Reliable Detection of Unknown Cell-Edge Users Via Canonical Correlation Analysis

Mohamed Salah Ibrahim, *Student Member, IEEE,* and Nicholas D. Sidiropoulos, *Fellow, IEEE*

***Abstract*—Providing reliable service to users close to the edge between cells remains a challenge in cellular systems, even as 5G deployment is around the corner. These users are subject to significant signal attenuation, which also degrades their uplink channel estimates. Even joint detection using base station (BS) cooperation often fails to reliably detect such users, due to near-far power imbalance, and channel estimation errors. Is it possible to bypass the channel estimation stage and design a detector that can reliably detect cell-edge user signals under significant near-far imbalance? This paper shows, perhaps surprisingly, that the answer is affirmative – albeit not via traditional multiuser detection. Exploiting that cell-edge user signals are weak but *common* to different base stations, while cell-center users are unique to their serving BS, this paper establishes an elegant connection between cell-edge user detection and canonical correlation analysis (CCA) of the associated space-time baseband-equivalent matrices. It proves that CCA identifies the common subspace of these matrices, even under significant intra- and inter-cell interference. The resulting mixture of cell-edge user signals can subsequently be unraveled using a well-known algebraic signal processing technique. Interestingly, the proposed approach does not even require that the signals from the different base stations are synchronized – the right synchronization can be automatically determined as well. Experimental results demonstrate that the proposed approach achieves order of magnitude BER improvements compared to 'oracle' multiuser detection that assumes perfect knowledge of the cell-center user channels.**

I. INTRODUCTION

AT the dawn of 5G, providing reliable high-speed service to users on the edge between cells remains a challenge that has persisted through several generations of cellular wireless systems. In 4G and legacy systems, the problem is usually tackled using aggressive power control, multiuser detection, and dynamic base station (BS) assignment / hand-off [2], [3]. Multiuser detection (MUD) is computationally complex (optimal MUD is NP-hard) [4], [5], requires accurate channel estimates for all users, and while it can tolerate power imbalance, practically tractable multiuser detection does not work well in near-far scenarios, especially when the channels for the far users are not accurately known. The so-called sphere decoder (SD – a branch-and-bound type implementation of the maximum likelihood detector) features significantly lower complexity than naive implementations at moderately high signal to noise ratios (SNRs), albeit worst-case and average complexities remain exponential [6], [7]. Semidefinite relaxation (SDR) is a polynomial-time alternative to SD, in the low to moderate SNR regime where it yields better error rates and lower complexity than SD [8], [9]. The complexity of SDR remains high for practical implementation [10].

Minimum mean square error (MMSE) [11], and the zero-forcing (ZF – also known as the decorrelating) detector are low-complexity linear detectors, whose performance remains far from optimal in general. ZF and MMSE detectors can be further improved by successively canceling the strong user signals once they are decoded – a technique referred to as successive interference cancellation (SIC), decision feedback (DF) [12], [13], or 'turbo' (iterative) interference cancellation [14].

Although all of the aforementioned detectors have been proven successful in many applications, their detection performance is contingent on the availability of accurate channel estimates. In wireless cellular systems, accurate channel estimates may be acquired for cell-center (strong) users, however, cell-edge (weak) user signals are received at low SNR due to the inverse power law relationship between received signal power and distance. This and the intra- and inter-cell interference (particularly prominent for the cell-edge users) together induce high uncertainty in the cell-edge user channel estimates, degrading their detection performance and even leading to connection drops [15], [16]. Furthermore, the frequent hand-offs of such users further complicate their situation [17]. While power control [18] and scheduling algorithms [19], [20] serve as two possible candidates that can considerably enhance cell-edge user detection performance, this comes at the expense of significantly reducing the rates of cell-center users. These are the ones with the best channels, so throttling their rate has a serious impact on the overall sum rate of the system.

This begs the question whether it is possible to reliably detect cell-edge user signals without knowing their channels or sacrificing cell-center user rates?

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 user signals can be reliably decoded under mild conditions, even at low SNR and when buried under heavy intra-cell and inter-cell interference. Exploiting the fact that cell-edge user signals are *weak but common* to both base stations, while

Original manuscript submitted to IEEE Transactions on Wireless Communications, July 25, 2019; revised January 3, 2020; accepted March 5, 2020. Preliminary version of part of this work [1] was presented at the 20th IEEE International Workshop on Signal Processing Advances in Wireless Communications (SPAWC), Cannes, France, July 2019. Supported by NSF ECCS-1807660.
Mohamed Salah Ibrahim, and Nicholas D. Sidiropoulos are with the Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, VA, 22904 USA (e-mail (mi6cw,nikos@virginia.edu))

users close to a base station are unique to that base station, reliable detection is enabled by Canonical Correlation Analysis (CCA) [21], [22] – a machine learning technique that reliably estimates a common subspace using eigendecomposition, even in the presence of strong interference.

Our approach is very different from multi-user detection using base station cooperation [23], as it capitalizes on CCA. CCA has been employed in several signal processing, communications, and machine learning applications, including array processing [24], multiple-input multiple-output (MIMO) equalization [25], [26], direction-of-arrival (DoA) estimation [27], radar anti-jamming [28] and blind source separation [29]–[32], and multi-view learning [33], 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 [34]–[36], also incorporating various constraints.

A. Contributions

This paper proposes a two-stage learning based approach that leverages base station cooperation to reliably detect cell-edge user signals without knowing their channels. The idea relies on connecting canonical correlation analysis with cell-edge user detection. In the first stage, CCA is invoked to find the common subspace of two space-time matrices, containing the baseband-equivalent signals received at two base stations. A basis for this common subspace is a mixture of the cell-edge user signals. In the second stage, this mixture is unraveled in an unsupervised fashion, using a classical algebraic technique from array signal processing, namely (R)ACMA [37]. (R)ACMA exploits constant modulus structure in the transmitted cell-edge signals, owing to digital binary/M-ary phase shift keying (BPSK or MPSK) modulation, to recover the individual cell-edge signals. Judicious experiments demonstrate that the proposed approach works remarkably well without any power control under realistic levels of intra-cell and inter-cell interference (following the urban macro scenario from the 3GPP 38.901 standard), delivering order of magnitude error rate improvements compared to 'oracle' multiuser detection that assumes perfect knowledge of the cell-center user channels. Furthermore, the proposed approach does not even require that the signals from the different base stations are synchronized – the right synchronization can be automatically determined as well.

Beyond these compelling contributions to the particular application in cellular communications considered herein, this paper makes two notable theoretical contributions of broader interest. First, it proves that CCA identifies the common subspace between two matrices, under a rather general (and purely deterministic) linear generative model. Second, it includes a performance analysis which shows that CCA works even in the non-ideal case where there is background noise and 'leakage' of the individual components to the other matrix view – e.g., the case where there is thermal noise and realistic adjacent-cell interference from non-cell-edge users that cannot be neglected, in the context of our application herein.

The overall complexity of the proposed method depends on the cost incurred in solving CCA and RACMA. Fortunately, both admit relatively simple algebraic solution via eigenvalue decomposition [22], [37]. This renders the overall approach computationally efficient even when the base station is equipped with a large number of antennas and is serving a large number of users.

A preliminary version of part of this work was presented at IEEE SPAWC 2019 [1]. This journal version includes performance analysis that was missing from [1], a new section showing how the common cell-edge signals can be used to synchronize the signals from the two base stations even if they were asynchronously acquired (thus alleviating the need for symbol-level synchronization), and a new comprehensive suite of experiments to demonstrate the superior performance of the proposed method in more practical scenarios.

B. Outline of the Paper

The rest of this paper is organized as follows. After a succinct introduction to CCA in Section II, Section III describes the system model and gives a brief review on cell-edge user detection. The proposed detector is presented in Section IV, while Section V explains how our detector can be used to resolve symbol synchronization between the two base stations. Simulation results are provided in Section VI, and conclusions are drawn in Section VII. Long proofs and derivations are relegated to the Appendix.

C. Notation

In this work, we use upper and lower case bold letters to denote matrices and column vectors, respectively. For any general matrix $\mathbf{G}$, we use $\mathbf{G}^T$, $\mathbf{G}^H$, $\mathbf{G}^{-1}$, $\mathbf{G}^\dagger$ and $\mathrm{Tr}(\mathbf{G})$ to denote the transpose, the conjugate-transpose, the inverse, the pseudo-inverse, and the trace of $\mathbf{G}$, respectively. Scalars are represented in the normal face, while calligraphic font is reserved for sets. $\|.\|_2$ 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. Canonical Correlation Analysis

Consider T samples of the pair $(\mathbf{y}_1, \mathbf{y}_2)$, where $\mathbf{y}_1 \in \mathbb{R}^{M_1}$ and $\mathbf{y}_2 \in \mathbb{R}^{M_2}$ are two "views" of the same entity. For example, $\mathbf{y}_1$ could contain a set of economic indicators, while $\mathbf{y}_2$ could contain crime, corruption, or social welfare data corresponding to the same country or municipality, and we have data for T countries or municipalities. Or, $\mathbf{y}_1$ could be the electroencephalogram (EEG) of a person and $\mathbf{y}_2$ could be the voxels of a functional magnetic resonance (fMRI) scan; or $\mathbf{y}_1$ could be a person's consumer record, while $\mathbf{y}_2$ could reflect his/her social network connections, and we have data for T people. We are interested in discovering what is common between these two views of the same set of entities. Is there a particular 'latent' factor that affects both the economy and crime, for example? Towards this end, we would like to derive 'meta-variables', one from each view, which are strongly correlated with each other. How can we do this?

Let $\mathbf{y}_1[t]$ and $\mathbf{y}_2[t]$ denote the t-th observation of $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively, corresponding to the t-th entity, for

$t \in \{1, \cdots, T\}$. Assume that both $\mathbf{y}_1$ and $\mathbf{y}_2$ are zero-mean, otherwise the sample mean can be subtracted as a pre-processing step. In its simplest form, CCA seeks to find a pair of linear combinations of the variables in the two respective views which are highly correlated to each other – ideally, perfectly correlated. Mathematically, CCA seeks to weight vectors $\mathbf{q}_1 \in \mathbb{R}^{M_1}$ and $\mathbf{q}_2 \in \mathbb{R}^{M_2}$ such that the correlation coefficient between $\mathbf{Y}_1^T\mathbf{q}_1$ and $\mathbf{Y}_2^T\mathbf{q}_2$ is maximized, where $\mathbf{Y}_\ell := [\mathbf{y}_\ell[1], \cdots, \mathbf{y}_\ell[T]] \in \mathbb{R}^{M_\ell \times T}$ and $\ell \in \{1, 2\}$. In an optimization framework, this can be expressed as

$$\max_{\mathbf{q}_1, \mathbf{q}_2} \frac{\mathbf{q}_1^T \mathbf{Y}_1 \mathbf{Y}_2^T \mathbf{q}_2}{\sqrt{\mathbf{q}_1^T \mathbf{Y}_1 \mathbf{Y}_1^T \mathbf{q}_1}\sqrt{\mathbf{q}_2^T \mathbf{Y}_2 \mathbf{Y}_2^T \mathbf{q}_2}} \tag{1}$$

Let $\mathbf{R}_{\mathbf{y}_\ell \mathbf{y}_\ell} := \frac{1}{T}\mathbf{Y}_\ell \mathbf{Y}_\ell^T$ and $\mathbf{R}_{\mathbf{y}_1 \mathbf{y}_2} := \frac{1}{T}\mathbf{Y}_1 \mathbf{Y}_2^T$ denote the sample auto-correlation matrix of $\mathbf{y}_\ell$ and the sample cross-correlation matrix of $\mathbf{y}_1$ and $\mathbf{y}_2$, respectively. Then, (1) can be equivalently written as

$$\max_{\mathbf{q}_1, \mathbf{q}_2} \ \mathbf{q}_1^T \mathbf{R}_{\mathbf{y}_1 \mathbf{y}_2} \mathbf{q}_2 \tag{2a}$$

$$\text{s.t.} \quad \mathbf{q}_\ell^T \mathbf{R}_{\mathbf{y}_\ell \mathbf{y}_\ell} \mathbf{q}_\ell = 1, \ \ell = 1, 2 \tag{2b}$$

where the constraints in (2b) arise from the fact that the objective of (1) is not affected by re-scaling $\mathbf{q}_1$ and/or $\mathbf{q}_2$. Using the Lagrange duality theorem, a solution of (2) can be provided in closed-form. The Lagrangian of (2) is

$$\mathcal{L}(\mathbf{q}_1, \mathbf{q}_2, \lambda_1, \lambda_2) = \mathbf{q}_1^T \mathbf{R}_{\mathbf{y}_1 \mathbf{y}_2} \mathbf{q}_2 - \sum_{\ell=1}^{2} \frac{\lambda_\ell}{2}(\mathbf{q}_\ell^T \mathbf{R}_{\mathbf{y}_\ell \mathbf{y}_\ell} \mathbf{q}_\ell - 1) \tag{3}$$

By taking the derivatives with respect to $\mathbf{q}_1$ and $\mathbf{q}_2$, we obtain

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}_1} = \mathbf{R}_{\mathbf{y}_1 \mathbf{y}_2} \mathbf{q}_2 - \lambda_1 \mathbf{R}_{\mathbf{y}_1 \mathbf{y}_1} \mathbf{q}_1 = 0 \tag{4}$$

$$\frac{\partial \mathcal{L}}{\partial \mathbf{q}_2} = \mathbf{R}_{\mathbf{y}_2 \mathbf{y}_1} \mathbf{q}_1 - \lambda_2 \mathbf{R}_{\mathbf{y}_2 \mathbf{y}_2} \mathbf{q}_2 = 0 \tag{5}$$

By left multiplying (4) and (5) with $\mathbf{q}_1^T$ and $\mathbf{q}_2^T$, respectively, we have

$$\mathbf{q}_1^T \mathbf{R}_{\mathbf{y}_1 \mathbf{y}_2} \mathbf{q}_2 = \lambda_1 \mathbf{q}_1^T \mathbf{R}_{\mathbf{y}_1 \mathbf{y}_1} \mathbf{q}_1 \tag{6}$$

$$\mathbf{q}_2^T \mathbf{R}_{\mathbf{y}_2 \mathbf{y}_1} \mathbf{q}_1 = \lambda_2 \mathbf{q}_2^T \mathbf{R}_{\mathbf{y}_2 \mathbf{y}_2} \mathbf{q}_2 \tag{7}$$

which together with the constraints in (2b) imply that $\lambda_1 = \lambda_2 = \lambda$. By assuming that the matrix $\mathbf{R}_{\mathbf{y}_2 \mathbf{y}_2}$ is invertible, the optimal solution, $\mathbf{q}_2^*$, of (5) is given by

$$\mathbf{q}_2^* = \frac{1}{\lambda} \mathbf{R}_{\mathbf{y}_2 \mathbf{y}_2}^{-1} \mathbf{R}_{\mathbf{y}_2 \mathbf{y}_1} \mathbf{q}_1^* \tag{8}$$

Then by substituting in (4), the optimal solution, $\mathbf{q}_1^*$, can be obtained by solving the following generalized eigenvalue problem

$$\mathbf{R}_{\mathbf{y}_1 \mathbf{y}_2} \mathbf{R}_{\mathbf{y}_2 \mathbf{y}_2}^{-1} \mathbf{R}_{\mathbf{y}_2 \mathbf{y}_1} \mathbf{q}_1 = \lambda^2 \mathbf{R}_{\mathbf{y}_1 \mathbf{y}_1} \mathbf{q}_1 \tag{9}$$

It can be easily seen from (4) that the maximum eigenvalue λ^* of (9) is nothing but the square of the correlation coefficient, ρ_1, associated with the canonical pair $(\mathbf{q}_1^*, \mathbf{q}_2^*)$.

Considering the generalization to $N \leq \min(M_1, M_2)$ canonical pairs, $\{(\mathbf{q}_1[n], \mathbf{q}_2[n])\}_{n=1}^N$. After identifying $\mathbf{q}_1^*[1] = \mathbf{q}_1^*$ and $\mathbf{q}_2^*[1] = \mathbf{q}_2^*$, we can iteratively solve the following problem

$$\max_{\mathbf{q}_1[n], \mathbf{q}_2[n]} \quad \mathbf{q}_1^T[n] \mathbf{R}_{\mathbf{y}_1 \mathbf{y}_2} \mathbf{q}_2[n] \tag{10a}$$

$$\text{s.t.} \quad \mathbf{q}_\ell^T[n] \mathbf{R}_{\mathbf{y}_\ell \mathbf{y}_\ell} \mathbf{q}_\ell[n] = 1, \ \ell = 1, 2 \tag{10b}$$

$$\mathbf{q}_\ell^T[n] \mathbf{R}_{\mathbf{y}_\ell \mathbf{y}_\ell} \mathbf{q}_\ell[j] = 0, \ j = 1, \cdots, n-1 \tag{10c}$$

for $n = \{2, \cdots, N\}$. Instead of solving N sub-problems of type (10), we can instead solve one joint problem. Let us stack the vectors $\{\mathbf{q}_\ell[n]\}_{n=1}^N$ in the matrix $\mathbf{Q}_\ell \in \mathbb{R}^{M_\ell \times N}$, for $\ell \in \{1, 2\}$, and rewrite (10) in the following compact form

$$\max_{\mathbf{Q}_1, \mathbf{Q}_2} \ \text{Tr}(\mathbf{Q}_1^T \mathbf{R}_{\mathbf{y}_1 \mathbf{y}_2} \mathbf{Q}_2) \tag{11a}$$

$$\text{s.t.} \quad \mathbf{Q}_\ell^T \mathbf{R}_{\mathbf{y}_\ell \mathbf{y}_\ell} \mathbf{Q}_\ell = \boldsymbol{I}, \ \ell = 1, 2 \tag{11b}$$

which yields simultaneously multiple canonical pairs. Following the same procedures for solving (2), it can be shown that the optimal solution $\mathbf{Q}_1^*$ should satisfy the following generalized eigenvalue equation

$$\mathbf{R}_{\mathbf{y}_1 \mathbf{y}_2} \mathbf{R}_{\mathbf{y}_2 \mathbf{y}_2}^{-1} \mathbf{R}_{\mathbf{y}_2 \mathbf{y}_1} \mathbf{Q}_1 = \mathbf{R}_{\mathbf{y}_1 \mathbf{y}_1} \mathbf{Q}_1 \Lambda^2 \tag{12}$$

where $\Lambda = \text{Diag}([\rho_1, \cdots, \rho_N])$ with ρ_ℓ be the ℓ-th correlation coefficient associated with the ℓ-th canonical pair, for $\ell = \{1, \cdots, N\}$. Note that the optimal solution $\mathbf{Q}_2^*$ can be directly obtained from (8) after solving (12).

The two-view CCA problem in (11) can be equivalently formulated as a distance minimization between the low dimensional representations $\mathbf{Y}_1^T \mathbf{Q}_1$ and $\mathbf{Y}_2^T \mathbf{Q}_2$ [22], [38], where the distance is measured by the Frobenius norm, i.e.,

$$\min_{\mathbf{Q}_1, \mathbf{Q}_2} \ \|\mathbf{Y}_1^T \mathbf{Q}_1 - \mathbf{Y}_2^T \mathbf{Q}_2\|_F^2 \tag{13a}$$

$$\text{s.t.} \quad \mathbf{Q}_\ell^T \mathbf{Y}_\ell \mathbf{Y}_\ell^T \mathbf{Q}_\ell = \boldsymbol{I}, \quad \ell = 1, 2 \tag{13b}$$

Note that by expanding the objective in (13), the equivalence between (12) and (13) can be readily verified. In what follows, we will see how the CCA approach can be utilized to handle the problem of cell-edge user detection in a multi-cell multi-user system.

III. Problem Statement

A. System Model

Consider a multi-cell multi-user MIMO system comprising two hexagonal cells with a single base station (BS) located at the center of each cell, as shown in Figure 1. The ℓ-th BS is equipped with M_ℓ antennas and serves K_ℓ single-antenna users, for $\ell \in \{1, 2\}$. Let $K_e = K_{e_1} + K_{e_2}$ denote the total number of cell-edge users located around the common edge between the two cells, where $K_{e_\ell} < K_\ell$ represents the number of cell-edge users served by the ℓ-th BS. Let $\mathbf{h}_{\ell kj}$ model path-loss and small scale fading between the k-th user in the j-th cell and the ℓ-th BS, given by

$$\mathbf{h}_{\ell kj} = \sqrt{\alpha_{\ell kj}} \mathbf{z}_{\ell kj} \tag{14}$$

where $\mathbf{z}_{\ell kj} \in \mathbb{C}^{M_\ell \times 1}$ represents the small scale fading between user k in cell j and BS ℓ, while $\alpha_{\ell kj} \in \mathbb{R}$ models the large scale fading that accounts for the path loss between BS ℓ and user k in cell j. Throughout this work, it is assumed that the uplink channel vectors $\mathbf{h}_{\ell kj}$ for the cell-edge users are not *a priori* known at any BS.

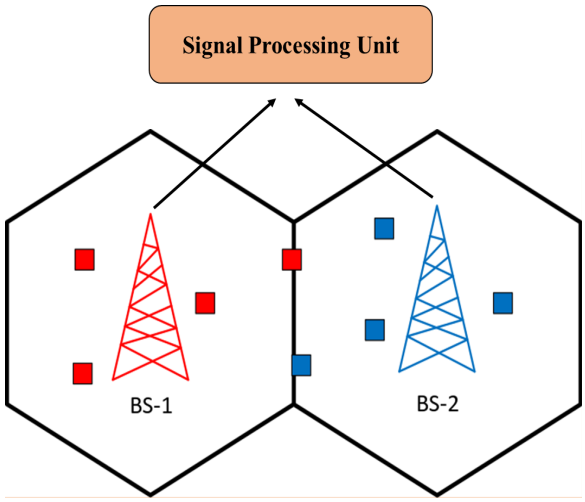


Fig. 1: System Model

B. Uplink transmission

Consider uplink transmission from the users to the BSs where each user aims at transmitting its data to its serving BS. We assume that all users access the same channel without any (sub-) channel allocation or coordination mechanism, thereby creating intra- and inter-cell interference. Define $\mathbf{s}_{kj} \in \mathbb{R}^{T\times 1}$ to 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 variations in the second stage). The received signal, $\mathbf{Y}_\ell \in \mathbb{C}^{M_\ell \times T}$, at the ℓ-th BS can be expressed as

$$\mathbf{Y}_\ell = \sum_{j=1}^{2}\sum_{k=1}^{K_j} \sqrt{\beta}_{kj}\mathbf{h}_{\ell kj}\mathbf{s}_{kj}^T + \mathbf{W}_\ell \tag{15}$$

where $\mathbf{h}_{\ell kj} \in \mathbb{C}^{M_\ell \times 1}$ is the uplink channel response vector defined in (14), $\mathbf{W}_\ell \in \mathbb{C}^{M_\ell \times T}$ contains independent identically distributed (i.i.d.) complex Gaussian entries of zero mean and variance σ^2, and β_{kj} represents the transmit power of the k-th user in the j-th cell.

Throughout this paper, we assume that each BS forwards its received signal to a central signal processing unit (CSPU). Although BSs cooperation has been considered before for the sake of mitigating inter-cell interference [39], cooperation here is assumed for a very different purpose. That is, we leverage the joint processing of the BSs signals at the CSPU to provide reliable detection of cell-edge user signals at low SNR, without knowledge of their channels. Furthermore, in contrast to prior cooperation strategies that assume perfect synchronization of the received signals from different BSs [23], [40], this work deals with BS asynchrony as well, rendering the approach more practical. Specifically, it will be shown in Section V how the proposed method can detect the cell-edge user signals even if there exists a time delay between $\mathbf{Y}_1$ and $\mathbf{Y}_2$.

C. Cell-Edge User Detection

Let us denote the cell-edge user transmitted signals by $\mathbf{S}_c \in \mathbb{R}^{T\times K_e}$ (where the subscript c stands for *common*), and those of the cell-center served by the ℓ-th BS as $\mathbf{S}_{p_\ell} \in \mathbb{R}^{T\times(K_\ell - K_{e_\ell})}$ (where the subscript p stands for *private*). Furthermore, let $\widetilde{\mathbf{W}}_\ell$ represent the noise at the ℓ-th BS plus the inter-cell interference caused by the cell-center users in cell j, where $j \neq \ell$. Therefore, (15) can be expressed as follows

$$\mathbf{Y}_\ell = \mathbf{H}_{\ell p_\ell}\mathbf{S}_{p_\ell}^T + \mathbf{H}_{\ell c}\mathbf{S}_c^T + \widetilde{\mathbf{W}}_\ell \tag{16}$$

where the matrices $\mathbf{H}_{\ell c} \in \mathbb{C}^{M_\ell \times K_e}$ and $\mathbf{H}_{\ell p_\ell} \in \mathbb{C}^{M_\ell \times (K_\ell - K_{e_\ell})}$ 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. Moreover, absorb the transmitted signal power, β_{kj}, of the k-th user in the j-th cell in its respective channel vectors, $\forall\ k, j$.

In general, to guarantee reliable detection performance for each cell-edge user, its serving BS requires accurate knowledge about its channel state information (CSI) [41]–[43]. However, due to the fact that cell-edge user signals are often received intermittently at very low signal to interference plus noise ratio (SINR) and SNR, their channel estimates are inaccurate [15], [16].

One possible approach to detect cell-edge user signals is to apply zero-forcing successive interference cancellation (ZF-SIC) [12], which is based on successively removing the cell-center (strong) user signals once they are detected using ZF. Applying SIC after ZF improves the detection performance of the cell edge user signals as it (ideally) cancels the strong interference that stems from the transmissions of cell-center users, i.e., intra-cell interference. However, cell-center user detection is imperfect, which can lead to error propagation, and in-cell SIC does not address the inter-cell interference, which is particularly prominent for the cell-edge users. In the absence of power control [18] and/or scheduling [19] , cell-edge user detection performance is severely affected by the intra-cell interference from cell-center users. In what follows, we present a novel *blind* detector that can reliably decode cell-edge user signals at low received SNR and without knowing their channels.

IV. Cell-Edge User Detection Via CCA

In this section, it is assumed that the base stations signals are perfectly synchronized at the CSPU. We will explain how to deal with asynchrony later. The goal of the proposed detector is to decode cell-edge user signals $\mathbf{S}_c$ from the received signals $\mathbf{Y}_1$ and $\mathbf{Y}_2$. As a pre-processing step, the signals are transformed to the real domain by forming the matrix $\overline{\mathbf{Y}}_\ell := [\mathbf{Y}_\ell^{(r)}; \mathbf{Y}_\ell^{(i)}] \in \mathbb{R}^{2M_\ell \times T}$, where $\mathbf{Y}_\ell^{(r)} = \mathbb{R}\mathrm{e}\{\mathbf{Y}_\ell\}$ and $\mathbf{Y}_\ell^{(i)} = \mathbb{I}\mathrm{m}\{\mathbf{Y}_\ell\}$ represent the real and imaginary components of the ℓ-th BS signal. Similarly, denote by $\overline{\mathbf{A}}_{\ell p_\ell} := [\mathbf{H}_{\ell p_\ell}^{(r)}; \mathbf{H}_{\ell p_\ell}^{(i)}] \in \mathbb{R}^{2M_\ell \times (K_\ell - K_{e_\ell})}$, $\overline{\mathbf{A}}_{\ell c} := [\mathbf{H}_{\ell c}^{(r)}; \mathbf{H}_{\ell c}^{(i)}] \in \mathbb{R}^{2M_\ell \times K_e}$ and $\overline{\mathbf{W}}_\ell = [\widetilde{\mathbf{W}}_\ell^{(r)}; \widetilde{\mathbf{W}}_\ell^{(i)}] \in \mathbb{R}^{2M_\ell \times T}$. Therefore, (16) can be equivalently expressed as

$$\overline{\mathbf{Y}}_\ell = \overline{\mathbf{A}}_{\ell p_\ell}\mathbf{S}_{p_\ell}^T + \overline{\mathbf{A}}_{\ell c}\mathbf{S}_c^T + \overline{\mathbf{W}}_\ell. \tag{17}$$

Remark 1. *Due to the broadcast nature of the wireless medium, each BS may (over)hear the transmitted signals of all users. However, due to the inverse relationship between*

power and distance, the received signal power of the cell-center users associated with the ℓ-th BS is high at the ℓ-th BS (the serving one) and low at the j-th BS (the non-serving one). This power imbalance renders the received SNR of these users to be high at the serving BS and very low at the non-serving one, and hence, one can think of these users as being "private" to their serving BS, as their received signals are around the noise floor at the non-serving BS. On the other hand, cell-edge users are approximately half-way between two different BSs, and so they are received at commensurate power at both BSs. In this sense, cell-edge users are "common" to both BSs. In what follows, we will show theoretically and experimentally that our proposed CCA-based approach can reliably recover these common cell-edge user signals under realistic conditions.

In what follows, the two-view CCA formulation in (13) is exploited to estimate the subspace containing the cell-edge user signals. For the sake of brevity, we refer to this subspace as the *common subspace*. Define the two matrices $\overline{\mathbf{Q}}_1 \in \mathbb{R}^{2M_1 \times N}$ and $\overline{\mathbf{Q}}_2 \in \mathbb{R}^{2M_2 \times N}$, where the n-th column of $\overline{\mathbf{Q}}_\ell$ represents the n-th canonical component of view $\overline{\mathbf{Y}}_\ell$, for $n \in \{1, \cdots, N\}$. The number of components (pairs) extracted, (N), depends on the minimum value of the correlation coefficient that needs to be considered.

An alternative formulation of (13) 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 $\overline{\mathbf{Q}}_1$ and $\overline{\mathbf{Q}}_2$, respectively. This can be written as

$$\min_{\overline{\mathbf{Q}}_1, \overline{\mathbf{Q}}_2, \mathbf{G}} \sum_{\ell=1}^{2} \|\overline{\mathbf{Y}}_\ell^T \overline{\mathbf{Q}}_\ell - \mathbf{G}\|_F^2 \tag{18a}$$

$$\text{s.t.} \quad \mathbf{G}^T\mathbf{G} = \boldsymbol{I} \tag{18b}$$

Problem (18) is known as the maximum-variance (MAX-VAR) formulation of CCA [38], [44], and, in the case of two views considered here, it is equivalent to (13) in the sense that both problems yield the same solution $\overline{\mathbf{Q}}_\ell^*$. In this section, we focus on the formulation in (18) as it facilitates our proof.

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

Theorem 1. *In the noiseless case, if matrix $\mathbf{B} := [\mathbf{S}_c, \mathbf{S}_{p_1}, \mathbf{S}_{p_2}] \in \mathbb{R}^{T \times (K_1+K_2)}$ is full column rank, and $\overline{\mathbf{A}}_\ell = [\overline{\mathbf{A}}_{\ell c}, \overline{\mathbf{A}}_{\ell p_\ell}] \in \mathbb{R}^{2M_\ell \times (K_e + K_\ell - K_{e_\ell})}$ is full column rank for $\ell \in \{1, 2\}$, then the optimal solution $\mathbf{G}^\star$ of problem (18) is given by $\mathbf{G}^\star = \mathbf{S}_c\mathbf{P}$, where $\mathbf{P}$ is a $K_e \times K_e$ non-singular matrix.*

Remark 2. *The full column rank condition on $\mathbf{B}$ requires T greater than or equal to $(K_1 + K_2)$, and the transmitted sequences from the different users to be linearly independent. For finite-alphabet signals, this occurs with very high probability for modest 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 greater than or equal to 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., $K_e = 1$ and each of $\mathbf{S}_c, \mathbf{G}$ and $\mathbf{Q}_\ell$ is a vector. In such setting (18) reduces to the following

$$\min_{\overline{\mathbf{q}}_1, \overline{\mathbf{q}}_2, \mathbf{g}} \sum_{\ell=1}^{2} \|\overline{\mathbf{Y}}_\ell^T \overline{\mathbf{q}}_\ell - \mathbf{g}\|_2^2 \tag{19a}$$

$$\text{s.t.} \quad \|\mathbf{g}\|_2^2 = 1 \tag{19b}$$

To solve the above problem, we need to find $(\overline{\mathbf{q}}_1^*, \overline{\mathbf{q}}_2^*, \mathbf{g}^*)$ that can together attain a zero-cost. In other words, we need the following two conditions to be satisfied simultaneously

$$\overline{\mathbf{Y}}_1^T \overline{\mathbf{q}}_1 = \mathbf{g} \tag{20a}$$

$$\overline{\mathbf{Y}}_2^T \overline{\mathbf{q}}_2 = \mathbf{g} \tag{20b}$$

Without loss of generality, we can let $\overline{\mathbf{q}}_\ell = \overline{\mathbf{A}}_\ell (\overline{\mathbf{A}}_\ell^T \overline{\mathbf{A}}_\ell)^{-1} \mathbf{u}_\ell$, where $\mathbf{u}_\ell$ is any vector in $\mathbb{R}^{K_e + K_\ell - K_{e_\ell}}$. The reason is that we can always decompose $\overline{\mathbf{q}}_\ell$ 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 $\overline{\mathbf{A}}_\ell^T$. Substituting in (20a) and (20b) and taking their difference, we obtain

$$\mathbf{B}\mathbf{u} = \mathbf{0}, \tag{21}$$

where $\mathbf{B} = [\mathbf{s}_c, \mathbf{S}_{p_1}, \mathbf{S}_{p_2}] \in \mathbb{R}^{T \times (K_1+K_2)}$ and $\mathbf{u} = [\mathbf{u}_1(1) - \mathbf{u}_2(1), \mathbf{u}_1(2 : \text{end}), -\mathbf{u}_2(2 : \text{end})]^T \in \mathbb{R}^{(K_1+K_2)}$, 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 $\mathbf{u} = \mathbf{0}_{(K_1+K_2)\times 1}$ is the only possible solution of (21). 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 (20), $\mathbf{g}^\star = \alpha \mathbf{s}_c / \|\mathbf{s}_c\|_2$, with $\alpha = \pm 1$, will be the only possible solution for problem (19).

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

$$\mathbf{U} := \begin{bmatrix} & \mathbf{U}_1(1 : K_e, :) - \mathbf{U}_2(1 : K_e, :) \\ & \mathbf{U}_1(K_e + 1 : \text{end}, :) \\ - & \mathbf{U}_2(K_e + 1 : \text{end}, :) \end{bmatrix} \in \mathbb{R}^{(K_1+K_2) \times K_e},$$

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

$$\mathbf{B}\mathbf{U} = \mathbf{0}, \tag{22}$$

and when $\mathbf{B}$ is full column rank the solution is unique: $\mathbf{U} = \mathbf{0}$, and therefore $\mathbf{U}_1(1 : K_e, :) = \mathbf{U}_2(1 : K_e, :) =: \mathbf{P}$, $\mathbf{U}_1(K_e + 1 : \text{end}, :) = \mathbf{0}$, $\mathbf{U}_2(K_e + 1 : \text{end}, :) = \mathbf{0}$, and therefore $\mathbf{G}^\star = \mathbf{S}_c\mathbf{P}$, where $\mathbf{P}$ is $K_e \times K_e$ non-singular such that the orthonormality constraint (18b) is satisfied. Note that if the signals themselves are (approximately) orthogonal, then $\mathbf{P}$ will be orthogonal as well, which helps with the next (RACMA) stage. □

Remark 3. *Theorem 1 provides results for an idealized scenario, where at each BS we ignore other-cell signals coming from users that are not close to the given BS's cell-edge. This is a reasonable approximation that we use to prove that the cell-edge signals can be recovered even when buried under very strong cell-center signals. In other words, Theorem 1 says that if we have two multi-antenna signal "views" that include very strong but private components (in our context, the received signals of each group of cell-center users at their serving BS, respectively) and weak but common components (in our context, the received signals of the cell-edge users between the two BSs), then CCA will exactly recover the subspace of the common components irrespective of their relatively low power. We will later present an elegant analysis which shows that what matters is the power (im)balance: signals received at roughly the same power at the two BSs are "common" and recovered via CCA, and signals received at high SNR at one BS and low SNR at the other BS are "private", and cannot be recovered by CCA.*

The next step is to extract the cell-edge user sequences $\mathbf{S}_c$ from $\mathbf{G}^\star = \mathbf{S}_c\mathbf{P}$. This problem can be viewed as a bi-linear 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

$$\min_{\overline{\mathbf{S}}_c, \overline{\mathbf{P}}} \ \|\mathbf{G}^\star - \overline{\mathbf{S}}_c\overline{\mathbf{P}}\|_F^2 \tag{23a}$$

$$\text{s.t.} \quad \overline{\mathbf{S}}_c(i,j) \in \Omega \tag{23b}$$

In [37], van der Veen proposed an algebraic algorithm called Real Analytical Constant Modulus Algorithm (RACMA) for this problem. RACMA does not claim to optimally solve (23), which is NP-hard even if $\mathbf{P}$ is known. Instead, RACMA assumes that noise is small, and reduces (23) 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 column-wise (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 signal (and sign) via correlation with the identification sequence.

The following Algorithm describes the two-step procedure for cell-edge users detection via CCA followed by RACMA.

Algorithm 1 CCA for Cell-Edge User Detection

Input: $\overline{\mathbf{Y}}_1, \overline{\mathbf{Y}}_2$

1) Solve problem (13) for $\overline{\mathbf{Q}}_\ell$ as explained in Section II
2) Compute $\mathbf{G}_\ell = \overline{\mathbf{Y}}_\ell^T\overline{\mathbf{Q}}_\ell \in \mathbb{R}^{T\times K_e}, \ \forall \ell = 1,2$
3) Construct $\mathbf{G} = [\mathbf{G}_1; \mathbf{G}_2] \in \mathbb{R}^{2T\times K_e}$ and pass it to RACMA
4) Compute the BER of cell-edge users by comparing the output of RACMA with the original cell-edge user transmitted sequences

Notice that the second step in Algorithm 1 stems out from the fact that the zero-cost solution of problem (18) is not guaranteed in the noisy case, and therefore, $\overline{\mathbf{Y}}_1^T\overline{\mathbf{Q}}_1$ is not equal to $\overline{\mathbf{Y}}_2^T\overline{\mathbf{Q}}_2$ in general. Then, it turns out that feeding RACMA with both $\overline{\mathbf{Y}}_1^T\overline{\mathbf{Q}}_1$ and $\overline{\mathbf{Y}}_2^T\overline{\mathbf{Q}}_2$ simultaneously results in much better BER as we will see in Section VI.

The overall complexity of the proposed method comes from solving problems (13) and (23). Fortunately, similar to (13), (23) also admits simple algebraic solution via eigenvalue decomposition [37]. This means that our end-to-end method requires solving two eigenvalue problems, i.e., the overall complexity is of $O(M^3)$, with $M = \max\{M_1, M_2\}$.

It is important to emphasize that, in the noisy case and under inter-cell interference (i.e., users close to base station B can be overheard at base station A), it turns out that our method can still identify the common subspace, even at low SNR values. In order to show this, we follow a very different path from that described in the proof of Theorem 1.

First, let us define two channel matrices $\mathbf{H}_1 := [\mathbf{H}_{1c}, \mathbf{H}_{1p_1}, \mathbf{H}_{1p_2}] \in \mathbb{C}^{M_1\times(K_1+K_2)}$ and $\mathbf{H}_2 := [\mathbf{H}_{2c}, \mathbf{H}_{2p_1}, \mathbf{H}_{2p_2}] \in \mathbb{C}^{M_2\times(K_1+K_2)}$, where $\mathbf{H}_\ell$ holds in its columns all the channel vectors from all users to the ℓ-th BS, for $\ell = 1,2$. Note that one can factor $\mathbf{H}_\ell = \mathbf{Z}_\ell\mathbf{P}_\ell^{1/2}$, where the columns of $\mathbf{Z}_\ell$ are the channel vectors representing small scale fading between the ℓ-th BS and all users. Accordingly, the diagonal matrix $\mathbf{P}_\ell$ incorporates the transmitted power and the path-loss between each user and the ℓ-th BS. We have the following result.

Proposition 1. *In the noisy and inter-cell interference case, if* $\frac{1}{T}\mathbf{B}^T\mathbf{B} \approx \mathbf{I}$ *and* $\mathbf{Z}_\ell^H\mathbf{Z}_\ell \approx \mathbf{I}$*, then under certain conditions on the relative SNRs of cell-center and cell-edge users (see the Appendix), the optimal solution* $\mathbf{Q}_\ell^\star$ *of problem* (13) *is given by* $\mathbf{Q}_\ell^\star = \mathbf{Z}_\ell\mathbf{V}\mathbf{M}_\ell$*, where* $\mathbf{V}$ *contains the first* K_e *columns of an* $(K_1+K_2)\times(K_1+K_2)$ *identity matrix, and* $\mathbf{M}_\ell$ *is a* $K_e \times K_e$ *non-singular matrix.*

The approximate semi-$\perp$ constraint on the matrix $\mathbf{B}$ posits that the transmitted sequences of different users are approximately orthogonal. For binary signals, this occurs with high probability for large enough T, since the user transmissions are independent. The approximate orthonormality constraint on $\mathbf{Z}_\ell$ requires the number of base station antennas to be greater than the total number of users assigned to both base stations, and is satisfied if, for example, the entries of $\mathbf{Z}_\ell$ are drawn from an i.i.d. zero-mean complex Gaussian distribution with variance $1/M_\ell$ (which is often assumed in the case of rich scattering).

Remark 4. *Recall that, in Theorem 1 we considered the MAX-VAR formulation in* (18)*, and showed that under certain conditions the optimal solution* $\mathbf{G}^*$ *is the subspace containing the cell-edge user signals. In Proposition 1, however, instead of directly estimating the common subspace, we will consider the generalized eigenvalue problem in* (12) *to solve for* $\mathbf{Q}_\ell^*$*, for* $\ell = 1,2$*. Then, we will show in the Appendix that upon applying the resulting* $\mathbf{Q}_\ell^*$ *to the corresponding received signals at BS* ℓ*, we obtain the common subspace corrupted by reduced noise (see* (42) *in the Appendix). It is worth mentioning that when the noise is absent,* (42) *boils down to the result we have in the Theorem.*

Proof. The proof is relegated to Appendix A. □

V. Synchronization

In Section IV, we proposed a learning-based approach that can identify cell-edge user signals. However, this was under the assumption that the received signals from both BSs are perfectly synchronized at the CSPU. One natural question that can be posed is what if there exists a time delay τ_d between $\overline{\mathbf{Y}}_1$ and $\overline{\mathbf{Y}}_2$ at the CSPU. It turns out that our proposed method not only can recover cell-edge user signals in the synchronized case, it can even detect the time difference, τ_d, between the two signals, re-synchronize the signals and then decode them as explained in Section IV.

Assume that the CSPU has received two long sequences $\widetilde{\mathbf{Y}}_1 \in \mathbb{R}^{2M_1 \times \widetilde{T}}$ and $\widetilde{\mathbf{Y}}_2 \in \mathbb{R}^{2M_2 \times \widetilde{T}}$, where $\widetilde{T} > T$, and that the sequence length T is known or has been estimated [45] at the CSPU. The goal is to find the correct delay between the signals $\widetilde{\mathbf{Y}}_1$ and $\widetilde{\mathbf{Y}}_2$ so that we can extract the desired signal $\overline{\mathbf{Y}}_\ell$ from $\widetilde{\mathbf{Y}}_\ell$, and then apply Algorithm 1 to identify cell-edge user signals. Exploiting the fact that communication signals are uncorrelated in time, and thus two copies of the same signal shifted by even one symbol are already uncorrelated, common user signals cannot be extracted via CCA if the $\widetilde{T}$ symbols are misaligned. The correlation coefficient, ρ_n, associated with each pair of canonical directions of $\overline{\mathbf{Q}}_1$ and $\overline{\mathbf{Q}}_2$ will not be at its maximum in this case, $\forall n \in \{1, \cdots, K_e\}$. Based on this key observation, we develop a CCA based algorithm that can re-synchronize and then recover cell-edge user signals.

Define $\widetilde{\mathbf{Y}}_1(\tau_1) := \widetilde{\mathbf{Y}}_1(:, \tau_1 : T + \tau_1 - 1)$ and $\widetilde{\mathbf{Y}}_2(\tau_2) := \widetilde{\mathbf{Y}}_2(:, \tau_2 : T + \tau_2 - 1)$. Furthermore, let us define a search window of size $[w_L, w_R]$ symbols. Upon setting $\tau_2 = 1$, the CSPU solves problem (13) using the signals $\widetilde{\mathbf{Y}}_1(\tau_1)$ and $\widetilde{\mathbf{Y}}_2(\tau_2)$ to obtain $\overline{\mathbf{q}}_1^* := \mathbf{Q}_1^*(:, 1)$ and $\overline{\mathbf{q}}_2^* := \overline{\mathbf{Q}}_2^*(:, 1)$. Then, the CSPU computes and stores the corresponding correlation coefficient ρ_1 between $\widetilde{\mathbf{Y}}_1^T(\tau_1)\overline{\mathbf{q}}_1^*$ and $\widetilde{\mathbf{Y}}_2^T(\tau_2)\overline{\mathbf{q}}_2^*$. If $\tau_2 \leq w_s$, increment τ_2 and repeat, where $w_s := w_R - w_L + 1$ is the window size. Finally, pick the value τ_2^* that gives the highest ρ_1. This procedure is summarized in Algorithm 2.

Remark 5. *Note that as the locations of the T symbols are not generally known, the value of τ_1 is chosen such that $\widetilde{\mathbf{Y}}_1(\tau_1)$ includes a sufficient part of $\overline{\mathbf{Y}}_1$. This is guaranteed with a very high probability as long as $w_s << T$. One possible choice is to set $\tau_1 = \widetilde{T}/2 - T/2$ so that one can assure the existence of enough samples from all users in $\widetilde{\mathbf{Y}}_1(\tau_1)$.*

Algorithm 2 CCA SYNC

Input: $\widetilde{\mathbf{Y}}_1 \in \mathbb{R}^{M_1 \times \widetilde{T}}$, $\widetilde{\mathbf{Y}}_2 \in \mathbb{R}^{M_2 \times \widetilde{T}}$
Initialization: $\tau_1 = \widetilde{T}/2 - T/2, \tau_2 := 1$
while $\tau_2 \leq w_s$ **do**
  Compute ρ_1 after solving (13) using $\widetilde{\mathbf{Y}}_1(\tau_1)$ and $\widetilde{\mathbf{Y}}_2(\tau_2)$
  Store (τ_2, ρ_1) in a stack
  Set $\tau_2 := \tau_2 + 1$
end
Selection: pick the $\tau_2^* := \tau_2$ corresponding to the highest ρ_1.

While Algorithm 2 returns the correct shift, $\tau_d = \tau_2^* - \tau_1$, between the two sequences, the common part in both $\widetilde{\mathbf{Y}}(\tau_1)$ and $\widetilde{\mathbf{Y}}(\tau_2^*)$ is not necessarily of length T since $\widetilde{\mathbf{Y}}(\tau_1)$ may not be equal to $\overline{\mathbf{Y}}_1$. However, from the practical point of view, each user has its own identification sequence as a preamble, so once we know the correct relative delay, we can run algorithm 1 on $\widetilde{\mathbf{Y}}(\tau_1)$ and $\widetilde{\mathbf{Y}}(\tau_2^*)$, and then simply find the sample index at which $\overline{\mathbf{Y}}_1$ starts via correlation with the identification sequence of any of the cell-edge users. It is worth mentioning that the computational complexity of Algorithm 2 is that of solving for the principal component (canonical pair) of (13) a number of times (equal to the search window). The first canonical pair can be cheaply computed via a power iteration.

VI. Experimental Results

To evaluate the performance of our proposed method, we consider a scenario with two hexagonal cells; each with radius $R = 500$ meters. Cell edge users are dropped randomly around the common edge between the two cells, i.e., the locations of cell-edge users were chosen randomly between $0.95R$ and $1.05R$. On the other hand, cell-center users are randomly dropped within distance zR from their serving BS, and we vary the value of z to see the effect of inter-cell interference on the proposed method. The transmitted power β_{kj} is set to 25dBm, $\forall k, j$, i.e., power control is not employed. Furthermore, the transmitted sequence length T is 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 power of cell-edge users. This enables us to see what values of SNR should cell-edge users have to achieve a specific BER. Furthermore, all results are averaged over 1000 channel realizations assuming different user locations in each realization. The uplink channel response vectors $\{\mathbf{h}_{\ell kj}^H\}$ are modeled as

$$\mathbf{h}_{\ell kj}^H = \sqrt{\frac{1}{M_\ell}} \sum_{n=1}^{L} \sqrt{\alpha_{\ell kj}^{(n)}} \mathbf{a}_r(\theta^{(n)})^H \tag{24}$$

where L is the number of paths between the ℓ-th BS and the k-th user in cell j, $\forall \{\ell, j\} \in \{1, 2\}$ and $k \in \{1, \cdots, K_\ell\}$. We use the path-loss model of the urban macro (UMa) scenario from the 3GPP 38.901 standard to compute the complex path gain $\alpha_{\ell kj}^{(n)}$, $\forall n, \ell, j, k$. Cell-center users were allowed to have a line of sight (LOS) path according to the LOS probability in the 3GPP 38.901 standard, however, all cell-edge users were non-LOS. 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

$$\mathbf{a}_r(\theta) = [1, \exp^{ikd\cos(\theta)}, \cdots, \exp^{ikd(M-1)\cos(\theta)}] \tag{25}$$

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 adopted three baselines. First, we implemented zero-forcing successive interference cancellation (ZF-SIC) where the channels of the cell-center users were assumed to be perfectly known at their serving BSs. Specifically, each BS decodes its cell-center users signals using ZF, encodes them again and then subtracts them from its received signal.

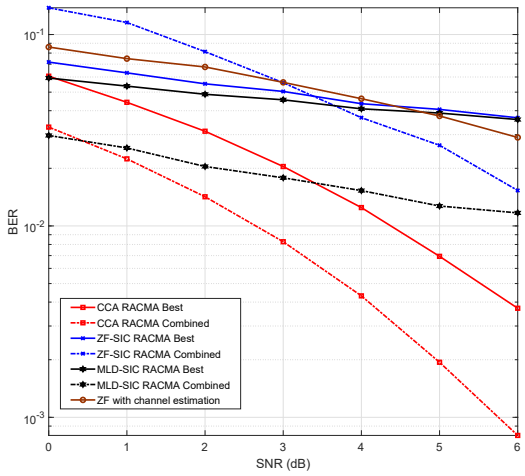


Fig. 2: BER vs. SNR of cell-edge users, $M_1 = M_2 = 10$, $K_1 = K_2 = 8$ and $K_e = 2$, distance of cell-center users $<$ 0.3R

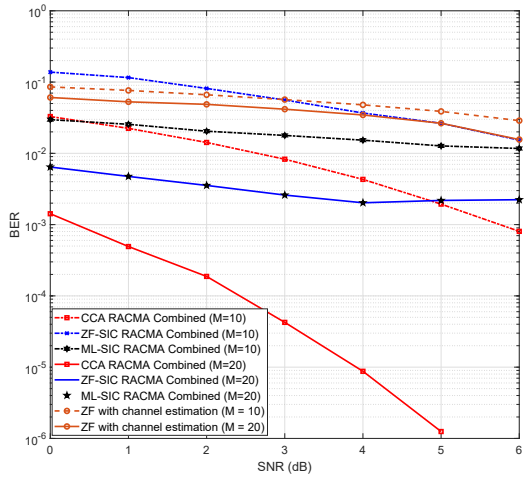


Fig. 3: BER vs. SNR of cell-edge users, $M_1 = M_2 = M$, $K_1 = K_2 = 8$ and $K_e = 2$, distance of cell-center users $<$ 0.3R

Afterwards, the residual signal from each BS will be passed to RACMA [37] in order to identify cell-edge user signals. Finally, the bit error rate (BER) of the cell-edge users is computed at both BSs and the best of the two BERs is reported. Furthermore, in order to guarantee fairness, since we have assumed joint processing of the BSs received signals, both residual signals from both BSs are further sent simultaneously to RACMA and the resulting BER (from RACMA with "double measurements") is also reported. Second, we estimated the channels of cell-center users and cell-edge users via transmitting orthogonal pilot sequences of length 300 each, then we used a ZF detector at each BS and reported the best of the two BERs computed at the two BSs. Third, we implemented maximum likelihood successive interference cancellation (ML-SIC) to decode and subtract cell-center users signals assuming perfect knowledge of their channels at their serving BS. However, since in the worst-case the ML detector requires enumeration over all possible sequences of cell-center users, we only used this baseline when the number of cell-center users is small. The CCA approach (first stage) was implemented in MATLAB, while the MATLAB codes written by A.-J. van der Veen [37] were utilized for the RACMA (second stage) implementation. In a preliminary experiment, we consider a scenario with $K_1 = K_2 = 8$, $M_1 = M_2 = 10$, $K_e = 2$ and cell-center users are dropped randomly up to distance zR, with $z = 0.3$. The numerical results for BER versus SNR of the cell-edge users is shown in Figure 2. It is obvious that our method significantly outperforms ZF-SIC and ML-SIC which assume perfect CSI of the cell-center users, whereas our method does not. For instance, more than one order of magnitude improvement using our CCA-RACMA method is observed at SNR= 6dB. Furthermore, the bad performance of ZF with channel estimation reflects how the inaccurate channel estimates of cell-edge users severely degrade their detection performance.

In order to see the effect of increasing the number of

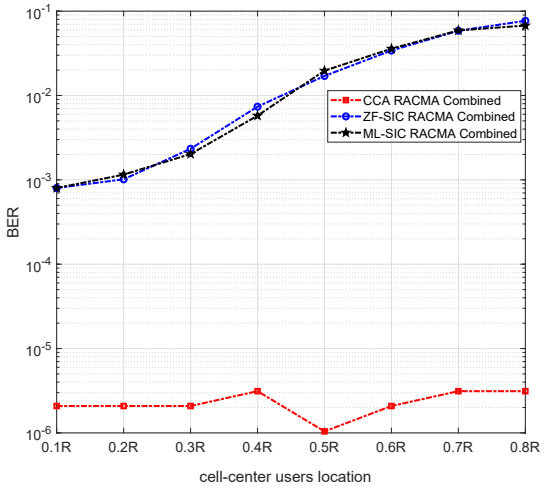


Fig. 4: BER vs. distance of cell-center users from their serving BS, with $M_1 = M_2 = 20$, $K_1 = K_2 = 8$ and $K_e = 2$, SNR $= 5$dB

antennas on the performance of the proposed method, we considered the same setting of the previous experiment, however, we increased the number of antennas at each base station to 20, i.e., $M_1 = M_2 = 20$. Figure 3 shows that doubling the number of antennas at each base station improves the BER of cell-edge users obtained by all methods. However, a significant improvement gap in the BER obtained by our "blind" method is observed compared to that of ZF-SIC and ML-SIC. For instance, while ZF-SIC achieves an order of magnitude reduction in BER with $M = 20$, CCA-RACMA attains more than three orders of magnitudes improvement in BER at SNR $= 5$dB. Furthermore, Figure 3 shows that our approach does yield measurable BER when the SNR of cell-edge users exceeds 5dB. The reason is that CCA can aggressively suppress the inter-cell interference when the

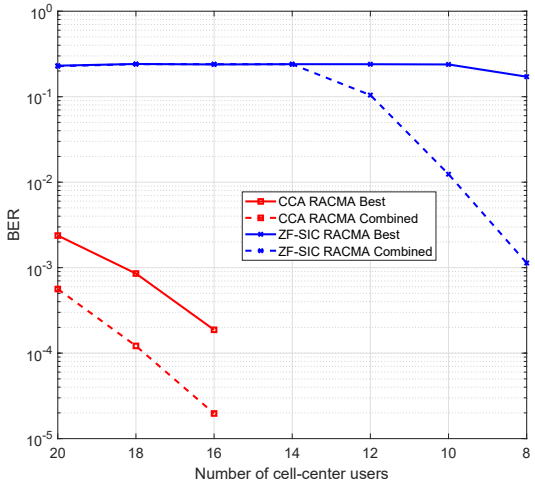


Fig. 5: BER vs. number of cell-center users at each BS, $M_1 = M_2 = 30$, SNR $= 4$dB and $K_e = 2$, distance of cell-center users < 0.6R

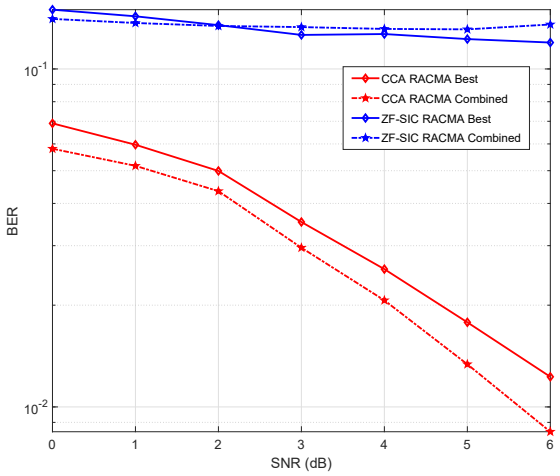


Fig. 6: BER vs. SNR, with $M_1 = M_2 = 25$, $K_1 = K_2 = 15$ and $K_e = 3$, distance of cell-center users < 0.8R

number of antennas exceeds the total number of users, as explained in Appendix A.

To test the effect of inter-cell interference, we vary the locations of cell-center users in their cell from $0.1R$ to $0.8R$, and for each setting we measure the BER attained by all methods at SNR $= 5$dB. Figure 4 demonstrates that the proposed CCA-RACMA approach still exhibits a favorable performance relative to that of ZF-SIC and ML-SIC. In particular, two orders of magnitude increase in the BER attained by ZF-SIC and ML-SIC is observed when the cell-center users are spread up to $0.8R$ compared to $0.1R$ from their serving BS, however, a very slight degradation in the performance of CCA-RACMA is observed, even for high spreads. Notice that, while the two baselines assume perfect knowledge of the cell-center user channels, this assumption becomes less realistic when "cell-center" users are in fact fully scattered throughout the cell. This therefore give a big advantage to the baselines over our method; notwithstanding, our method still works the best, even in this case.

We now consider another experiment with $M_1 = M_2 = 30$, $K_e = 2$ and SNR $= 5$dB. Assuming fixed user positions, we vary the number of cell-center users in each cell from 20 to 8, and for each given number of cell-center users we compute the BER of cell-edge users. In this experiment, all cell-center users are randomly dropped up to distance $0.5R$ from their serving BS. In Figure 5, we observe that our proposed blind method can attain BER that is below the detectable threshold for this simulation when the number of cell-center users per cell is less than 16 while the ZF-SIC detector is severely affected by the cancellation errors from cell-center users. This shows how the proposed approach can handle dense scenarios, and hence, it is expected to work well in the case of multiple BSs (more than two).

We next consider $M_1 = M_2 = M = 25$, $K_1 = K_2 = 15$, $Ke = 3$ and cell-center users are randomly located at distance less than $0.8R$ from their serving BS. As shown in Figure 6,

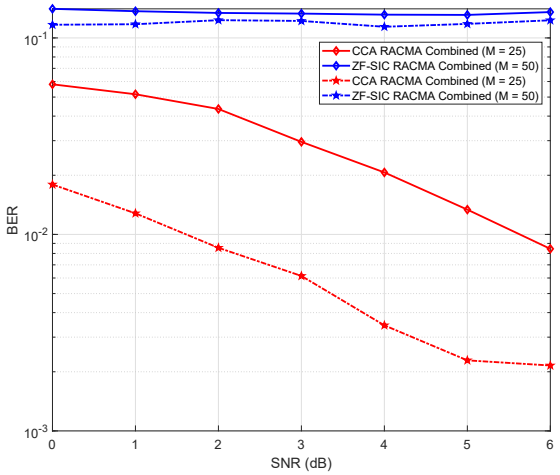


Fig. 7: BER vs. SNR, with $M_1 = M_2 = M$, $K_1 = K_2 = 15$ and $K_e = 3$, distance of cell-center users < 0.8R

jointly injecting more users (cell-center and cell-edge) and allowing them to be more spread, yields a noticeable degradation in the BER of cell-edge users achieved by all 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. On the other hand, our method also exhibits some degradation in the performance because adding more users creates more intercell 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 cell-center CSI. For example, our method still has more than an order of magnitude lower BER at different SNR values.

We further simulate the previous scenario with double the number of antennas at each base station. As Figure 7 depicts, doubling the number of antennas at each base station yields an

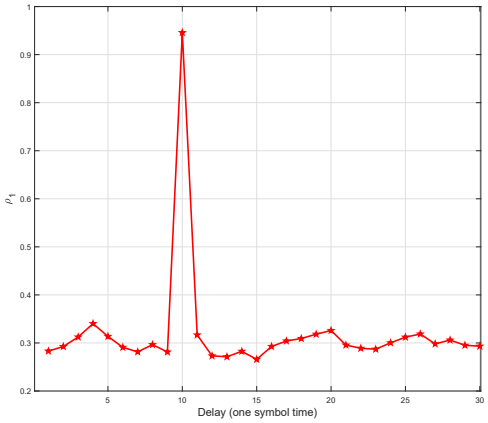


Fig. 8: Correlation coefficient of the first pair-wise canonical component ρ_1 vs. delay

order of magnitude improvement in the BER of our method, while only slightly improving the BER of ZF-SIC.

Finally, to show how CCA can still detect cell-edge user signals even when the received signals at the two BSs are not perfectly synchronized, we consider a scenario with $K_1 = K_2 = 8$, $M_1 = M_2 = 10$, $K_e = 2$, SNR $= 3$dB, $T = 800$ and cell-center users are dropped randomly up to distance $0.5R$, and we assume that the received signal at the ℓ-th BS is $\widetilde{\mathbf{Y}}_\ell \in \mathbb{R}^{M_\ell \times \widetilde{T}}$, where $\widetilde{T}$ was set to 830. Then, we applied Algorithm 2 on $\widetilde{\mathbf{Y}}_1$ and $\widetilde{\mathbf{Y}}_2$, and observed the correlation coefficient of the first pair-wise canonical components as a function of the relative shift. Figure 8 shows how CCA can clearly identify the correct delay, and hence, detect cell-edge user signals as explained before. Clearly when the BS signals are not synchronized, there is no meaningful common subspace – even the first pair of canonical components exhibits low correlation. When we hit the correct delay, on the other hand, there are common components and the correlation coefficient ρ_1 is very high, as shown in Figure 8.

VII. Conclusion and Future Work

This paper has considered cell-edge user signal detection in the uplink of a multi-cell multi-user MIMO system. The goal is to design a detector that can reliably demodulate cell-edge user signals in the presence of strong intra-cell interference from users close to the base station, without resorting to power control or scheduling algorithms that throttle the cell-center user rates. This paper proposed a two-stage approach that leverages base stations cooperation to reliably identify cell-edge user signals at low SNR, without even knowing their channels. First, two-view CCA was brought in to estimate the subspace containing the cell-edge user signals shared by both base stations under the assumption that BS signals are synchronized. Then, an efficient analytical method called RACMA that guarantees the identifiability of binary signals from well-conditioned mixtures was utilized to extract the cell-edge user signals from the resulted mixture. We presented theoretical analysis of common subspace identifiability, in both ideal and realistic scenarios that include noise and inter-cell interference.

Furthermore, we developed an algorithm that can identify cell-edge user signals in the case when BS signals are not fully synchronized at the CSPU. Extensive simulations using a realistic path-loss model were carried out to show the superiority of the proposed learning-based method. It was shown that our blind CCA method achieves more than an order of magnitude improvement in the cell-edge user BER compared to the 'oracle' zero forcing and maximum likelihood cell-center multiuser detection followed by interference cancellation of the cell-center users before detecting the cell-edge users.

In the future, it is interesting to study the more general setting of the problem that considers (possibly many) more than two 'entangled' base stations. This introduces much more interference on cell-edge users that renders the problem much more challenging. In such scenarios, one can resort to generalized canonical correlation analysis to detect unknown cell-edge users whose signals are received with relatively equal power at multiple BSs. More importantly, one can investigate how many base stations should cooperate to provide the best detection performance for cell-edge users, and construct a performance-complexity blueprint. In addition, it is crucial to develop an efficient algorithm that can resolve the synchronization issue in the case of cooperation amongst a large number of BSs.

Appendix A
Proof of Proposition 1

In order to see how CCA can identify cell-edge user signals in the noisy and inter-cell interference case, let us first rewrite the received signal at the ℓ-th BS as

$$\mathbf{Y}_\ell = \mathbf{H}_{\ell p_\ell}\mathbf{S}_{p_\ell}^T + \mathbf{H}_{\ell c}\mathbf{S}_c^T + \mathbf{H}_{\ell p_j}\mathbf{S}_{p_j}^T + \mathbf{W}_\ell \qquad (26)$$

where $\ell, j \in \{1,2\}$ and $j \neq \ell$. Recall that, from (14) and (15), one can easily see that the channel matrix $\mathbf{H}_{\ell p_\ell}$ in (26) can be factored into $\mathbf{Z}_{\ell p_\ell}\mathbf{P}^{1/2}$ where $\mathbf{Z}_{\ell p_\ell} \in \mathbb{C}^{M_\ell \times (K_\ell - K_{e_\ell})}$ holds in its columns the small-scale fading vectors defined in (14) while $\mathbf{P}_{\ell p_\ell} \in \mathbb{R}^{K_\ell - K_{e_\ell}}$ is a diagonal matrix whose entries model the received signal power (product of path loss and transmitted signal power) for each of the cell-center users served by the ℓ-th BS, and likewise for $\mathbf{H}_{\ell c}$ and $\mathbf{H}_{\ell p_j}$. Therefore, (26) can be equivalently written as

$$\mathbf{Y}_\ell = \mathbf{Z}_{\ell p_\ell}\mathbf{P}_{\ell p_\ell}^{1/2}\mathbf{S}_{p_\ell}^T + \mathbf{Z}_{\ell c}\mathbf{P}_{\ell p_c}^{1/2}\mathbf{S}_c^T + \mathbf{Z}_{\ell p_j}\mathbf{P}_{\ell p_j}^{1/2}\mathbf{S}_{p_j}^T + \mathbf{W}_\ell \quad (27)$$

Let us first consider a simple scenario with two cell-center users (one at each BS) and one cell-edge user located at the common edge between the two BSs. We define β_p, β_e and β_f as the received signal power (RSP) of the ℓ-th cell-center user at the ℓ-th BS, the RSP of the cell-edge user at the ℓ-th BS, and the RSP of the j-th cell-center user at the ℓ-th BS, respectively, for $\ell \neq j$. Furthermore, for the sake of simplicity, we assume here that the cell-edge user signal is received with equal power at both BSs, i.e., the cell-edge user is exactly on

the edge between the two BSs. Then (27) can be expressed as

$$\mathbf{Y}_\ell = \mathbf{Z}_\ell \mathbf{P}_\ell^{1/2} \mathbf{B}^T + \mathbf{W}_\ell \tag{28}$$

where $\mathbf{B} = [\mathbf{s}_e, \mathbf{s}_{p_1}, \mathbf{s}_{p_2}]$ holds in its columns the temporal signals of the three users, $\mathbf{P}_1 = \text{Diag}([\beta_e, \beta_p, \beta_f])$, $\mathbf{P}_2 = \text{Diag}([\beta_e, \beta_f, \beta_p])$ and $\mathbf{B} = [\mathbf{s}_e, \mathbf{s}_{p_1}, \mathbf{s}_{p_2}]$, where $\mathbf{D} = \text{Diag}(\mathbf{d})$ is a diagonal matrix with the vector $\mathbf{d}$ on its diagonal. The entries of $\mathbf{Z}_\ell$ represent the small scale fading between each user and the antennas at BS ℓ. $\mathbf{Z}_\ell$ is modeled as i.i.d. circularly symmetric zero-mean Gaussian with variance $1/M_\ell$ (corresponding to a rich scattering scenario).

We will now compute the cross- and auto-correlation matrices $\mathbf{R}_{\mathbf{y}_1\mathbf{y}_2}$, $\mathbf{R}_{\mathbf{y}_\ell\mathbf{y}_\ell}$ as follows. Since the cross correlation matrix, $\mathbf{R}_{\mathbf{y}_1\mathbf{y}_2}$, is given by $\frac{1}{T}\mathbf{Y}_1\mathbf{Y}_2^H$, then it follows that $\mathbf{R}_{\mathbf{y}_1\mathbf{y}_2}$ is given by

$$\begin{aligned}\mathbf{R}_{\mathbf{y}_1\mathbf{y}_2} &= \frac{1}{T}(\mathbf{Z}_1\mathbf{P}_1^{1/2}\mathbf{B}^T + \mathbf{W}_1)(\mathbf{Z}_2\mathbf{P}_2^{1/2}\mathbf{B} + \mathbf{W}_2)^H \\ &= \mathbf{Z}_1\mathbf{P}_{12}\mathbf{H}_2^H\end{aligned} \tag{29}$$

where $\mathbf{P}_{12} = (\mathbf{P}_2\mathbf{P}_1)^{1/2}$. Note that, in (29), in addition to the assumption that $\frac{1}{T}\mathbf{B}^T\mathbf{B} = \mathbf{I}$, we exploited the fact that, for large T, $\frac{1}{T}\mathbf{W}_\ell\mathbf{W}_j^H \approx 0$ and $\frac{1}{T}\mathbf{B}^T\mathbf{W}_j^H \approx 0$, for $j, \ell \in \{1, 2\}$. Similarly, the auto-correlation matrix of the received signal of the ℓ-th BS can be expressed as

$$\mathbf{R}_{\mathbf{y}_\ell\mathbf{y}_\ell} = \mathbf{Z}_\ell\mathbf{P}_\ell\mathbf{Z}_\ell^H + \sigma^2\mathbf{I} \tag{30}$$

Now, we substitute with (29) and (30) in (9) to obtain

$$\begin{aligned}\mathbf{Z}_1\mathbf{P}_{12}\mathbf{Z}_2^H(\mathbf{Z}_2\mathbf{P}_2\mathbf{Z}_2^H + \sigma^2\mathbf{I})^{-1}\mathbf{Z}_2\mathbf{P}_{12}\mathbf{Z}_1^H\mathbf{q}_1 \\ = \lambda^2(\mathbf{Z}_1\mathbf{P}_1\mathbf{Z}_1^H + \sigma^2\mathbf{I})\mathbf{q}_1\end{aligned} \tag{31}$$

which can be equivalently written as

$$\begin{aligned}\mathbf{Z}_1\mathbf{\Gamma}_{12}\mathbf{Z}_2^H(\mathbf{Z}_2\mathbf{\Gamma}_2\mathbf{Z}_2^H + \mathbf{I})^{-1}\mathbf{Z}_2\mathbf{\Gamma}_{12}\mathbf{Z}_1^H\mathbf{q}_1 \\ = \lambda^2(\mathbf{Z}_1\mathbf{\Gamma}_1\mathbf{Z}_1^H + \mathbf{I})\mathbf{q}_1\end{aligned} \tag{32}$$

where $\mathbf{\Gamma}_1 = \text{Diag}([\gamma_e, \gamma_p, \gamma_f])$, $\mathbf{\Gamma}_2 = \text{Diag}([\gamma_e, \gamma_f, \gamma_p])$ and $\mathbf{\Gamma}_{12} = (\mathbf{\Gamma}_2\mathbf{\Gamma}_1)^{1/2}$, with $\gamma_e = \beta_e/\sigma^2$ be the received SNR of the cell-edge user, $\gamma_p = \beta_p/\sigma^2$ be the received SNR of each cell-center user at its serving BS, and $\gamma_f = \beta_f/\sigma^2$ be the received SNR of each cell-center at the other (non-serving) BS. By left multiplying the two sides of (32) by $\mathbf{H}_1^\dagger$, we obtain

$$\begin{aligned}\mathbf{\Gamma}_{12}\mathbf{Z}_2^H(\mathbf{Z}_2\mathbf{\Gamma}_2\mathbf{Z}_2^H + \mathbf{I})^{-1}\mathbf{Z}_2\mathbf{\Gamma}_{12}\mathbf{Z}_1^H\mathbf{q}_1 \\ = \lambda^2(\mathbf{\Gamma}_1\mathbf{Z}_1^H + \mathbf{Z}_1^\dagger)\mathbf{q}_1\end{aligned} \tag{33}$$

By substituting with $\mathbf{Z}_1^\dagger = (\mathbf{Z}_1^H\mathbf{Z}_1)^{-1}\mathbf{Z}_1^H$, and by letting $\mathbf{v} = \mathbf{Z}_1^H\mathbf{q}_1$, (33) can be expressed as

$$\mathbf{\Gamma}_{12}\mathbf{Z}_2^H(\mathbf{Z}_2\mathbf{\Gamma}_2\mathbf{Z}_2^H + \mathbf{I})^{-1}\mathbf{Z}_2\mathbf{\Gamma}_{12}\mathbf{v} = \lambda^2(\mathbf{\Gamma}_1 + (\mathbf{Z}_1^H\mathbf{Z}_1)^{-1})\mathbf{v} \tag{34}$$

By defining the matrix $\mathbf{Z} := \mathbf{Z}_2^H(\mathbf{Z}_2\mathbf{\Gamma}_2\mathbf{Z}_2^H + \mathbf{I})^{-1}\mathbf{Z}_2$, it then follows that $\mathbf{Z}$ can be simplified as

$$\begin{aligned}\mathbf{Z} &= \mathbf{Z}_2^H(\mathbf{Z}_2\mathbf{\Gamma}_2\mathbf{Z}_2^H + \mathbf{I})^{-1}\mathbf{Z}_2 && \text{(35a)}\\ &= \mathbf{Z}_2^H(\mathbf{Z}_2^\dagger(\mathbf{Z}_2\mathbf{\Gamma}_2\mathbf{Z}_2^H + \mathbf{I}))^\dagger && \text{(35b)}\\ &= \mathbf{Z}_2^H(\mathbf{Z}_2^H)^\dagger(\mathbf{\Gamma}_2 + (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1})^{-1} && \text{(35c)}\\ &= (\mathbf{\Gamma}_2 + (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1})^{-1} && \text{(35d)}\end{aligned}$$

Note that in (35b) and (35c), we have exploited the following two properties of the pseudoinverse

P 1. For any square matrix $\mathbf{A}$, if $\mathbf{A}$ is invertible, its pseudoinverse is its inverse, i.e., $\mathbf{A}^\dagger = \mathbf{A}^{-1}$

P 2. $(\mathbf{BA})^\dagger = \mathbf{A}^\dagger\mathbf{B}^\dagger$

By substituting with (35d) in (34), we obtain

$$\mathbf{\Gamma}_{12}(\mathbf{\Gamma}_2 + (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1})^{-1}\mathbf{\Gamma}_{12}\mathbf{v} = \lambda^2(\mathbf{\Gamma}_1 + (\mathbf{Z}_1^H\mathbf{Z}_1)^{-1})\mathbf{v} \tag{36}$$

which can be equivalently expressed as

$$\mathbf{F}\mathbf{v} = \lambda^2\mathbf{v} \tag{37}$$

where $\mathbf{F}$:$=(\mathbf{\Gamma}_1 + (\mathbf{Z}_1^H\mathbf{Z}_1)^{-1})^{-1}\mathbf{\Gamma}_{12}(\mathbf{\Gamma}_2 + (\mathbf{Z}_2^H\mathbf{Z}_2)^{-1})^{-1}\mathbf{\Gamma}_{12}$ is an $K_s \times K_s$ matrix, and $K_s = 3$ for the particular scenario considered here. For ease of exposition, we will assume here that the number of antennas M_ℓ is large enough so that $(\mathbf{Z}_\ell^H\mathbf{Z}_\ell)^{-1}$ is approximately identity. Thus, matrix $\mathbf{F}$ can be expressed as

$$\mathbf{F} := \begin{bmatrix} (\frac{\gamma_e}{\gamma_e+1})^2 & 0 & 0 \\ 0 & \frac{\gamma_f\gamma_c}{(\gamma_f+1)(\gamma_c+1)} & 0 \\ 0 & 0 & \frac{\gamma_f\gamma_c}{(\gamma_f+1)(\gamma_c+1)} \end{bmatrix} \in \mathbb{R}^{K_s \times K_s},$$

If each cell-center user is close to its serving BS, then $\gamma_f << 1$ and $\gamma_c >> 1$. Therefore, the term $\frac{\gamma_f\gamma_c}{(\gamma_f+1)(\gamma_c+1)}$ will be approximately equal to γ_f. Then, it can be easily seen that the maximum eigenvalue of the matrix $\mathbf{F}$ is equal to $(\frac{\gamma_e}{\gamma_e+1})^2$ and the other two eigenvalues will be approximately equal to γ_f. Since the maximum eigenvalue of the matrix $\mathbf{F}$ is nothing but the square of the correlation coefficient that is associated with the vectors $\mathbf{Y}_1^T\mathbf{q}_1$ and $\mathbf{Y}_2^T\mathbf{q}_2$. Then, it turns out that the maximum correlation coefficient is given by

$$\rho_{\max} = \frac{\gamma_e}{\gamma_e + 1} \tag{38}$$

Now, we need to compute the eigenvectors $\mathbf{q}_1$ and $\mathbf{q}_2$. Since the maximum eigenvector of the diagonal matrix $\mathbf{F}$ is given by

$$\mathbf{v} = [\pm 1, 0, 0]^T \tag{39}$$

the eigenvector $\mathbf{q}_1$ can be obtained by solving the following system of linear equations

$$\mathbf{v} = \mathbf{Z}_1^H\mathbf{q}_1 \tag{40}$$

Without loss of generality, we can let $\mathbf{q}_1^* = \mathbf{Z}_1(\mathbf{Z}_1^H\mathbf{Z}_1)^{-1}\mathbf{v}$. The reason is that we can always find two components to the vector $\mathbf{q}_1^*$; one in the subspace spanned by $\mathbf{Z}_1$ and one orthogonal to it, however, the latter will vanish after multiplication with $\mathbf{Z}_1^H$. By substituting with $\mathbf{q}_1^*$ in (8), it can be easily proved that the corresponding canonical component of the second view $\mathbf{q}_2^* = \mathbf{Z}_2(\mathbf{Z}_2^H\mathbf{Z}_2)^{-1}\mathbf{v}$. Define $\hat{\mathbf{s}}_{\ell c} := \mathbf{Y}_\ell^H\mathbf{q}_\ell^*$ and substitute with $\mathbf{q}_\ell^*$, we get the following

$$\hat{\mathbf{s}}_{\ell c} = \sqrt{\beta_e}c\mathbf{s}_c + \mathbf{n}_\ell \tag{41}$$

where $\mathbf{n}_\ell = \mathbf{W}_\ell^H\mathbf{q}_\ell^* \in \mathbb{C}^T$ and $c = \pm 1$. This means that, in the case of single cell-edge user, the proposed detector can efficiently recover cell-edge user signals at low SNR even in the presence of inter-cell interference.

The generalization to $K_e > 1$ and $K_\ell - K_e > 1$ now follows directly. In that case, the matrix F will have the vector $\mathbf{f} \in \mathbb{R}^{K_s}$ on its diagonal, where

$$\mathbf{f}(j) = \begin{cases} (\frac{\gamma_{e_j}}{\gamma_{e_j}+1})^2, & j \in \{1, \cdots, K_e\} \\ \frac{\gamma_{f_j}\gamma_{p_j}}{(\gamma_{f_j}+1)(\gamma_{p_j}+1)} & j \in \{K_e+1, \cdots, K_s\} \end{cases}$$

Assume that $\gamma_{f_j} << 1, \forall j \in \{K_e+1, \cdots, K_s\}$. Then it can be easily seen that the largest K_e eigen vectors are the first K_e columns of an $K_s \times K_s$ identity matrix. Upon letting $\mathbf{V} = \mathbf{I}(:, 1:K_e)$, the optimal solution $\mathbf{Q}_\ell^* = \mathbf{H}_\ell(\mathbf{Z}_\ell^H\mathbf{Z}_\ell)^{-1}\mathbf{V}\mathbf{M}_\ell$, where $\mathbf{M}_\ell$ is any $K_e \times K_e$ non singular matrix that satisfies the ℓ-th orthonormality constraint in (13). Define $\hat{\mathbf{S}}_{\ell c} := \mathbf{Y}_\ell^H\mathbf{Q}_\ell^*$ and substitute in (27), we obtain

$$\hat{\mathbf{S}}_{\ell c} = \mathbf{S}_c\mathbf{P}_c^{1/2}\mathbf{M}_\ell + \mathbf{N}_\ell \tag{42}$$

where $\mathbf{P}_c = \text{Diag}([\beta_{e_1}, \cdots, \beta_{e_{K_e}}])$, and $\mathbf{N}_\ell = \mathbf{W}_\ell^H\mathbf{Q}_\ell^*$. Note that, after obtaining $\hat{\mathbf{S}}_{\ell c}$, we pass it to RACMA to identify the cell-edge user signals $\mathbf{S}_c$.

REFERENCES

[1] M. S. Ibrahim and N. Sidiropoulos, "Cell-edge interferometry: Reliable detection of unknown cell-edge users via canonical correlation analysis," in *IEEE International Conference on Signal Processing Advances in Wireless Communications (SPAWC)*, France, July 2019, pp. 1–5.

[2] S. G. Kiani, G. E. Oien, and D. Gesbert, "Maximizing multicell capacity using distributed power allocation and scheduling," in *IEEE Wireless Communications and Networking Conference (WCNC)*, China, March 2007, pp. 1690–1694.

[3] A. Tolli, M. Codreanu, and M. Juntti, "Cooperative MIMO-OFDM cellular system with soft handover between distributed base station antennas," *IEEE Transactions on Wireless Communications*, vol. 7, no. 4, pp. 1428–1440, 2008.

[4] S. Verdu *et al.*, *Multiuser detection*. Cambridge University Press, 1998.

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

[6] H. Vikalo and B. Hassibi, "The expected complexity of sphere decoding, part I: Theory, part II: Applications," *IEEE Transactions on Signal Processing*, vol. 53, no. 8, pp. 2819–2834, 2005.

[7] 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.

[8] P. H. Tan and L. K. Rasmussen, "The application of semidefinite programming for detection in CDMA," *IEEE Journal on Selected Areas in Communications*, vol. 19, no. 8, pp. 1442–1449, 2001.

[9] M. Abdi, H. El Nahas, A. Jard, and E. Moulines, "Semidefinite positive relaxation of the maximum-likelihood criterion applied to multiuser detection in a CDMA context," *IEEE Signal Processing Letters*, vol. 9, no. 6, pp. 165–167, 2002.

[10] Z. Luo, W. Ma, A. M. So, Y. Ye, and S. Zhang, "Semidefinite relaxation of quadratic optimization problems," *IEEE Signal Processing Magazine*, vol. 27, no. 3, pp. 20–34, May 2010.

[11] 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.

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

[13] A. Duel-Hallen, "Decorrelating decision-feedback multiuser detector for synchronous code-division multiple-access channel," *IEEE Transactions on Communications*, vol. 41, no. 2, pp. 285–290, 1993.

[14] P. D. Alexander, M. C. Reed, J. A. Asenstorfer, and C. B. Schlegel, "Iterative multiuser interference reduction: Turbo CDMA," *IEEE Transactions on Communications*, vol. 47, no. 7, pp. 1008–1014, 1999.

[15] 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.

[16] 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.

[17] S. Agarwal, S. De, S. Kumar, and H. M. Gupta, "Qos-aware downlink cooperation for cell-edge and handoff users," *IEEE Transactions on Vehicular Technology*, vol. 64, no. 6, pp. 2512–2527, 2014.

[18] 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.

[19] H. R. Chayon, K. B. Dimyati, H. Ramiah, and A. W. Reza, "Enhanced quality of service of cell-edge user by extending modified largest weighted delay first algorithm in LTE networks," *Symmetry*, vol. 9, no. 6, p. 81, 2017.

[20] H. Xu and P. Ren, "Joint user scheduling and power control for cell-edge performance improvement in backhaul-constrained network MIMO," in *IEEE 24th Annual International Symposium on Personal, Indoor, and Mobile Radio Communications (PIMRC)*, London, UK, Sept. 2013, pp. 1342–1346.

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

[22] 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.

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

[24] 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 (ICASSP)*, Taiwan, Apr. 2009, pp. 2113–2116.

[25] 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.

[26] 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.

[27] 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.

[28] Z. Bai, G. Huang, and L. Yang, "A radar anti-jamming technology based on canonical correlation analysis," in *International Conference on Neural Networks and Brain*, vol. 1, China, Oct. 2005, pp. 9–12.

[29] 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.

[30] 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.

[31] W. Liu, D. P. Mandic, and A. Cichocki, "Analysis and online realization of the CCA approach for blind source separation," *IEEE Transactions on Neural Networks*, vol. 18, no. 5, pp. 1505–1510, 2007.

[32] 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.

[33] 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.

[34] 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 *IEEE 16th International Conference on Data Mining (ICDM)*, Dec 2016, pp. 871–876.

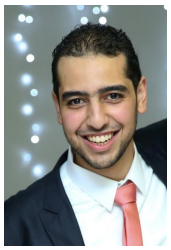
[35] X. Fu, K. Huang, M. Hong, N. D. Sidiropoulos, and A. M. So, "Scalable and flexible multiview MAX-VAR canonical correlation analysis," *IEEE Transactions on Signal Processing*, vol. 65, no. 16, pp. 4150–4165, Aug 2017.

[36] 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.

[37] 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.

[38] 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.

[39] T. Mayer, H. Jenkac, and J. Hagenauer, "Turbo base-station cooperation for intercell interference cancellation," in *IEEE International Conference on Communications (ICC)*, vol. 11, Turkey, June 2006, pp. 4977–4982.
[40] C. Hoymann, L. Falconetti, and R. Gupta, "Distributed uplink signal processing of cooperating base stations based on IQ sample exchange," in *IEEE International Conference on Communications (ICC)*, Germany, June 2009, pp. 1–5.
[41] E. Björnson, E. G. Larsson, and M. Debbah, "Massive MIMO for maximal spectral efficiency: How many users and pilots should be allocated?" *IEEE Transactions on Wireless Communications*, vol. 15, no. 2, pp. 1293–1308, 2016.
[42] X. Li, E. Bjornson, E. G. Larsson, S. Zhou, and J. Wang, "A multi-cell MMSE detector for massive MIMO systems and new large system analysis," in *IEEE Global Communications Conference (GLOBECOM)*, CA, USA, Dec. 2015, pp. 1–6.
[43] H. Q. Ngo, M. Matthaiou, and E. G. Larsson, "Performance analysis of large scale MU-MIMO with optimal linear receivers," in *Swedish Communication Technologies Workshop (Swe-CTW)*, Sweden, Oct. 2012, pp. 59–64.
[44] P. Horst, "Generalized canonical correlations and their applications to experimental data," *Journal of Clinical Psychology*, vol. 17, no. 4, pp. 331–347, 1961.
[45] M. Wax and T. Kailath, "Detection of signals by information theoretic criteria," *IEEE Transactions on Acoustics, Speech, and Signal Processing*, vol. 33, no. 2, pp. 387–392, 1985.



Mohamed Salah Ibrahim received the B.Sc. (with highest honors) in Electrical Engineering from Alexandria University, Alexandria, Egypt, in 2013, and the M.Sc. degree in wireless technologies from Nile University, Giza, Egypt, in 2016. In July 2017, he earned the M.E. degree in Mathematics and Engineering Physics, Faculty of Engineering, Alexandria, Egypt. Since September 2017, he has been working towards his Ph.D. degree in the Department of Electrical and Computer Engineering, University of Virginia, Charlottesville, Virginia, USA. He was a finalist of the Best Student Paper Competition at IEEE SPAWC 2018. His research interests include signal processing, optimization, machine learning and wireless communications.

Nicholas D. Sidiropoulos (F'09) earned the Diploma in Electrical Engineering from Aristotle University of Thessaloniki, Greece, and M.S. and Ph.D. degrees in Electrical Engineering from the University of Maryland at College Park, in 1988, 1990 and 1992, respectively. He has served on the faculty of the University of Virginia, University of Minnesota, and the Technical University of Crete, Greece, prior to his current appointment as Louis T. Rader Professor and Chair of ECE at UVA. From 2015 to 2017 he was an ADC Chair Professor at the University of Minnesota. His research interests are in signal processing, communications, optimization, tensor decomposition, and factor analysis, with applications in machine learning and communications. He received the NSF/CAREER award in 1998, the IEEE Signal Processing Society (SPS) Best Paper Award in 2001, 2007, and 2011, served as IEEE SPS Distinguished Lecturer (2008-2009), and as Vice President - Membership of IEEE SPS (2017-2019). He received the 2010 IEEE Signal Processing Society Meritorious Service Award, and the 2013 Distinguished Alumni Award from the University of Maryland, Dept. of ECE. He is a Fellow of IEEE (2009) and a Fellow of EURASIP (2014).