Cell-edge Interferometry: Reliable Detection of Unknown Cell-edge Users via Canonical Correlation Analysis

Mohamed Salah Ibrahim and Nicholas D. Sidiropoulos
Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA
Email: {mi6cw,nikos@virginia.edu}

***Abstract*—A key challenge in 4G and emerging 5G systems is that of reliably detecting the uplink transmissions of users close to the edge between cells. These users are subject to significant signal attenuation due to path loss, and frequent hand-off from one cell to the other, making channel estimation very challenging. Even multiuser detection using base station cooperation often fails to detect such users, due to channel estimation errors and the sensitivity of multiuser detection to near-far power imbalance. Is it even possible to reliably decode the cell-edge users' signals under these circumstances? This paper shows, perhaps surprisingly, that with a suitable base station 'interferometry' strategy, the cell-edge users' signals can be reliably decoded at low SNR under mild conditions. Exploiting the fact that cell-edge users' signals are weak but *common* to both base stations, while users close to a base station are unique to that base station, reliable detection is enabled by Canonical Correlation Analysis (CCA) – a machine learning technique that reliably estimates a common subspace, even in the presence of strong individual interference. Free from cell-center interference, the resulting mixture of cell-edge signals can then be unraveled using well-known algebraic signal processing techniques. Simulations demonstrate that the proposed detector achieves order of magnitude BER improvement compared to an 'oracle' zero-forcing with successive interference cancellation that assumes perfect knowledge of all channels. The paper also includes proof of common subspace identifiability for the assumed generative model, which was curiously missing from the machine learning / CCA literature.**

I. Introduction

Multi-user (MU) MIMO detection for uplink reception aims at demodulating multiple users' signals in the presence of multi-access interference [1], [2]. The optimum maximum likelihood detector (MLD) requires solving an NP-Hard combinatorial problem of complexity that grows exponentially with the number of users [1]. Sphere decoding offers MLD performance at lower complexity, especially in the high signal to noise regime [3], however its complexity is still exponential [4]. Minimum mean square error (MMSE) [5] and zero-forcing (ZF) are low-complexity linear detectors, whose performance can be improved by successively removing the strong users' signals once they are decoded – a technique referred to as decision feedback (DF) or successive interference cancellation (SIC) [6]. All of the aforementioned detectors have been successful in many applications as long as accurate channel estimates are available, and power control is used to mitigate near-far power imbalances.

Supported in part by NSF ECCS-1807660.

In cellular systems, accurate channel estimation can be accomplished for cell-center (strong) users, however, cell-edge (weak) user signals are received at low signal-to-noise ratio (SNR) because attenuation follows an inverse power law as a function of distance. This also degrades their channel estimates [7], [8], and hence, their detection performance. Whereas power control can be used to mitigate path loss [8], [26], this comes at the expense of significantly reducing the rate of the users that are close to the base station. Cell-edge users are also subject to frequent hand-offs, which further aggrevate the situation.

The situation begs the question whether it is possible to reliably decode weak cell-edge users' signals under such circumstances? This paper shows that with a suitable base station 'interferometry' strategy inspired from machine learning, together with a well-known algebraic signal processing tool, the cell-edge users' signals can be reliably decoded under mild conditions, even at low SNR and when buried under heavy cell-specific interference from the cell-center users. Exploiting the fact that cell-edge users' signals are weak but *common* to both base stations, while users close to a base station are unique to that base station, reliable detection is enabled by Canonical Correlation Analysis (CCA) [9], [10] – a machine learning technique that reliably estimates a *common subspace* using eigendecomposition, even in the presence of strong individual-cell interference. We show that CCA yields the correct subspace containing the cell-edge users' signals. Free from cell-center interference, the resulting mixture of cell-edge signals can then be unraveled by exploiting their finite-alphabet structure using RACMA [11]. The paper includes an algebraic proof of the main claim in the case where thermal noise and adjacent-cell interference from non-cell-edge users can be neglected (i.e., users close to base station B are not overhead at base station A).

With a realistic path-loss model (that includes interference from all users) the proposed detector achieves order of magnitude BER improvement compared to an 'oracle' cooperative zero-forcing with successive interference cancellation that assumes perfect knowledge of all channels. Intuitively, this is because the concept of common versus individual signals translates to equipowered versus imbalanced signals in that case. The more the power imbalance, the more a given signal tends to 'belong' to one base station (and hence become 'individual'). The idealized case considered in our proof emerges in the limit.

Our approach is very different from multiuser detection using base station cooperation [12], as it capitalizes on CCA. In the signal processing literature, CCA has been used in direction-of-arrival (DoA) estimation [13], equalization [14], [15], array processing [16], blind source separation [17]–[19], and multi-view learning [20], to name a few applications; but not anywhere close to our present context. Scalable algorithms for generalized (multi-view) CCA were recently developed by the authors' group [21]–[23], also incorporating various constraints. From a methodological point of view, the key contribution of this paper is the proof of common subspace identifiability for the assumed generative model, which was – surprisingly – missing from the machine learning / CCA literature. From the application point of view, the main contribution is to showcase the power of CCA in solving a practical engineering problem that seems insurmountable otherwise.

Matrices (vectors) are denoted by upper- (lower-) case boldface letters, and $(\cdot)^T$ transpose. Scalars are represented in the normal face. $\|.\|_2$ and $\|.\|_F$ denote the ℓ_2-norm and the frobenius norm, respectively. Finally, $\mathbf{I}_N$ and $\mathbf{0}_{N\times M}$ denote the $N \times N$ identity matrix and the $N \times M$ zero matrix, respectively.

II. Preliminaries

A. System Model

Consider an uplink wireless transmission in a MU-MIMO system with two cells, as shown in Figure 1. The base station (BS) in each cell is equipped with M antennas and serves K single-antenna users. Our approach works for different K for each base station, but we assume common K for simplicity of exposition. Let $K_c = K_{e_1} + K_{e_2}$ denotes the total number of cell-edge users located around the common edge between the two cells, where $K_{e_\ell} < K$ represents the number of cell-edge users served by the ℓ-th BS, for $\ell \in \{1, 2\}$. Let $\mathbf{s}_{kj} \in \mathbb{R}^{T\times 1}$ be the vector containing symbols transmitted by the k-th user in cell j, where each entry of $\mathbf{s}_{kj}$ belongs to the finite alphabet $\Omega = \{\pm 1\}$ (our approach works for general PSK and other alphabets, with some changes in the second stage). The received signal, $\mathbf{Y}_\ell \in \mathbb{C}^{M\times T}$, at the ℓ-th BS can be expressed as

$$\mathbf{Y}_\ell = \sum_{j=1}^{2}\sum_{k=1}^{K} \sqrt{\beta}\mathbf{h}_{\ell kj}\mathbf{s}_{kj}^T + \mathbf{W}_\ell \tag{1}$$

where $\mathbf{W}_\ell \in \mathbb{C}^{M\times T}$ contains independent identically distributed (i.i.d) entries with each element drawn from a complex Gaussian distribution with zero mean and variance σ^2, and β models the transmitted signal power. The term $\mathbf{h}_{\ell kj}$ models independent small scale fading and path-loss attenuation between the k-th user in the j-th cell and the ℓ-th BS, and is given by

$$\mathbf{h}_{\ell kj} = \sqrt{\alpha_{\ell kj}}\mathbf{g}_{\ell kj} \tag{2}$$

where $\mathbf{g}_{\ell kj} \in \mathbb{C}^{M\times 1}$ represents the small scale fading between user k in cell j and BS ℓ. In our simulations, its entries

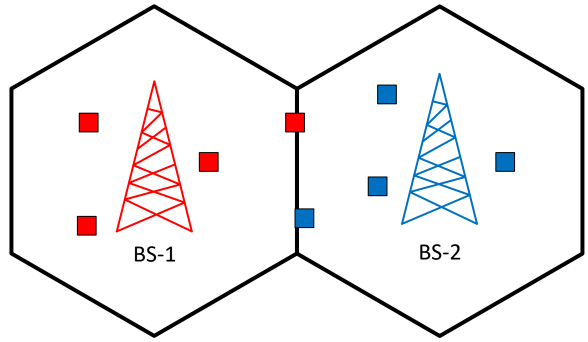

Fig. 1: System Model

are assumed to be i.i.d complex zero-mean Gaussian random variables. On the other hand, $\alpha_{\ell kj} \in \mathbb{C}$ is the large scale fading coefficient that accounts for the channel attenuation (path-loss) between user k in cell j and BS ℓ. It is assumed that $\mathbf{h}_{\ell kj}$'s is not known *apriori* at the ℓ-th BS.

B. Cell-Edge Users Detection

Let us collect the signals of cell-edge users and cell-center users served by BS ℓ in the matrices $\mathbf{S}_c \in \mathbb{R}^{T\times K_c}$ and $\mathbf{S}_{p_\ell} \in \mathbb{R}^{T\times(K-K_{e_\ell})}$, respectively, for $\ell \in \{1, 2\}$. Furthermore, let $\tilde{\mathbf{W}}_\ell$ represent the noise at the ℓ-th BS plus the interference caused by the cell-center users in cell j, where $j \neq \ell$. Therefore, (2) can be expressed as follows

$$\mathbf{Y}_\ell = \mathbf{H}_{\ell p_\ell}\mathbf{S}_{p_\ell}^T + \mathbf{H}_{\ell c}\mathbf{S}_c^T + \tilde{\mathbf{W}}_\ell \tag{3}$$

where the matrices $\mathbf{H}_{\ell c} \in \mathbb{C}^{M\times K_c}$ and $\mathbf{H}_{\ell p_\ell} \in \mathbb{C}^{M\times(K-K_{e_\ell})}$ hold on their columns all the channel vectors from cell-edge users to the ℓ-th BS, and the channel vectors from cell-center users to their serving BS, respectively. Note that the subscripts c and p stand for 'common' and 'private', respectively. In addition, absorb the transmitted signal power, β, of each user in the respective channel vectors.

One possible approach to detect cell-edge users' signals is to apply zero-forcing successive interference cancellation (ZF-SIC) [6], which is based on successively removing the cell-center (strong) users' signals once they are detected using ZF. Afterwards, the approach applies ZF to detect the cell-edge (weak) users. However, ZF requires accurate channel estimates to provide reliable detection performance. Although this can be guaranteed for cell-center users, cell-edge users' signals are received intermittently at low SNR which results in poor channel estimates [7], [8]. In what follows, we present a novel detector that can reliably decode cell-edge users' signals at low received SNR and strong interference, without the availability of CSI.

III. Cell-Edge Users Detection Using CCA

In this section, the signals received by both base stations are jointly processed by a central signal processing unit. The goal of the proposed detector is to detect the cell-edge users' signals $\mathbf{S}_c$ from the received signals $\mathbf{Y}_1$ and $\mathbf{Y}_2$. First, the signals are transformed to the real domain by forming the matrix $\overline{\mathbf{Y}}_\ell := [\mathbf{Y}_\ell^{(r)}; \mathbf{Y}_\ell^{(i)}] \in \mathbb{R}^{2M\times T}$, where $\mathbf{Y}_\ell^{(r)} = \mathbb{R}\mathrm{e}\{\mathbf{Y}_\ell\}$ and $\mathbf{Y}_\ell^{(i)} = \mathbb{I}\mathrm{m}\{\mathbf{Y}_l\}$ represent the real and imaginary components of the ℓ-th BS signal. Similarly, denote by $\overline{\mathbf{A}}_{\ell p_\ell} := [\mathbf{H}^{(r)}_{\ell p_\ell}; \mathbf{H}^{(i)}_{\ell p_\ell}] \in \mathbb{R}^{2M\times(K-K_{e_\ell})}$, $\overline{\mathbf{A}}_{\ell c} := [\mathbf{H}^{(r)}_{\ell c}; \mathbf{H}^{(i)}_{lc}] \in \mathbb{R}^{2M\times K_c}$ and $\overline{\mathbf{W}} = [\tilde{\mathbf{W}}^{(r)}_l; \tilde{\mathbf{W}}^{(i)}_l] \in \mathbb{R}^{2M\times T}$. Therefore, (3) can be equivalently written as

$$\overline{\mathbf{Y}}_\ell = \overline{\mathbf{A}}_{\ell p_\ell}\mathbf{S}^T_{p_\ell} + \overline{\mathbf{A}}_{\ell c}\mathbf{S}^T_c + \overline{\mathbf{W}}_\ell. \tag{4}$$

Next, the two-view CCA formulation [10] is utilized to estimate the subspace containing the cell-edge users' signals. For the sake of brevity, we refer to this subspace as the *common subspace*. The two-view CCA seeks to find a common structure of the views through a linear transformation of the data received at each view (BS) [24]. In other words, it finds two matrices $\mathbf{Q}_1 \in \mathbb{R}^{2M\times N}$ and $\mathbf{Q}_2 \in \mathbb{R}^{2M\times N}$, with $N < \min\{M, T\}$, such that the correlation between the projections of $\overline{\mathbf{Y}}_1$ and $\overline{\mathbf{Y}}_2$ onto these matrices is maximized. In an optimization framework, this can be mathematically expressed as [25]

$$\min_{\mathbf{Q}_1,\mathbf{Q}_2} \|\overline{\mathbf{Y}}^T_1\mathbf{Q}_1 - \overline{\mathbf{Y}}^T_2\mathbf{Q}_2\|^2_F \tag{5a}$$

$$\text{s.t.} \quad \mathbf{Q}^T_\ell\overline{\mathbf{Y}}_\ell\overline{\mathbf{Y}}^T_\ell\mathbf{Q}_\ell = \boldsymbol{I}, \quad \ell \in \{1, 2\} \tag{5b}$$

Note that the n-th column of $\mathbf{Q}_\ell$ represents the n-th canonical component of view $\overline{\mathbf{Y}}_\ell$, for $n \in \{1, \cdots, N\}$, and the correlation coefficient ρ_n of the n-th pair of columns (one from each matrix) is computed. The number of components (pairs) extracted, (N), depends on the minimum value of ρ that needs to be considered. Problem (5) can be optimally solved using generalized eigenvalue decomposition [9], [25]. An alternative formulation of (5) is to search for an orthogonal representation $\mathbf{G} \in \mathbb{R}^{T\times N}$ that is maximally correlated after the linear projections of $\overline{\mathbf{Y}}_1$ and $\overline{\mathbf{Y}}_2$ on $\mathbf{Q}_1$ and $\mathbf{Q}_2$, respectively. This can be written as

$$\min_{\mathbf{Q}_1,\mathbf{Q}_2,\mathbf{G}} \sum_{\ell=1}^{2} \|\overline{\mathbf{Y}}^T_\ell\mathbf{Q}_\ell - \mathbf{G}\|^2_F \tag{6a}$$

$$\text{s.t.} \quad \mathbf{G}^T\mathbf{G} = \boldsymbol{I} \tag{6b}$$

Problem (6) is known as the MAX-VAR formulation of the CCA [24] and, in the case of two views considered here, it is equivalent to (5) in the sense that both problems yield the same solutions $\mathbf{Q}^\star_\ell$. In this work, we focus on the formulation in (6) as it facilitates our proof.

Assume that we are interested in the first K_c canonical components of the matrices $\mathbf{Q}_1$ and $\mathbf{Q}_2$, i.e., $N = K_c$. We have the following result.

Theorem: In the noiseless case, if matrix $\mathbf{B} := [\mathbf{S}_c, \mathbf{S}_{p_1}, \mathbf{S}_{p_2}] \in \mathbb{R}^{T\times 2K}$ is full column rank, and $\overline{\mathbf{A}}_\ell = [\overline{\mathbf{A}}_{\ell c}, \overline{\mathbf{A}}_{\ell p_\ell}] \in \mathbb{R}^{2M\times(K_c+K-K_{e_\ell})}$ is a full column rank for $\ell \in \{1, 2\}$, then the optimal solution $\mathbf{G}^\star$ of problem (6) is given by $\mathbf{G}^\star = \mathbf{S}_c\mathbf{P}$, where $\mathbf{P}$ is a $K_c \times K_c$ non-singular matrix.

Remark. *The full column rank condition on $\mathbf{B}$ requires $T \geq 2K$, and the transmitted sequences from the different users to be linearly independent. For finite-alphabet signals, this happens with very high probability for moderate T, since the different user transmissions are independent. The more restrictive condition is full column rank of $\overline{\mathbf{A}}_\ell$, which relates the number of base station antennas and signals impinging on each base station. We thus need two times the number of antennas in each base station to be $\geq$ the number of users assigned to that base station, plus any cell-edge users assigned to the other base station. Other than this dimensionality constraint though, if the channel vectors are drawn from a jointly continuous distribution, the latter condition will be satisfied with probability one.*

Proof: First, let us start with the single cell-edge user case, i.e., $K_c = 1$ and each of $\mathbf{S}_c, \mathbf{G}$ and $\mathbf{Q}_\ell$ is a vector. In such setting (6) relaxes to the following

$$\min_{\mathbf{q}_1,\mathbf{q}_2,\mathbf{g}} \sum_{\ell=1}^{2} \|\overline{\mathbf{Y}}^T_\ell\mathbf{q}_\ell - \mathbf{g}\|^2_2 \tag{7a}$$

$$\text{s.t.} \quad \|\mathbf{g}\|^2_2 = 1 \tag{7b}$$

To solve the above problem, we need to find $(\mathbf{q}^\star_1, \mathbf{q}^\star_2, \mathbf{g}^\star)$ that can together attain a zero-cost. In other words, we need the following two conditions to be satisfied simultaneously

$$\overline{\mathbf{Y}}^T_1\mathbf{q}_1 = \mathbf{g} \tag{8a}$$

$$\overline{\mathbf{Y}}^T_2\mathbf{q}_2 = \mathbf{g} \tag{8b}$$

Without loss of generality, we can let $\mathbf{q}_\ell = \overline{\mathbf{A}}_\ell(\overline{\mathbf{A}}^T_\ell\overline{\mathbf{A}}_\ell)^{-1}\mathbf{u}_\ell$, where $\mathbf{u}_\ell$ is any vector in $\mathbb{R}^{K_c+K-K_{e_\ell}}$. The reason is that we can always decompose $\mathbf{q}_\ell$ into a component in the subspace spanned by $\overline{\mathbf{A}}_\ell$ and one orthogonal to it. The latter is annihilated anyway after multiplication with $\overline{\mathbf{A}}^T_\ell$. Substituting in (8a) and (8b) and taking their difference, we obtain

$$\mathbf{B}\mathbf{u} = \mathbf{0}, \tag{9}$$

where $\mathbf{B} = [\mathbf{s}_c, \mathbf{S}_{p_1}, \mathbf{S}_{p_2}] \in \mathbb{R}^{T\times 2K}$ and $\mathbf{u} = [\mathbf{u}_1(1) - \mathbf{u}_2(1), \mathbf{u}_1(2 : \text{end}), -\mathbf{u}_2(2 : \text{end})]^T \in \mathbb{R}^{2K}$, where $\mathbf{u}_1(2 : \text{end})$ is the vector containing all except the first element of $\mathbf{u}$. It can be easily seen that if $\mathbf{B}$ is full column rank, then $\mathbf{u} = \mathbf{0}_{2K\times 1}$ is the only possible solution of (9). This means that $\mathbf{u}_1 = \mathbf{u}_2 = c\mathbf{e}_1$, where c is any constant and $\mathbf{e}_1$ is the first column of the identity matrix. Consequently, from (8), $\mathbf{g}^\star = \alpha\mathbf{s}_c/\|\mathbf{s}_c\|_2$, with $\alpha = \pm 1$, will be the only possible solution for problem (7).

The generalization to $K_c > 1$ now follows naturally. Letting $\mathbf{Q}_\ell = \overline{\mathbf{A}}_\ell(\overline{\mathbf{A}}^T_\ell\overline{\mathbf{A}}_\ell)^{-1}\mathbf{U}_\ell$, and defining

$$\mathbf{U} := \begin{bmatrix} \mathbf{U}_1(1 : K_c, :) - \mathbf{U}_2(1 : K_c, :) \\ \mathbf{U}_1(K_c + 1 : \text{end}, :) \\ -\ \mathbf{U}_2(K_c + 1 : \text{end}, :) \end{bmatrix} \in \mathbb{R}^{2K\times K_c},$$

where $\mathbf{U}_1(1 : K_c, :)$ means rows 1 to K_c and all columns of $\mathbf{U}_1$, we obtain

$$\mathbf{B}\mathbf{U} = \mathbf{0}, \tag{10}$$

and when $\mathbf{B}$ is full column rank the solution is unique: $\mathbf{U} = \mathbf{0}$, and therefore $\mathbf{U}_1(1 : K_c, :) = \mathbf{U}_2(1 : K_c, :) =: \mathbf{P}$, $\mathbf{U}_1(K_c + 1 : \text{end}, :) = \mathbf{0}$, $\mathbf{U}_2(K_c + 1 : \text{end}, :) = \mathbf{0}$, and therefore $\mathbf{G}^\star = \mathbf{S}_c\mathbf{P}$, where $\mathbf{P}$ is $K_c \times K_c$ non-singular such that the orthonormality constraint (6b) is satisfied. Note that if the signals themselves are (approximately) orthogonal,

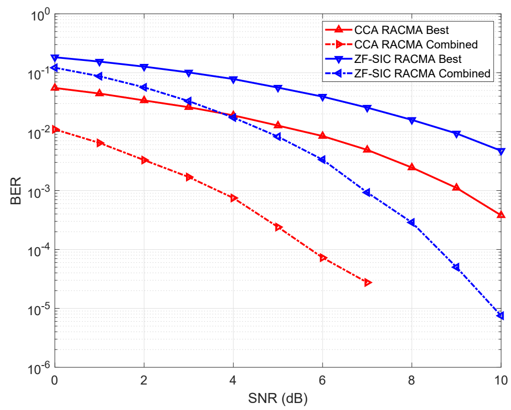

Fig. 2: BER vs. SNR of cell-edge users, with $M = 12$, $K = 10$ and $K_c = 2$, distance of cell-center users < 0.6R

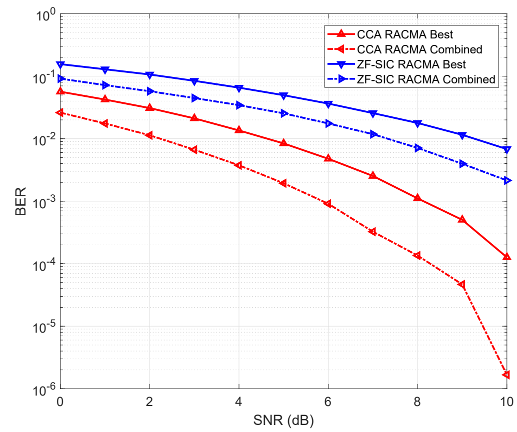

Fig. 3: BER vs. SNR of cell-edge users, with $M = 20$, $K = 15$ and $K_c = 3$, distance of cell-center users $< 0.75R$

then $\mathbf{P}$ will be orthogonal as well, which helps with the next (RACMA) stage. ■

The next step is to extract the cell-edge users' sequences $\mathbf{S}_c$ from $\mathbf{G}^\star = \mathbf{S}_c\mathbf{P}$. This problem can be viewed as a bilinear factorization of the matrix $\mathbf{G}^\star$ to its factors $\mathbf{P}$ and $\mathbf{S}_c$ under the constraint that the entries of $\mathbf{S}_c$ belong to the finite alphabet $\Omega = \pm 1$. This can be mathematically posed as an optimization problem as follows

$$\min_{\overline{\mathbf{S}}_c, \overline{\mathbf{P}}} \|\mathbf{G}^\star - \overline{\mathbf{S}}_c\overline{\mathbf{P}}\|_F^2 \tag{11a}$$

$$\text{s.t.} \quad \overline{\mathbf{S}}_c(i,j) \in \Omega \tag{11b}$$

In [11], van der Veen proposed an algebraic algorithm called Real Analytical Constant Modulus Algorithm (RACMA) for this problem. RACMA does not claim to optimally solve (11), which is NP-hard even if $\mathbf{P}$ is known. Instead, RACMA assumes that noise is small, and reduces (11) to a generalized eigenvalue problem. The solution is subject to sign and user permutation ambiguity This means that the original $\mathbf{S}_c$ can be identified up to permutations and columnwise (user) scaling by ± 1. From the practical point of view, each user has its unique identification sequence, so once the users' signals are received correctly each BS can identify each user's signal (and sign) via correlation with the identification sequence.

Remark. *It is important to emphasize that, in the noisy case and if the cell-center users in each cell are randomly dropped up to a certain distance from their serving BS, it turns out that our method can still identify the common subspace at low SNR values as we will see in the next section.*

IV. Experimental Results

To assess the performance of our proposed method, we consider a scenario with two hexagonal cells; each with radius $R = 500$ meter. Cell-edge users' locations were generated randomly by following a uniform distribution, however, they were confined to be around the common edge between the two cells, i.e., the locations of cell-edge users were chosen between $0.95R$ and $1.05R$. On the other hand, the distance between each cell-center user and its serving BS is at most $0.6R$. The transmitted power β was set to 25dBm for all users. Furthermore, the transmitted sequence's length T was fixed to 800. Additive white Gaussian noise is assumed with variance σ^2 so that the SNR is P_e/σ^2, where P_e is the average received signal received power of cell-edge users. In fact, this enables us to see what values of SNR should cell-edge users have to achieve a specific BER. Furthermore, all results were averaged over 500 Monte-Carlo trials. The uplink channels $\{\mathbf{h}_{lkj}^H\}$ are modeled as

$$\mathbf{h}_{lkj}^H = \sqrt{\frac{1}{M}} \sum_{n=1}^{L} \sqrt{\alpha_{lkj}^{(n)}} \mathbf{a}_r(\theta^{(n)})^H \tag{12}$$

where L is the number of paths between the l-th BS and the k-th user in cell j, $\forall \{l,j\} \in \{1,2\}$ and $k \in \{1, \cdots, K\}$. We use the path-loss model of the urban macro (UMa) scenario from 3GPP 38.901 standard to compute the complex path gain $\alpha_{lkj}^{(n)}$, $\forall n, l, j, k$. The term $\mathbf{a}_r(.)$ is the array response vector at the BS, and $\theta^{(n)} \sim \mathcal{U}[-\pi, \pi]$ denotes the azimuth angle of arrival of the n-th path. Assuming the BS is equipped with a uniform linear array, then

$$\mathbf{a}_r(\theta) = [1, \exp^{ikd\cos(\theta)}, \cdots, \exp^{ikd(M-1)\cos(\theta)}] \tag{13}$$

where $k = 2\pi/\lambda$, λ is the carrier wavelength and $d = \lambda/2$ is the spacing between antenna elements.

In order to benchmark the performance of our proposed method, we implemented the zero-forcing successive interference cancellation (ZF-SIC) where the channels of cell-center users were assumed to be perfectly known at their serving BSs. Note that after canceling the interference of cell-center users from the received signal, we pass on the residual signals to RACMA [11] to identify cell-edge users' signals. Afterwards, the bit error rate (BER) of cell-edge users was computed at both BSs and the best was reported. Furthermore, in order to guarantee fairness, since we have assumed that both BSs' signals are received at the processing center, both residual signals from both BSs were sent simultaneously to RACMA and the combined BER was reported. Similarly, due to the presence of noise and inter-cell interference that can affect the estimation of the common subspace, we sent $\overline{\mathbf{Y}}_1^T \mathbf{Q}_1$ and $\overline{\mathbf{Y}}_2^T \mathbf{Q}_2$ simultaneously to RACMA to solve for the cell-edge signals and compute the BER.

The performance of BER versus SNR of cell-edge users, for $M = 12$, $K = 10$ and $K_c = 2$, is depicted in Figure 2. It is obvious that our blind method achieves a considerable improvement in the BER compared to the ZF-SIC with full CSI for cell-center users. For instance, more than one order of magnitude improvement using our blind learning method was observed. Furthermore, Figure 2 shows that our approach attained a zero BER when the SNR of cell-edge users exceeds 7dB.

Moreover, we carried out another experiment with $M = 20, K = 15, K_c = 3$ and cell-center users were randomly located at distance less than $0.75R$ from their serving BS. Figure 3 shows that injecting more users and allowing them to be more spreaded, affected the BER of cell-edge users obtained by both methods. This makes sense because, for ZF-SIC, there exists a higher chance that the detection performance of some cell-center users will be affected by the interference of cell-edge users resulting in cancellation errors from SIC, while our method also exhibits some degradation in the performance because adding more users creates more inter-cell interference that can contaminate the common subspace estimated by CCA. However, our approach can still achieve much better performance to that obtained by ZF-SIC with perfect CSI. For example, our method still has more than an order of magnitude enhancement in the BER at different SNR values.

V. Conclusions

In this work, cell-edge user detection in the uplink of a multi-cell multiuser MIMO system was considered. The goal is to design a detector that can reliably demodulate cell-edge user signals in the presence of strong in-cell interference from users close to the base station, without resorting on power control that throttles the users that are close to the base station. This paper proposed a two-stage based approach that managed to reliably identify cell-edge users' signals at low SNR, without even knowing their channels. First, two-view CCA was brought in to estimate the subspace containing the cell-edge users' signals shared by both base stations. Then, an efficient analytical method called RACMA that guarantees the identifiability of binary signals from well-conditioned mixtures was exploited to extract the cell-edge users' signals from the subspace. Simulations revealed that our blind method achieves more than an order of magnitude improvement in the BER compared to the 'oracle' zero forcing successive interference cancellation with perfect channel state information.

References

[1] S. Verdu *et al.*, *Multiuser detection*. Cambridge university press, 1998.

[2] J. G. Proakis and M. Salehi, *Digital communications*. McGraw-hill New York, 2001, vol. 4.

[3] H. Vikalo and B. Hassibi, "The expected complexity of sphere decoding, part i: Theory, part ii: Applications," *IEEE Transactions on Signal Processing*, 2003.

[4] J. Jaldén and B. Ottersten, "On the complexity of sphere decoding in digital communications," *IEEE transactions on signal processing*, vol. 53, no. 4, pp. 1474–1484, 2005.

[5] U. Madhow and M. L. Honig, "MMSE interference suppression for direct-sequence spread-spectrum CDMA," *IEEE transactions on communications*, vol. 42, no. 12, pp. 3178–3188, 1994.

[6] S. Moshavi, "Multi-user detection for DS-CDMA communications," *IEEE communications magazine*, vol. 34, no. 10, pp. 124–136, 1996.

[7] G. Boudreau, J. Panicker, N. Guo, R. Chang, N. Wang, and S. Vrzic, "Interference coordination and cancellation for 4G networks," *IEEE Communications Magazine*, vol. 47, no. 4, 2009.

[8] N. Himayat, S. Talwar, A. Rao, and R. Soni, "Interference management for 4G cellular standards [WIMAX/LTE UPDATE]," *IEEE Communications Magazine*, vol. 48, no. 8, 2010.

[9] H. Hotelling, "Relations between two sets of variates," *Biometrika*, vol. 28, no. 3/4, pp. 321–377, 1936.

[10] J. D. Carroll, "Generalization of canonical correlation analysis to three or more sets of variables," in *Proceedings of the 76th annual convention of the American Psychological Association*, vol. 3, 1968, pp. 227–228.

[11] A.-J. van der Veen, "Analytical method for blind binary signal separation," *IEEE Transactions on Signal Processing*, vol. 45, no. 4, pp. 1078–1082, 1997.

[12] S. Khattak, *Base Station Cooperation Strategies for Multi-user Detection in Interference Limited Cellular Systems*. Jörg Vogt Verlag, 2008.

[13] Q. Wu and K. M. Wong, "Un-music and un-cle: An application of generalized correlation analysis to the estimation of the direction of arrival of signals in unknown correlated noise," *IEEE Transactions on Signal Processing*, vol. 42, no. 9, pp. 2331–2343, 1994.

[14] A. Dogandzic and A. Nehorai, "Finite-length MIMO equalization using canonical correlation analysis," *IEEE Transactions on Signal Processing*, vol. 50, no. 4, pp. 984–989, 2002.

[15] J. Via and I. Santamaria, "Adaptive blind equalization of SIMO systems based on canonical correlation analysis," in *IEEE 6th Workshop on Signal Processing Advances in Wireless Communications*, NY, USA, June 2005, pp. 318–322.

[16] H. Ge, I. P. Kirsteins, and X. Wang, "Does canonical correlation analysis provide reliable information on data correlation in array processing?" in *IEEE International Conference on Acoustics, Speech and Signal Processing*, 2009, pp. 2113–2116.

[17] M. Borga and H. Knutsson, "A canonical correlation approach to blind source separation," *Report LiU-IMT-EX-0062 Department of Biomedical Engineering, Linkping University*, 2001.

[18] Y.-O. Li, T. Adali, W. Wang, and V. D. Calhoun, "Joint blind source separation by multiset canonical correlation analysis," *IEEE Transactions on Signal Processing*, vol. 57, no. 10, pp. 3918–3929, 2009.

[19] A. Bertrand and M. Moonen, "Distributed canonical correlation analysis in wireless sensor networks with application to distributed blind source separation," *IEEE Transactions on Signal Processing*, vol. 63, no. 18, pp. 4800–4813, 2015.

[20] R. Arora and K. Livescu, "Multi-view learning with supervision for transformed bottleneck features," in *IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, Italy, May 2014, pp. 2499–2503.

[21] X. Fu, K. Huang, E. E. Papalexakis, H. Song, P. P. Talukdar, N. D. Sidiropoulos, C. Faloutsos, and T. Mitchell, "Efficient and distributed algorithms for large-scale generalized canonical correlations analysis," in *2016 IEEE 16th International Conference on Data Mining (ICDM)*, Dec 2016, pp. 871–876.

[22] X. Fu, K. Huang, M. Hong, N. D. Sidiropoulos, and A. M. So, "Scalable and flexible multiview max-var canonical correlation analysis," *IEEE Transactions on Signal Processing*, vol. 65, no. 16, pp. 4150–4165, Aug 2017.

[23] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and M. Hong, "Structured sumcor multiview canonical correlation analysis for large-scale data," *IEEE Transactions on Signal Processing*, vol. 67, no. 2, pp. 306–319, Jan 2019.

[24] P. Horst, "Generalized canonical correlations and their applications to experimental data," *Journal of Clinical Psychology*, vol. 17, no. 4, pp. 331–347, 1961.

[25] D. R. Hardoon, S. Szedmak, and J. Shawe-Taylor, "Canonical correlation analysis: An overview with application to learning methods," *Neural computation*, vol. 16, no. 12, pp. 2639–2664, 2004.

[26] C. U. Castellanos, D. L. Villa, C. Rosa, K. I. Pedersen, F. D. Calabrese, P.-H. Michaelsen, and J. Michel, "Performance of uplink fractional power control in utran LTE," in *Vehicular Technology Conference (VTC)*, Calgary, Canada, Sept. 2008, pp. 2517–2521.