel f_1 to this group

TABLE I: Computational Complexity, Run Time (RT), and Performance Index (PI) of three competing ICA methods for five parameters: N = no. of sources, M = no. of mixtures, T = no. of data samples, Q = no. of iterations, and I = no. of executed sweep; in the simulations, we considered $N = M = 5$.

Method	Computational Complexity (flops)	T=2000		T=4000		T=8000	
		RT [s]	PI	RT [s]	PI	RT [s]	PI
JADE [9]	$\min\{TM^2/2 + 4M^3/3 + NMT, 2KM^2\} + 3TN(N+1)(N^2+N+2)/8 + TN^2 + \min\{4N^6/3, 8N^3(N^2+3)\} + IN(N-1)(75+21N+4N^2)/2$	0.03445	7.08145	0.04657	2.10674	0.07535	0.43901
Complex ICA-EBM [10]	$\min\{TM^2/2 + 4M^3/3 + NMT, 2TM^2\} + 4M^3/3 + (Q-1)M^3/2 + IN(N-1)(17(M^2-1) + 75 + 4N + 4N(M^2-1))/2$	1.05756	8.07938	1.50168	2.13367	1.69157	0.35794
Complex FastICA [11]	$\min\{TM^2/2 + 4M^3/3 + NMT, 2TM^2\} + (2(N-1)(N+T) + 5TN(N+1))/2K$	0.45037	9.33160	0.75995	5.77939	1.38608	1.78655

of MSs. We repeat this procedure for the remaining cells and MSs until when all the MSs have been allocated a channel. Algorithm 1 presents the pseudo-code of our *Channel-Select* allocation strategy: here, lines 6 and 8 make the diagonal and non-diagonal component of matrix \mathbf{A} , respectively, higher and lower, which forces the maximum $D(\mathbf{A}_l)$ to be as small as possible, as required by the objective function in (5). Figure 4 compares our *Channel-Select* strategy with random select and optimal solution for different number of MS's. As we expect, for a large number of MSs the possible combination of MSs increases and, as a result, the average $D(\mathbf{A})$ of our *Channel-Select* strategy and optimal solution decreases. So, unlike in random select, *Channel-Select* is able to increase the SNR (by decreasing the negative effect of background noise) of estimated source signals by exploiting diagonal dominant mixing channel coefficient matrices. From Fig. 3(a), we see that the difference between *Channel-Select* and optimal is not significant as the increase in BER is in order of 10^{-4} .

It should be mentioned that our solution, *Cloud-BSS*, is transparent to the MSs. The only over-the-air signaling that it requires is the one for conveying the channel allocation decision (made by *Channel-Select*) to the MSs; and this uplink-channel-allocation signaling is already part of current as well as next-generation OFDMA-based cellular systems.

D. Cluster Size and Computational Complexity

One of the main requisites for LTE is the requirement of very low level of latency. So, it is necessary to explore the computational complexity and run time of our *Cloud-BSS* solution for different cluster sizes. The cluster size dictates the computational complexity and the accuracy of the ICA methods, where the former depends on the no. of sources (N), no. of mixtures (M), no. of data samples (T), no. of iterations (Q), and no. of sweeps²(I). As the cluster size, i.e., the no. of RRHs (mixtures) and MS signals (sources), increase the computational complexity increases. The computational complexity of an algorithm is measured by the required floating point operations (flops) to execute it, where a flop corresponds to a multiplication followed by an addition.

Here, we briefly compare the complexity and accuracy of three well-known complex-valued ICA algorithms: JADE [9],

²A sweep is an iteration process over all principal 2×2 submatrices.

Algorithm 1 *Channel-Select* Strategy

Input: L = Total number of frequency channels available, M = Number of cells in a cluster

Output: MS_l = Set of MSs associated with the l^{th} channel

Description:

```

1: for  $l = 1; l \leq L; l++$  do
2:    $t = 1; K = l \bmod M;$ 
3:   if  $K = 0$  then
4:      $K = M;$ 
5:   end if
6:    $MS_l(t)$ =Find in cell # $K$  the nearest MS to RRH # $K$ ;
7:   for  $j = 1 : M$  &&  $j! = K$  do
8:      $MS_l(+t)$ =Find in cell # $j$  the farthest MS to RRH # $K$ ;
9:   end for
10:  Among  $MS_l(2 : L)$  keep the  $N - 1$  farthest MSs and remove the others, reducing the size of  $MS_l$  from  $M$  to  $N$ ;
11:  Allocate  $f_l$  to the remaining  $MS_l$ s;
12:  if All MSs have been allocated with a channel then
13:    return;
14:  end if
15: end for

```

Complex ICA-EBM [10], and Complex FastICA [11]. Complex ICA-EBM adopts a line-search optimization procedure using a projected conjugate gradient, while Complex FastICA finds independent components by separately maximizing the negentropy of each mixture. To compare the separation quality of these algorithms, we use the Performance Index (PI), which measures the difference between the mixing and estimated separating matrix, defined as,

$$PI = \sum_{i=1}^N \left[\left(\sum_{k=1}^N \frac{|p_{ik}|^2}{\max_j |p_{ij}|^2} - 1 \right) + \left(\sum_{k=1}^N \frac{|p_{ki}|^2}{\max_j |p_{ji}|^2} - 1 \right) \right], \quad (6)$$

where p_{ij} is the $(i, j)^{th}$ element of the matrix $\mathbf{P} = \mathbf{B}\mathbf{A}$, and $\max_j |p_{ij}|$ and $\max_j |p_{ji}|$ are the maximum absolute value in the i^{th} row and column of matrix \mathbf{P} , respectively. As the PI increases, the difference between \mathbf{B} and \mathbf{A}^{-1} increases, so the separation quality is correspondingly poorer. If the

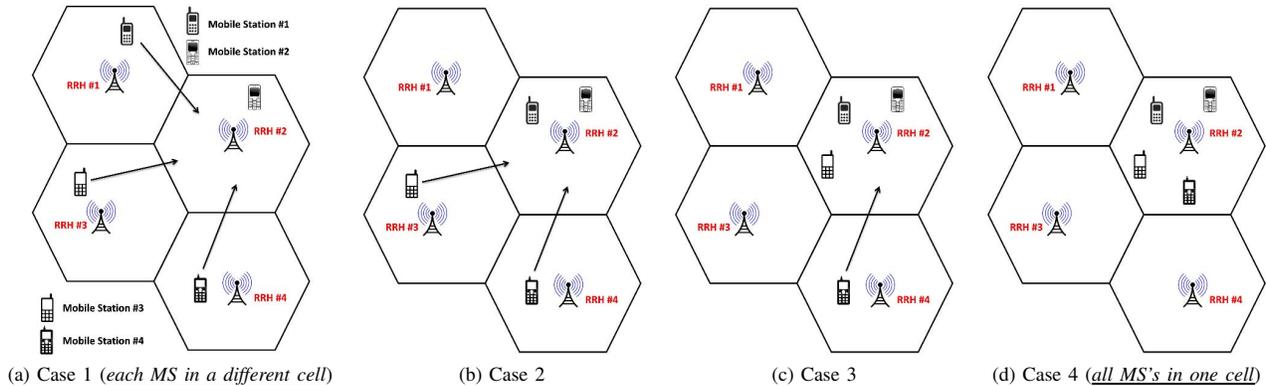


Fig. 5: **Four cases** of mobile network in which all the MS's are using the same frequency channel.

separation is perfect, then the separating matrix is the inverse of the mixing matrix and the PI is zero. Table I lists the computational complexity, separation quality, and run time³ of the aforementioned ICA algorithmic methods for different numbers of data samples T . As we can see, the PI decreases as T increases from 2000 to 8000, although at the price of the Run Time (RT): this is because with more data samples more information about the statistics of the mixtures is available, which enables the ICA algorithms to estimate the separating matrix more accurately. From Table I it is clear that the execution time of our *Cloud-BSS* solution using JADE algorithm is in order of millisecond, which is feasible in LTE systems.

Moreover, Fig. 6 shows the increase in RT (left y-axis) and PI (right y-axis) with the increase in the number of sources and mixtures (here $N = M$). To sum up, increasing the cluster size leads to higher computational complexity and lower accuracy; however, it also brings a few advantages, which we study in the following section, calling for finding the right trade-off.

III. OTHER BENEFITS OF OUR PROPOSED CLOUD-BSS

Besides improving the system spectral efficiency, using BSS within C-RAN brings several other advantages, some of which are briefly discussed below.

TABLE II: Reduction of handovers by clustering the cells using our BSS-based solution; cell radius = 1 km, simulation area = 30×30 km², $s_{min} = 0$, $s_{max} = 30$ m/s, simulation time = 1 hr, no. of MSs = 1000, $\bar{d} = \pi$, $\bar{s} = 15$ m/s, $d_{x_{n-1}} \sim \mathcal{N}(\pi, 1)$, $s_{x_{n-1}} \sim \mathcal{N}(15, 3)$, no. of simulations = 100.

Mobility Model	Number of Handovers			
	Without Clustering	Cells/Cluster = 3	Cells/Cluster = 4	Cells/Cluster = 5
Random Waypoint	5716 \pm 5%	2318 \pm 4%	1268 \pm 5%	843 \pm 6%
Gauss-Markov	3673 \pm 0.6%	1682 \pm 1.1%	711 \pm 2.3%	457 \pm 2.9%

Fewer Handovers: As long as the active users stay in the same cluster, there is no need to perform costly handovers because when a MS moves from one cell to another all the

³The experiments were performed on an Intel Core2 Quad CPU 2.4-GHz PC with 8 GB of RAM.

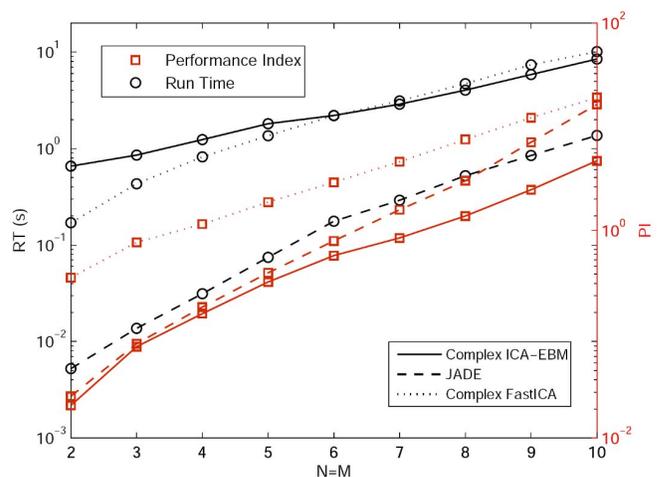


Fig. 6: Increase in Run Time (RT) and Performance Index (PI) of the three ICA algorithms with increase in the no. of sources and mixtures (here $T=8000$). The left y-axis reports the RT in seconds, while the right y-axis reports the PI as in (6).

RRHs within the cluster are still able to receive the mixtures of the transmitted signals. The number of handovers can be further reduced by increasing the size of the clusters. However, with an increase in both N and M , the complexity of the ICA methods and of the frequency-allocation algorithm, as well as the noise level in the estimated sources, will also increase.

To show the performance of our proposed solution in terms of number of handover sessions, we consider two mobility models: 1) Random Waypoint and 2) Gauss-Markov [14]. In the first model, a MS moves from its current location to a new one by choosing randomly a direction d [rd] and a speed s [m/s] from pre-defined ranges, e.g., $[0, 2\pi]$ and $[s_{min}, s_{max}]$, respectively. After choosing these parameters, a MS moves to its new location by traveling for a certain time or distance. The model also includes pause time between changes in direction and speed. The second mobility model is designed to adapt to different levels of randomness by means of a tuning parameter: the direction and speed at the n^{th} step are calculated based

on those at the $(n-1)^{th}$ step and on a random variable, as,

$$\begin{aligned} d_n &= \alpha d_{n-1} + (1-\alpha)\bar{d} + \sqrt{(1-\alpha^2)}d_{x_{n-1}} \\ s_n &= \alpha s_{n-1} + (1-\alpha)\bar{s} + \sqrt{(1-\alpha^2)}s_{x_{n-1}}, \end{aligned} \quad (7)$$

where d_n and s_n are the new direction and speed for the n^{th} step, α ($0 \leq \alpha \leq 1$) is the tuning parameter to vary the randomness, \bar{d} and \bar{s} are constants representing the mean value of direction and speed as $n \rightarrow \infty$, and $d_{x_{n-1}}$ and $s_{x_{n-1}}$ are random variables from a Gaussian distribution.

Table II represents the reduction in the number of handover sessions using *Cloud-BSS*. In the simulations, we performed an handover to the neighboring cell/cluster if both of the following conditions are met [15]: **(1)** If the signal strength from the neighboring cell/cluster exceeds that of the serving cell/cluster by an hysteresis (i.e., margin) level of at least 1 dB; **(2)** If the distance from the serving cell/cluster exceeds that of the neighboring cell/cluster by more than 1.1 km. It is clear that the number of handover sessions decreases with the increase of the cluster size. However, as we show in Figs. 3(a) and 6, the complexity of the ICA algorithms and the noise level of the estimated MS signals also increase.

Furthermore, we show how robust our solution is in terms of mobility and handover. As we observed previously, the achievable BER is highly dependent on the topology; so, in order to improve the BER performance, we introduced *Channel-Select*, a frequency-channel allocation heuristic. Now, we study how robust our solution is under user mobility without considering *Channel-Select*. To do this, we consider clusters of four cells (cell radius = 3 km), as shown in Fig. 5, and study four cases: in the first, all the MSs are in different cells while in the forth all the MS's have moved into cell #2. As we can see, in the first case, the topology is diagonal dominant and we expect an acceptable BER. However, from case #1 to case #4, $D(\mathbf{A})$ increases and we expect a decrease in the BER performance. Figure 7(a) shows the BER performance for these four cases, which corroborates our analysis.

Increased Reliability: We can trade capacity for improved BER performance. As mentioned earlier, we can assign each frequency channel to a maximum of as many as the total number of RRHs (recall that ICA algorithms require the no. of mixtures to be equal or greater than the no. of sources, i.e., $M \geq N$). The relationship between the *Net Channel Capacity* C_n of our proposed *Cloud-BSS* solution and the no. of admitted MSs N per frequency channel is,

$$C_n = N \cdot C_{ch} \cdot [1 - BER(N)], \quad (8)$$

where C_{ch} is the capacity of the frequency channel and $BER(N)$ indicates that the BER depends on N , as studied in Fig. 3(a). Figure 7(b) shows the increase of the net channel capacity with respect to the number of admitted MS's for the SNR ranging in [5, 20] dB and $D(\mathbf{A}) = 0.3$.

When the capacity is not the key issue, we can improve the BER by not using all the potential capacity of the network, i.e., by allocating each channel to fewer MSs so to induce diagonally dominant mixing matrices. This would lead to the reduction of the columns of matrix \mathbf{A} and to a higher degree

of freedom in making the mixing matrix diagonally dominant. To illustrate this intuition, we consider Topology 2 in Fig. 2(b), which is not diagonally dominant. However, we can make it so by allocating the channel to only two users instead of three, e.g., we can either remove MS #2 or MS #3. Figure 7(c) shows the BER performance in such scenario: interestingly, by removing MS #3 the performance is better than removing MS #2 as we obtain a lower $D(\mathbf{A})$, leading to a lower BER.

Interference Cancellation: Due to the orthogonality of subcarriers in LTE systems, the MSs have immunity to intra-cell interference. However, cell-edge users are known to face large ICI, especially, in a highly-loaded cellular environment. In *Cloud-BSS*, as BSS deals with mixed signals, intra-cluster interference is not a concern as it is a part of the received mixed signals and, as such, the ICA algorithms take care of it when separating the sources. Hence, there is no need for the BSs to go through costly inter-BS message exchange to execute coordinated interference cancellation. In fact, by applying the ICA algorithm to the received signals in a cluster, harmful interference can be turned into useful signal, which boosts performance by transforming a 'foe' into a 'friend'. Moreover, the *Channel-Select* strategy decreases the defective impact of inter-cluster interference (background noise) on the estimated signals by exploiting the diagonal dominant mixing channel coefficient matrices. With the increase in the cluster size, the average Signal to Interference plus Noise Ratio (SINR) also increases, which allows the cellular network to enjoy great spectral efficiency enhancement. Table III represents the improvement of spectral efficiency and throughput using *Cloud-BSS*. With a cluster size equal to 5, *Cloud-BSS* enhances the uplink average cell user and cell-edge user throughput by 49% and 93%, respectively.

TABLE III: Comparison of uplink Spectral Efficiency (SE) and Throughput (R) between non-cooperative traditional systems and *Cloud-BSS* (cell radius = 1 km).

Method	Average Cell User		Cell Edge User	
	SE [bps/Hz]	R [Mbps]	SE [bps/Hz]	R [Mbps]
Traditional System (Non-Cooperative)	2.17 ± 7.3%	10.85	0.73 ± 6.8%	3.65
<i>Cloud-BSS</i> ($M = 3$)	2.76 ± 6.8%	13.79	1.12 ± 6.2%	5.60
<i>Cloud-BSS</i> ($M = 4$)	3.07 ± 7.1%	15.35	1.27 ± 6.3%	6.35
<i>Cloud-BSS</i> ($M = 5$)	3.23 ± 7.1%	16.14	1.41 ± 7.8%	7.05

No Additional Overhead Compared to CoMP: Coordinated Multipoint Processing (CoMP) is another approach to mitigate the average interference, in which the BSs within a cluster exchange Channel State Information (CSI) [16]. However, this approach requires a pilot-symbol overhead (19% of the capacity) in order to estimate the channel coefficients [17]. Releasing this huge amount of capacity can increase the net bit rate, which may be used for other purposes like coding to increase reliability. Our BSS-based solution implicitly estimates the channel coefficients and therefore does not require pilot-data exchange as in CoMP.

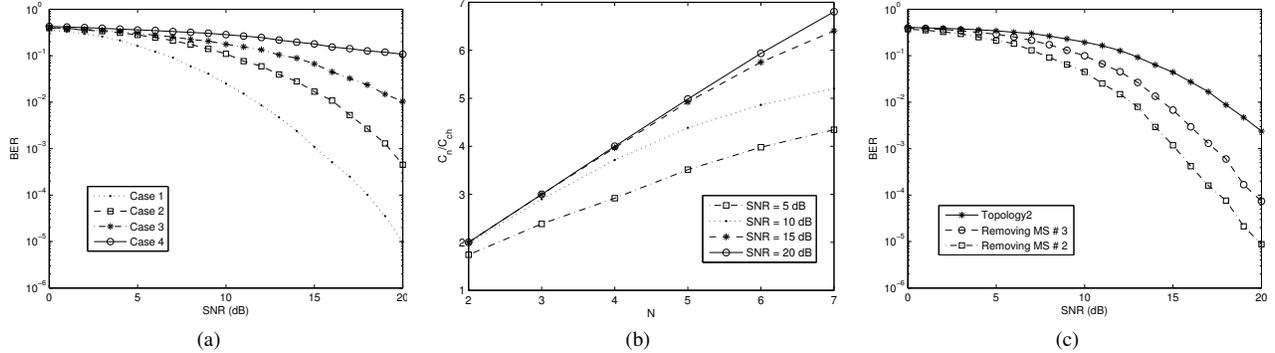


Fig. 7: (a) Increase of Bit Error Rate (BER) with Signal to Noise Ratio (SNR) by moving the MS's towards a common cell; (b) Increase in the Net Channel Capacity with increase in N , with $M = 7$, for different $SNR = 5, 10, 15, 20$ dB and $D(\mathbf{A}) = 0.3$; (c) Trading off capacity for improved BER: decrease of BER with SNR when removing MS # 2 or MS # 3 from Fig. 2(b).

IV. CONCLUSION AND FUTURE WORK

We presented a novel Blind Source Separation (BSS)-based solution, *Cloud-BSS*, that leverages the centralized characteristic of Cloud Radio Access Network (C-RAN) so to improve performance of highly mobile cellular networks. *Cloud-BSS* divides a set of neighboring cells into clusters that can use all of the frequency channels in the system band, thus increasing the system spectral efficiency, decreasing handovers, and eliminating the need for bandwidth-consuming channel estimation while mitigating interference. We discussed the effect of irregular topologies on *Cloud-BSS* performance in terms of BER and introduced a strategy, named *Channel-Select*, to improve the SNR.

Although *Cloud-BSS* is able to mitigate the intra-cluster interference, it can not mitigate the inter-cluster interference and we only are able to decrease the effect of inter-cluster interference on estimated source signals by exploiting the *Channel-Select* strategy. As a result, the achieved system capacity – while improved – is still not close to the interference-free-capacity upper bound, especially in environments with strong Co-Channel Interference (CCI). Therefore, in order to mitigate the inter-cluster interference and approach such upper bound, we are currently exploring the possibility to define different clusters for different frequency channels and to dynamically determine clusters' sizes based on the current and future positions of the mobile users.

Acknowledgments: This work was supported in part by the National Science Foundation grant no. CNS-1319945.

APPENDIX PROOF OF THEOREM 1

Since $\|\mathbf{C}^{-1}\|_{\infty}^{-1} = \inf_x \frac{\|\mathbf{C}\mathbf{x}\|_{\infty}}{\|\mathbf{x}\|_{\infty}}$, we only need to show that $\alpha\|\mathbf{x}\|_{\infty} \leq \|\mathbf{C}\mathbf{x}\|_{\infty}$ for all \mathbf{x} . Take some vector \mathbf{x} and let $\|x\|_{\infty} = |x_i|$; then, it follows that $0 < \alpha \leq |c_{ii}| - \sum_{j \neq i} |c_{ij}|$.

With some manipulation, we have,

$$0 < \alpha |x_i| \leq |c_{ii}x_i| - \sum_{j \neq i} |c_{ij}x_j| \leq |c_{ii}x_i| - \left| \sum_{j \neq i} c_{ij}x_j \right| \leq \left| \sum_j c_{ij}x_j \right| \leq \max_j \left| \sum_j c_{ij}x_j \right| = \|\mathbf{C}\mathbf{x}\|_{\infty}.$$

REFERENCES

- [1] D. Astély, E. Dahlman, A. Furuskar, Y. Jading, M. Lindstrom, and S. Parkvall, "Lte: The evolution of mobile broadband," *IEEE Communications Magazine*, vol. 47, no. 4, pp. 44–51, Apr. 2009.
- [2] C. M. R. Institute, "C-RAN: The Road Towards Green RAN," Oct. 2011.
- [3] A. Sang, X. Wang, M. Madhian, and R. D. Gitlin, "Coordinated load balancing, handoff/cell-site selection, and scheduling in multi-cell packet data systems," *Wireless Networks*, vol. 14, no. 1, pp. 103–120, 2008.
- [4] K. Gomadam, V. Cadambe, and S. Jafar, "Approaching the Capacity of Wireless Networks through Distributed Interference Alignment," in *Proc. of IEEE Global Telecommunications Conf. (GLOBECOM)*, Dec. 2008.
- [5] P. Gupta, A. Vishwanath, S. Kalyanaraman, and Y. Lin, "Unlocking Wireless Performance with Co-operation in Co-located Base Station Pools," in *Proc. of the Intl. Conf. on Communication Systems and Networks (COMSNETS)*, Jan. 2010.
- [6] Y. Lin, L. Shao, Z. Zhu, Q. Wang, and R. Sabhikhi, "Wireless Network Cloud: Architecture and System Requirements," *IBM Journal of Research and Development*, vol. 54, no. 1, pp. 1–4, 2010.
- [7] M. Madhavan, P. Gupta, and M. Chetlur, "On Quantifying Multiplexing Gains in a Wireless Network Cloud," Jun. 2012.
- [8] A. Hyvärinen and E. Oja, "Independent component analysis: Algorithms and applications," *Neural networks*, vol. 13, no. 4, pp. 411–430, 2000.
- [9] J. Cardoso and A. Souloumiac, "Blind beamforming for non-gaussian signals," in *Radar and Signal Processing, IEE Proceedings F*, vol. 140, no. 6, 1993, pp. 362–370.
- [10] X. Li and T. Adali, "Complex independent component analysis by entropy bound minimization," *IEEE Transactions on Circuits and Systems I: Regular Papers*, vol. 57, no. 7, pp. 1417–1430, 2010.
- [11] E. Bingham and A. Hyvärinen, "A fast fixed-point algorithm for independent component analysis of complex valued signals," *Intl. Journal of Neural Systems*, vol. 10, no. 01, pp. 1–8, 2000.
- [12] R. Varga, "On diagonal dominance arguments for bounding," *Linear Algebra and its Applications*, vol. 14, no. 3, pp. 211–217, 1976.
- [13] J. Varah, "A lower bound for the smallest singular value of a matrix," *Linear Algebra and Its Applications*, vol. 11, no. 1, pp. 3–5, 1975.
- [14] T. Camp, J. Boleng, and V. Davies, "A survey of mobility models for ad hoc network research," *Wireless communications and mobile computing*, vol. 2, no. 5, pp. 483–502, Aug. 2002.
- [15] K. Itoh, S. Watanabe, J. Shih, and T. Sato, "Performance of handoff algorithm based on distance and rssi measurements," *IEEE Transactions on Vehicular Technology*, vol. 51, no. 6, pp. 1460–1468, Nov. 2002.
- [16] D. Lee, H. Seo, B. Clerckx, E. Hardouin, D. Mazzarese, S. Nagata, and K. Sayana, "Coordinated multipoint transmission and reception in lte-advanced: deployment scenarios and operational challenges," *Communications Magazine, IEEE*, vol. 50, no. 2, pp. 148–155, 2012.
- [17] M. Simko, D. Wu, C. Mehlhüner, J. Eilert, and D. Liu, "Implementation aspects of channel estimation for 3gpp lte terminals," in *European Wireless Conf. on Sustainable Wireless Technologies (European Wireless)*, 2011, pp. 1–5.