

Optimizing Multi-Cell Massive MIMO for Spectral Efficiency: How Many Users Should Be Scheduled?

Emil Björnson^{*†‡}, Erik G. Larsson^{*}, and Mérouane Debbah[‡]

^{*}Department of Electrical Engineering (ISY), Linköping University, Linköping, Sweden

[†]SUPELEC, Gif-sur-Yvette, France, and [‡]ACCESS, Dept. of Signal Processing, KTH, Stockholm, Sweden

Abstract—Massive MIMO is a promising technique to increase the spectral efficiency of cellular networks, by deploying antenna arrays with hundreds or thousands of active elements at the base stations and performing coherent beamforming. A common rule-of-thumb is that these systems should have an order of magnitude more antennas, N , than scheduled users, K , because the users' channels are then likely to be quasi-orthogonal. However, it has not been proved that this rule-of-thumb actually maximizes the spectral efficiency. In this paper, we analyze how the optimal number of scheduled users, K^* , depends on N and other system parameters. The value of K^* in the large- N regime is derived in closed form, while simulations are used to show what happens at finite N , in different interference scenarios, and for different beamforming.

Index Terms—Massive MIMO, pilot contamination, user scheduling.

I. INTRODUCTION

Cellular communication networks are continuously evolving to keep up with the rapidly increasing demand for wireless data services. Higher area throughput (in bit/s per unit area) has traditionally been achieved by a combination of three multiplicative factors [1]: more frequency spectrum (Hz), higher cell density (more cells per unit area), and higher spectral efficiency (bit/s/Hz/cell). This paper considers the latter and especially the massive multiple-input multiple-output (MIMO) concept, proposed in [2], which is identified as a key to increase spectral efficiency (SE) by orders of magnitude [3]–[5].

The massive MIMO concept is based on equipping base stations (BSs) with hundreds or thousands of antenna elements which, unlike conventional cellular technology, are operated in a coherent fashion. This can provide unprecedented array gains and a spatial resolution that allows for multi-user MIMO communication to tens or hundreds of user equipments (UEs) per cell, while maintaining robustness to inter-user interference. The research on massive MIMO has so far focused on establishing the fundamental physical layer properties; in particular, that the acquisition of channel state information (CSI) is limited by the channel coherence and how this impacts SEs and the ability to mitigate inter-cell interference [2], [6], [7]. The improvements in energy efficiency were analyzed in [8], [9], while [10], [11] predict that the hardware impairments of practical transceivers has small impact on the SE. In contrast, the research community has only briefly touched on resource allocation problems in the medium access control layer (e.g., user scheduling)—although the truly achievable SE can only be understood if the layers are jointly optimized.

The importance of resource allocation for massive MIMO was described in [12], where initial guidelines were given. A main insight is that the limited number of orthogonal pilot sequences need to be allocated among the UEs to limit interference—by capitalizing on pathloss differences [13], [14] and spatial correlation [15], [16].

In this paper, we consider a related resource allocation question: *how many UEs should be scheduled per cell to maximize the spectral efficiency?* This question has, to the best of our knowledge, not been

answered for multi-cell systems.¹ We show how the coherence block length, number of antennas, and other system parameters determine the answer. To this end, we derive new SE expressions for the uplink with power control and for two different beamforming schemes.

II. SYSTEM MODEL

This paper considers a cellular network with universal time and frequency reuse. Each cell is given an index in the set \mathcal{B} , where the cardinality $|\mathcal{B}|$ is the number of cells. The BS in each cell is equipped with an array of N antennas and communicates in the uplink with K single-antenna UEs in the cell. We consider a massive MIMO topology where N is large and fixed, while K is chosen adaptively.

Each UE is picked at random from an arbitrary user distribution in the serving cell. The geometric position of UE $k \in \{1, \dots, K\}$ in cell $l \in \mathcal{B}$ is thus a random variable and denoted by $\mathbf{z}_{lk} \in \mathbb{R}^2$. The channels are modeled as block-fading; $\mathbf{h}_{jlk} \in \mathbb{C}^N$ denotes the channel response between the considered UE and BS j in a given coherence block of duration T channel uses. This channel varies independently between one coherence block and the next. The realizations are drawn from circularly symmetric complex Gaussian distributions,

$$\mathbf{h}_{jlk} \sim \mathcal{CN}(\mathbf{0}, d_j(\mathbf{z}_{lk})\mathbf{I}_N), \quad (1)$$

which is a reasonable model for arrays with both few and many antennas (see recent measurements reported in [17]). The deterministic function $d_j(\mathbf{z})$ gives the variance of the channel attenuation from BS j at any arbitrary UE position \mathbf{z} . We assume that $d_j(\mathbf{z}_{lk})$ is known at BS j for all l and k , while the exact UE positions are unknown.

The received signal $\mathbf{Y}_j \in \mathbb{C}^{N \times T}$ at BS $j \in \mathcal{B}$ in a coherence block is modeled, similarly to [2], [7], [8], as

$$\mathbf{Y}_j = \sum_{l \in \mathcal{B}} \sum_{k=1}^K \sqrt{p_{lk}} \mathbf{h}_{jlk} \mathbf{x}_{lk}^H + \mathbf{N}_j \quad (2)$$

where $\mathbf{x}_{lk}^H \in \mathbb{C}^{1 \times T}$ is the vector signal from UE k in cell l (normalized such that $\mathbb{E}\{[|\mathbf{x}_{lk}]_t|^2\} = 1$ for all elements $t \in \{1, \dots, T\}$), p_{lk} is the corresponding transmit power, and $\mathbf{N}_j \in \mathbb{C}^{N \times T}$ is additive noise with each element distributed independently as $\mathcal{CN}(0, \sigma^2)$.

Contrary to most previous works on massive MIMO, which assume fixed uplink power, we consider statistics-aware power control²; the signals from UE k in cell l are allocated the power $p_{lk} = \frac{\rho}{d_l(\mathbf{z}_{lk})}$, where $\rho > 0$ is a design parameter. This power-control policy inverts the average channel attenuation and has the merit of making the average effective channel gain the same for all UEs: $\mathbb{E}\{p_{lk} \|\mathbf{h}_{ljk}\|^2\} = N\rho$. This policy guarantees a uniform user experience, saves valuable energy at UEs, and avoids near-far blockage where the receiver's limited dynamic range makes weak signals drown in stronger signals.

¹A few results for single-cell systems are available, for example, in [8].

²Channel-aware power control was considered in [9], [18], but requires a feedback mechanism for instantaneous CSI. Since small-scale fading average out in massive MIMO, we expect statistical power control to be sufficient.

III. AVERAGE PER-CELL SPECTRAL EFFICIENCY

In this section, we derive and analyze the SE, discuss the impact of pilot allocation and user density, and study the behaviors as $N \rightarrow \infty$.

A. Pilot-Based Channel Estimation

BS j can use its multitude of antennas for adaptive receive beamforming, which can amplify useful signals and reject interfering signals. This requires, however, some knowledge of the effective user channels $\sqrt{p_{lk}}\mathbf{h}_{jlk}$ (for all l and k). Such CSI is typically acquired by pilot signaling, where the UEs send known signals. This is a challenging task in networks where the transmission resources are reused across cells, because the pilot signals are then affected by strong inter-cell interference. This so-called *pilot contamination* limits the quality of the acquired CSI and the ability to reject inter-cell interference (unless certain intricate algorithms can be used [14]).

The impact of pilot contamination is usually studied under the assumption that exactly the same pilot signals are used in all cells. In contrast, we now derive the main properties of massive MIMO systems (with power control) for arbitrary pilot allocation. The pilot signals span B channel uses, where $1 \leq B \leq T$, and are sent in the beginning of each coherence block.³ Each pilot signal can be represented by a deterministic vector $\mathbf{v} \in \mathbb{C}^B$ and the fixed per-symbol power implies that all entries have unit magnitude. We assume that all pilot signals originate from a predefined *pilot book*

$$\mathcal{V} = \{\mathbf{v}_1, \dots, \mathbf{v}_B\} \quad \text{where} \quad \mathbf{v}_{b_1}^H \mathbf{v}_{b_2} = \begin{cases} B, & b_1 = b_2, \\ 0, & b_1 \neq b_2. \end{cases} \quad (3)$$

Hence, the B pilot signals form an orthogonal basis and can, for example, be taken as the columns of a discrete Fourier transform (DFT) matrix [19]. The signal transmitted by UE k in cell l is partitioned as $\mathbf{x}_{lk}^H = [\mathbf{v}_{i_{lk}}^H \tilde{\mathbf{x}}_{lk}^H]$, where $i_{lk} \in \{1, \dots, B\}$ is the index of the pilot signal and $\tilde{\mathbf{x}}_{lk}^H \in \mathbb{C}^{1 \times (T-B)}$ is the information signal.

Using this notation, the received signal at BS j can be partitioned as $\mathbf{Y}_j = [\tilde{\mathbf{Y}}_j \hat{\mathbf{Y}}_j]$ where $\tilde{\mathbf{Y}}_j \in \mathbb{C}^{N \times B}$ is the received signal during pilot signaling and $\hat{\mathbf{Y}}_j \in \mathbb{C}^{N \times (T-B)}$ is received during information transmission. According to (2), these matrices are given by

$$\tilde{\mathbf{Y}}_j = \sum_{l \in \mathcal{B}} \sum_{k=1}^K \sqrt{p_{lk}} \mathbf{h}_{jlk} \mathbf{v}_{i_{lk}}^H + \tilde{\mathbf{N}}_j \quad (4)$$

$$\hat{\mathbf{Y}}_j = \sum_{l \in \mathcal{B}} \sum_{k=1}^K \sqrt{p_{lk}} \mathbf{h}_{jlk} \tilde{\mathbf{x}}_{lk}^H + \hat{\mathbf{N}}_j \quad (5)$$

where also the noise matrix is partitioned as $\mathbf{N}_j = [\tilde{\mathbf{N}}_j \hat{\mathbf{N}}_j]$.

Based on this system model, the next lemma derives the linear minimum mean-squared error (LMMSE) estimator of the effective power-controlled channels, which are defined as $\mathbf{h}_{jlk}^{\text{eff}} = \sqrt{p_{lk}}\mathbf{h}_{jlk}$.

Lemma 1. The LMMSE estimate at BS j of the effective power-controlled channel $\mathbf{h}_{jlk}^{\text{eff}}$, for any $l \in \mathcal{B}$ and $k \in \{1, \dots, K\}$, is

$$\hat{\mathbf{h}}_{jlk}^{\text{eff}} = \frac{d_j(\mathbf{z}_{lk})}{d_l(\mathbf{z}_{lk})} (\mathbf{v}_{i_{lk}}^H \Psi_j^{-1} \otimes \mathbf{I}_N) \text{vec}(\tilde{\mathbf{Y}}_j) \quad (6)$$

where \otimes denotes the Kronecker product, $\text{vec}(\cdot)$ is vectorization, and

$$\Psi_j = \sum_{l \in \mathcal{B}} \sum_{m=1}^K \frac{d_j(\mathbf{z}_{lm})}{d_l(\mathbf{z}_{lm})} \mathbf{v}_{i_{lm}} \mathbf{v}_{i_{lm}}^H + \frac{\sigma^2}{\rho} \mathbf{I}_B. \quad (7)$$

³The pilot signals need not be synchronized across cells as assumed herein, but there is little to gain from shifting pilot and information signals between cells; this leads to a mix of deterministic pilots and stochastic information at each channel use, but the average pilot contamination will not change in any substantial way [8, Remark 5]. The effective interference suppression concepts considered in this paper are also harder to implement in such cases.

The estimation error covariance matrix $\mathbf{C}_{jlk} \in \mathbb{C}^{N \times N}$ is given by

$$\begin{aligned} \mathbf{C}_{jlk} &= \mathbb{E} \left\{ (\mathbf{h}_{jlk}^{\text{eff}} - \hat{\mathbf{h}}_{jlk}^{\text{eff}})(\mathbf{h}_{jlk}^{\text{eff}} - \hat{\mathbf{h}}_{jlk}^{\text{eff}})^H \right\} \\ &= \rho \frac{d_j(\mathbf{z}_{lk})}{d_l(\mathbf{z}_{lk})} \left(1 - \frac{d_j(\mathbf{z}_{lk})}{d_l(\mathbf{z}_{lk})} \mathbf{v}_{i_{lk}}^H \Psi_j^{-1} \mathbf{v}_{i_{lk}} \right) \mathbf{I}_N \end{aligned} \quad (8)$$

and the mean-squared error (MSE) is $\text{MSE}_{jlk} = \text{tr}(\mathbf{C}_{jlk})$.

Proof: The expression for an LMMSE estimator is $\hat{\mathbf{h}}_{jlk}^{\text{eff}} = \mathbb{E}\{\mathbf{h}_{jlk}^{\text{eff}} \text{vec}(\tilde{\mathbf{Y}}_j)^H\} \left(\mathbb{E}\{\text{vec}(\tilde{\mathbf{Y}}_j) \text{vec}(\tilde{\mathbf{Y}}_j)^H\} \right)^{-1} \text{vec}(\tilde{\mathbf{Y}}_j)$ [20, Chapter 12], where the expectations are with respect to channel realizations. The lemma follows from direct algebraic computation. ■

There are two important differences between Lemma 1 and the channel estimators that are conventionally used in the massive MIMO literature: 1) we estimate the effective channels including power control; and 2) the estimator supports arbitrary pilot allocation.

Since the pilot signals in \mathcal{V} form an orthogonal basis, we note that

$$\mathbf{v}_{i_{lk}}^H \Psi_j^{-1} = \frac{1}{\sum_{\ell \in \mathcal{B}} \sum_{m=1}^K \frac{d_j(\mathbf{z}_{\ell m})}{d_\ell(\mathbf{z}_{\ell m})} \mathbf{v}_{i_{\ell m}}^H \mathbf{v}_{i_{\ell m}} + \frac{\sigma^2}{\rho}} \mathbf{v}_{i_{lk}}^H \quad (9)$$

which in conjunction with the error covariance matrix in (8) reveals how the estimation error depends on the allocation of pilot signals to the UEs (i.e., which of the products $\mathbf{v}_{i_{lk}}^H \mathbf{v}_{i_{\ell m}}$ that are non-zero).

Although Lemma 1 allows for estimation of all channel vectors, each BS can only resolve B different vectors since there are only B orthogonal pilot signals. To show this, we define the $N \times B$ matrix

$$\hat{\mathbf{H}}_{\mathcal{V},j} = \left[(\mathbf{v}_1^H \Psi_j^{-1} \otimes \mathbf{I}_N) \text{vec}(\tilde{\mathbf{Y}}_j), \dots, (\mathbf{v}_B^H \Psi_j^{-1} \otimes \mathbf{I}_N) \text{vec}(\tilde{\mathbf{Y}}_j) \right] \quad (10)$$

using all B pilot signals. The channel estimate in (6) for UE k in cell l can be written as $\hat{\mathbf{h}}_{jlk}^{\text{eff}} = \frac{d_j(\mathbf{z}_{lk})}{d_l(\mathbf{z}_{lk})} \hat{\mathbf{H}}_{\mathcal{V},j} \mathbf{e}_{i_{lk}}$, where \mathbf{e}_i denotes the i th column of the $B \times B$ identity matrix \mathbf{I}_B . This is the essence of pilot contamination; BSs cannot tell apart UEs that use the same pilot signal and thus cannot reject the corresponding interference.

B. Achievable Spectral Efficiency by Receive Beamforming

The channel estimates allow each BS to semi-coherently detect the information signals from its UEs. In particular, we assume that BS j applies a receive beamforming vector $\mathbf{g}_{jk} \in \mathbb{C}^N$ to the received signal (i.e., $\mathbf{g}_{jk}^H \tilde{\mathbf{Y}}_j$) to amplify the signal from its k th UE and reject inter-user interference. The interference rejection in massive MIMO systems can be either *passive* or *active*. The canonical example of passive rejection is maximum ratio combining (MRC), defined as

$$\mathbf{g}_{jk}^{\text{MRC}} = \hat{\mathbf{H}}_{\mathcal{V},j} \mathbf{e}_{i_{jk}} = \hat{\mathbf{h}}_{jjk}^{\text{eff}}, \quad (11)$$

which maximizes the gain of the useful signal and relies on that interfering channels are quasi-orthogonal to $\hat{\mathbf{h}}_{jjk}^{\text{eff}}$ when N is large.

In contrast, active rejection is achieved by making the receive beamforming as orthogonal to the interfering channels as possible. We propose a new *pilot-based zero-forcing combining* (P-ZFC), defined as

$$\mathbf{g}_{jk}^{\text{P-ZFC}} = \hat{\mathbf{H}}_{\mathcal{V},j} (\hat{\mathbf{H}}_{\mathcal{V},j}^H \hat{\mathbf{H}}_{\mathcal{V},j})^{-1} \mathbf{e}_{i_{jk}}. \quad (12)$$

In contrast to classical zero-forcing receivers, which only try to orthogonalize the K intra-cell channels, P-ZFC exploits that all the B estimated channel directions in $\hat{\mathbf{H}}_{\mathcal{V},j}$ are known at BS j and orthogonalize all of them to also mitigate parts of the inter-cell interference. The cost is a loss in array gain of B instead of K .

MRT and P-ZFC are the receive beamforming schemes considered in this paper, while other schemes are left for future work. The next theorem provides closed-form expressions for the achievable uplink SEs and effective signal-to-interference-and-noise ratios (SINRs).

$$\overline{\text{SINR}}_{jk}^{\text{MRC}} = \frac{B}{\left(\sum_{l \in \mathcal{B}} \mu_{jl}^{(1)} \frac{K}{N} + \frac{\sigma^2}{N\rho} \right) \left(\sum_{\ell \in \mathcal{B}_{m=1}}^K \mu_{j\ell}^{(1)} \mathbf{v}_{i_{jk}}^H \mathbf{v}_{i_{\ell m}} + \frac{\sigma^2}{\rho} \right) + \sum_{l \in \mathcal{B}_{m=1}}^K \left(\mu_{jl}^{(2)} + \frac{\mu_{jl}^{(2)} - (\mu_{jl}^{(1)})^2}{N} \right) \mathbf{v}_{i_{jk}}^H \mathbf{v}_{i_{lm}} - B} \quad (13)$$

$$\overline{\text{SINR}}_{jk}^{\text{P-ZFC}} = \frac{B}{\sum_{l \in \mathcal{B}_{m=1}}^K \left(\mu_{jl}^{(2)} - B \frac{\mu_{jl}^{(2)} - (\mu_{jl}^{(1)})^2}{N(N-B)} \right) \mathbf{v}_{i_{jk}}^H \mathbf{v}_{i_{lm}} + \left(\sum_{l \in \mathcal{B}_{m=1}}^K \mu_{jl}^{(1)} \left(1 - \frac{B \mu_{jl}^{(1)}}{\sum_{\ell \in \mathcal{B}_{m=1}}^K \mu_{j\ell}^{(1)} \mathbf{v}_{i_{lm}}^H \mathbf{v}_{i_{\ell m}} + \frac{\sigma^2}{\rho}} \right) + \frac{\sigma^2}{\rho} \right) \left(\frac{\sum_{\ell \in \mathcal{B}_{m=1}}^K \mu_{j\ell}^{(1)} \mathbf{v}_{i_{jk}}^H \mathbf{v}_{i_{\ell m}} + \frac{\sigma^2}{\rho}}{N-B} \right) - B} \quad (14)$$

Theorem 1. An ergodic achievable SE in cell j is

$$\text{SE}_j = \sum_{k=1}^K \left(1 - \frac{B}{T} \right) \log_2 \left(1 + \overline{\text{SINR}}_{jk} \right) \quad [\text{bit/s/Hz/cell}] \quad (15)$$

where the SINR for UE k , $\overline{\text{SINR}}_{jk}$, is given in (13) for MRC and (14) for P-ZFC at the top of the page. The following notation is used:

$$\mu_{jl}^{(\gamma)} = \mathbb{E}_{\mathbf{z}_{lm}} \left\{ \left(\frac{d_j(\mathbf{z}_{lm})}{d_l(\mathbf{z}_{lm})} \right)^\gamma \right\} \quad \text{for } \gamma = 1, 2. \quad (16)$$

Proof: The SEs follow from treating interference and channel uncertainty as worst-case Gaussian noise (cf. [2], [6], [7]) and taking variations in interference power (due to random UE positions) into account in the coding. Details are omitted due to limited space. ■

The closed-form SE expressions in Theorem 1 are only functions of the pilot allocation and the propagation parameters $\mu_{jl}^{(1)}, \mu_{jl}^{(2)}$. The latter are the average ratio between the channel variance to BS j and the channel variance to BS l , for an arbitrary UE in cell l , and the average second-order moment of this ratio, respectively. The SE expressions manifest the importance of pilot allocation, since all interference terms contain inner products of pilot signals that are either zero or B . Since it is hard to gain insights directly from the SINR expressions in (13) and (14), we look at the asymptotic limit.

Corollary 1. When $N \rightarrow \infty$ (with $K, B \leq T < \infty$), the SEs with MRT and P-ZFC approach the same limit:

$$\overline{\text{SINR}}_{jk}^{\text{MRC}}, \overline{\text{SINR}}_{jk}^{\text{P-ZFC}} \rightarrow \frac{B}{\sum_{l \in \mathcal{B}_{m=1}}^K \mu_{jl}^{(2)} \mathbf{v}_{i_{jk}}^H \mathbf{v}_{i_{lm}} - B}. \quad (17)$$

In order to maximize the asymptotic SINR in (17), we should allocate the pilot signals such that $\mathbf{v}_{i_{jk}}^H \mathbf{v}_{i_{lm}} = 0$ whenever $\mu_{jl}^{(2)}$ is large. If $\beta = \frac{B}{K}$ is an integer, this amounts to allocating orthogonal pilots among the UEs in each cell and making sure that only $\frac{1}{\beta}$ of the interfering cells reuse these pilots. We refer to this as *fractional pilot reuse*. An explicit example is provided in the next section for hexagonal cells. The following result holds for any cell structure.

Corollary 2. Let $B = K\beta$ for some integer β , assume orthogonal intra-cell pilot signals, and let $\mathcal{B}_j \subset \mathcal{B}$ be the set of cells that use the same pilots as cell j . The asymptotic SE for cell j becomes

$$\text{SE}_j^\infty = K \left(1 - \frac{K\beta}{T} \right) \log_2 \left(1 + \frac{1}{\sum_{l \in \mathcal{B}_j \setminus \{j\}} \mu_{jl}^{(2)}} \right), \quad (18)$$

when $N \rightarrow \infty$. This SE is maximized jointly for all cells by $K^* = \left\lfloor \frac{T}{2\beta} \right\rfloor$ or $K^* = \left\lceil \frac{T}{2\beta} \right\rceil$ scheduled users (one of the closest integers).

Proof: The logarithmic part of (18) is independent of K , while the concave pre-log factor $K \left(1 - \frac{K\beta}{T} \right)$ is maximized by $K = \frac{T}{2\beta}$. ■

Corollary 2 is a main contribution of this paper and proves that the number of scheduled UEs should be proportional to the coherence block length T (when N is large enough); for example, we get $K^* = \frac{T}{2}$ for $\beta = 1$ and $K^* = \frac{T}{6}$ for $\beta = 3$. Since both $T = 500$ and $T = 10000$ are reasonable coherence block lengths in practice, depending

on the UE mobility, this means that we should schedule hundreds or even thousands of simultaneous UEs in order to be optimal. This is only possible if the UE selection policy is simple and scalable. If $K^* = \frac{T}{2\beta}$ is an integer, the asymptotically optimal SE is

$$\text{SE}_j^\infty = \frac{T}{4\beta} \log_2 \left(1 + \frac{1}{\sum_{l \in \mathcal{B}_j \setminus \{j\}} \mu_{jl}^{(2)}} \right) \quad (19)$$

and increases linearly with T (in the large- N regime).

Interestingly, the optimal scheduling gives $B = \frac{T}{2}$ for any β , which means that *half* the coherence block is used for pilot transmission. This is an extraordinary fact that bears some similarity with results in [21], [22] for block-fading noncoherent point-to-point MIMO channels, where the maximal degrees of freedom (DoF) are $\frac{T}{4}$ and are achieved by having $\frac{T}{2}$ transmit/receive antennas and using pilots of the same length. The fundamental difference is that the DoF concept, where unbounded SE is achieved at high SNRs, does not apply to cellular networks [23]. Instead, the pre-log factor $\frac{T}{4\beta}$ in (19) may be interpreted as the relative improvement in SE that can be achieved by aggressive UE scheduling in massive MIMO systems.

We have now established the asymptotically optimal number of scheduled UEs, as $N \rightarrow \infty$. To investigate the impact on practical systems with finite N , the next section considers a certain network.

IV. OPTIMIZING NUMBER OF USERS IN HEXAGONAL NETWORKS

The concept of cellular communications has been around for decades [24]. Although practical deployments have irregular cells, it is common practice to establish general properties by analyzing symmetric networks where the cells are regular hexagons [25].

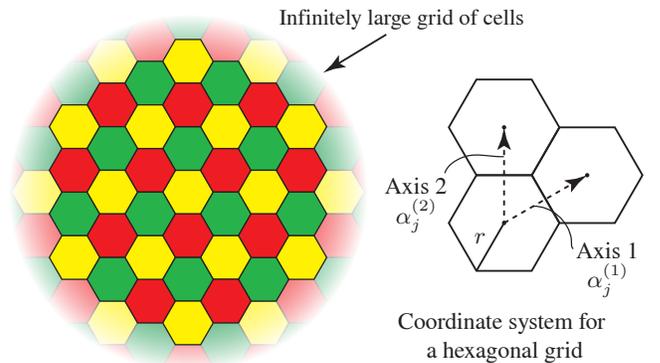
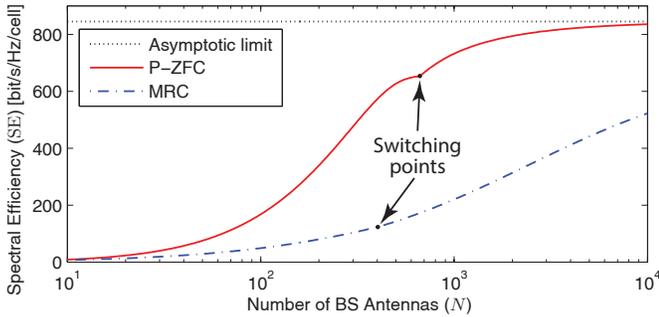


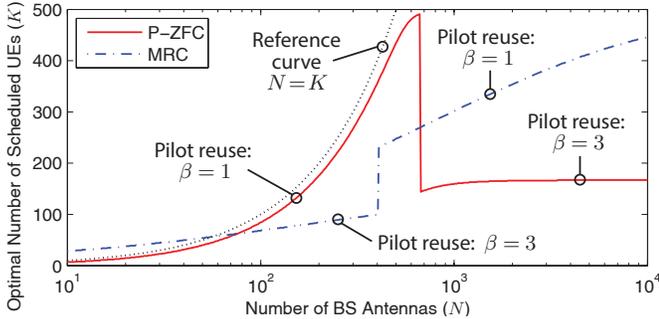
Fig. 1: Part of a hexagonal network, colored for pilot reuse $\beta = 3$.

In this section, we consider the symmetric network depicted in Fig. 1 with hexagonal cells of radius $r > 0$ (the distance from center to corners). The hexagonal grid is infinitely large, to avoid edge effect and giving all cells the same properties. Each cell can be uniquely indexed by a pair of integers $\alpha_j^{(1)}, \alpha_j^{(2)} \in \mathbb{Z}$, where \mathbb{Z} is the set of all integers. This integer pair specifies the location of BS j [24]:

$$\mathbf{b}_j = \begin{bmatrix} \sqrt{3}r/2 \\ r/2 \end{bmatrix} \alpha_j^{(1)} + \begin{bmatrix} 0 \\ \sqrt{3}r \end{bmatrix} \alpha_j^{(2)} \in \mathbb{R}^2. \quad (20)$$



(a) Optimized SE per cell.



(b) Optimal number of scheduled UEs, K^* , per cell.

Fig. 2: Simulation results with average inter-cell interference.

We assume a classic pathloss model where the variance of the channel attenuation in (1) is $d_j(\mathbf{z}) = \frac{C}{\|\mathbf{z} - \mathbf{b}_j\|^\kappa}$, where $\|\cdot\|$ is the Euclidean norm, $C > 0$ is a reference value, and $\kappa \geq 2$ is the pathloss exponent. These assumptions allow us to compute $\mu_{jl}^{(\gamma)}$ in (16) as

$$\mu_{jl}^{(\gamma)} = \mathbb{E}_{\mathbf{z}_{lm}} \left\{ \left(\frac{d_j(\mathbf{z}_{lm})}{d_l(\mathbf{z}_{lm})} \right)^\gamma \right\} = \mathbb{E}_{\mathbf{z}_{lm}} \left\{ \left(\frac{\|\mathbf{z}_{lm} - \mathbf{b}_l\|}{\|\mathbf{z}_{lm} - \mathbf{b}_j\|} \right)^{\kappa\gamma} \right\} \quad (21)$$

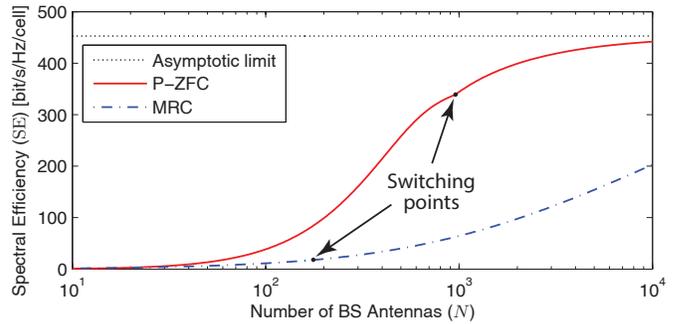
for any UE distributions in the cells. We notice that C and r cancel out in (21), if the UE distribution is invariant to the radius. Since the power control makes the SINRs in Theorem 1 independent of the UE's position, we only need to define the parameter ρ ; the average SNR (over fading) between any UE and any antenna at its serving BS. This parameter is set to 10 dB in the simulations, which is not particularly large but allows for decent channel estimation accuracy and convergence speed to the large- N regime; see [10].

Although all cells reuse all time/frequency resources, we consider pilot books of size $B = \beta K$ to allow for fractional pilot reuse that avoids pilot contamination from neighboring cells. The hexagonal grid only allows for some reuse factors: $\beta \in \{1, 3, 4, 7, \dots\}$ [24].

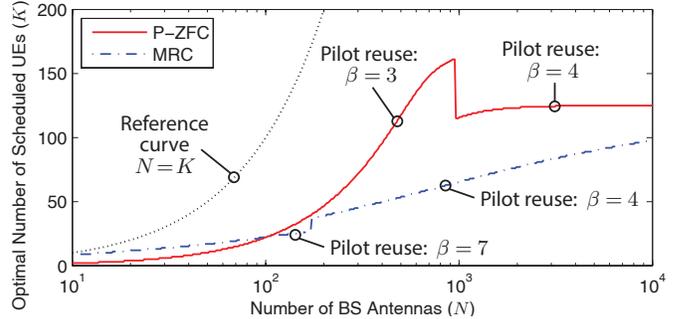
A. Simulation Results

We simulate the uplink SE in an arbitrary cell on the hexagonal grid in Fig. 1 and take all non-negligible interference into account. For each number of antennas, N , we optimize the SE with respect to the number of UEs K and pilot reuse factor β (which determine $B = \beta K$). We set the coherence block length to $T = 1000$ (e.g., achieved by 10 ms coherence time and 100 kHz coherence bandwidth), use the whole block for uplink, and pick $\kappa = 3.5$ as pathloss exponent.

We compute $\mu_{jl}^{(1)}$ and $\mu_{jl}^{(2)}$ in two different ways: 1) by averaging over uniform UE locations in all cells (restricting all UEs to be at least $0.14r$ from the BS); and 2) worst-case scenario where all UEs in other cells are at the cell edge closest to BS j (for each j). The former is optimistic since the variations in interference power from the average number can be large, while the latter is overly pessimistic



(a) Optimized SE per cell.



(b) Optimal number of scheduled UEs, K^* , per cell.

Fig. 3: Simulation results with worst-case inter-cell interference.

since the UEs cannot all be at the worst locations simultaneously. The average case is shown in Fig. 2 and the worst case in Fig. 3.

Interestingly, the results for MRC and P-ZFC are very different. MRC uses fractional pilot reuse at small N to reduce the inter-cell interference, while it reduces β at large N since the UEs' channels are asymptotically quasi-orthogonal. In contrast, P-ZFC increases the pilot reuse β with N since its active interference cancellation becomes more effective in this range. The performance gap between P-ZFC and MRC is large, particularly when hundreds or thousands of antennas are deployed, thus P-ZFC is preferable if we can afford the higher complexity of computing the beamforming. A few thousand antennas are needed to converge to the asymptotic limit in (19) using P-ZFC, while many more antennas are needed with MRC.

The number of scheduled UEs is very different for the two beamforming schemes and between the mean and worst-case interference. In Fig. 2, where the inter-cell interference is relatively mild, it can be preferable to schedule hundreds of UEs; P-ZFC schedules almost as many UEs as there are antennas, until we reach $K = \frac{T}{2}$ and the SE can only be increased by fractional pilot-reuse. For the worst-case interference in Fig. 3, it is preferable to schedule fewer UEs, but there can still be tens or a hundred UEs per cell. In the practical range of $100 \leq N \leq 1000$, we notice that the active interference cancellation in P-ZFC allows it to schedule many more UEs than MRC does.

V. CONCLUSION

This paper investigated how many UEs, K , that should be scheduled in massive MIMO (with power control) to maximize the SE per cell for a fixed N . The optimal K approaches $\frac{T}{2\beta}$ as $N \rightarrow \infty$, but can be different at finite N . The common rule-of-thumb of $N \gg K$ holds true when $N \gg \frac{T}{2}$, but not when the inter-cell interference is weak due to randomized UE locations. By simulations, a $100\times$ increase in SE over the IMT-Advanced requirement of 2.25 bit/s/Hz/cell was demonstrated. Taking the ratio between SE and K reveals that this is achieved by having an SE per UE in the common range (1-5 bit/s/Hz/user), but schedule unconventionally many UEs per cell.

REFERENCES

- [1] Nokia Siemens Networks, "2020: Beyond 4G radio evolution for the Gigabit experience," White Paper, Tech. Rep., 2011.
- [2] T. L. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.
- [3] R. Baldemair, E. Dahlman, G. Fodor, G. Mildh, S. Parkvall, Y. Selen, H. Tullberg, and K. Balachandran, "Evolving wireless communications: Addressing the challenges and expectations of the future," *IEEE Veh. Technol. Mag.*, vol. 8, no. 1, pp. 24–30, Mar. 2013.
- [4] F. Boccardi, R. Heath, A. Lozano, T. Marzetta, and P. Popovski, "Five disruptive technology directions for 5G," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 74–80, Feb. 2014.
- [5] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, vol. 52, no. 2, pp. 186–195, Feb. 2014.
- [6] J. Jose, A. Ashikhmin, T. L. Marzetta, and S. Vishwanath, "Pilot contamination and precoding in multi-cell TDD systems," *IEEE Trans. Commun.*, vol. 10, no. 8, pp. 2640–2651, Aug. 2011.
- [7] J. Hoydis, S. ten Brink, and M. Debbah, "Massive MIMO in the UL/DL of cellular networks: How many antennas do we need?" *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 160–171, Feb. 2013.
- [8] H. Q. Ngo, E. G. Larsson, and T. L. Marzetta, "Energy and spectral efficiency of very large multiuser MIMO systems," *IEEE Trans. Commun.*, vol. 61, no. 4, pp. 1436–1449, Apr. 2013.
- [9] E. Björnson, L. Sanguinetti, J. Hoydis, and M. Debbah, "Optimal design of energy-efficient multi-user MIMO systems: Is massive MIMO the answer?" *IEEE Trans. Wireless Commun.*, submitted, Available: <http://arxiv.org/abs/1403.6150>.
- [10] E. Björnson, J. Hoydis, M. Kountouris, and M. Debbah, "Massive MIMO systems with non-ideal hardware: Energy efficiency, estimation, and capacity limits," *IEEE Trans. Inf. Theory*, to appear.
- [11] E. Björnson, M. Matthaiou, and M. Debbah, "Massive MIMO with arbitrary non-ideal arrays: Hardware scaling laws and circuit-aware design," *IEEE Trans. Wireless Commun.*, submitted, Available: <http://arxiv.org/abs/1409.0875>.
- [12] H. Huh, G. Caire, H. Papadopoulos, and S. Ramprasad, "Achieving "massive MIMO" spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, Sept. 2012.
- [13] M. Li, Y.-H. Nam, B. Ng, and J. Zhang, "A non-asymptotic throughput for massive MIMO cellular uplink with pilot reuse," in *Proc. IEEE Globecom*, 2012.
- [14] R. Müller, M. Vehkaperä, and L. Cottarelli, "Blind pilot decontamination," in *Proc. ITG Workshop on Smart Antennas (WSA)*, 2013.
- [15] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE J. Sel. Areas Commun.*, vol. 31, no. 2, pp. 264–273, Feb. 2013.
- [16] M. Li, S. Jin, and X. Gao, "Spatial orthogonality-based pilot reuse for multi-cell massive MIMO transmission," in *Proc. WCSP*, 2013.
- [17] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO in real propagation environments," *IEEE Trans. Wireless Commun.*, 2014, submitted. [Online]. Available: <http://arxiv.org/abs/1403.3376>
- [18] K. Guo, Y. Guo, G. Fodor, and G. Ascheid, "Uplink power control with MMSE receiver in multi-cell MU-massive-MIMO systems," in *Proc. IEEE ICC*, 2014.
- [19] M. Biguesh and A. B. Gershman, "Downlink channel estimation in cellular systems with antenna arrays at base stations using channel probing with feedback," *EURASIP J. Appl. Signal Process.*, vol. 2004, no. 9, pp. 1330–1339, 2004.
- [20] S. M. Kay, *Fundamentals of Statistical Signal Processing: Estimation Theory*. Prentice Hall, 1993.
- [21] L. Zheng and D. Tse, "Communication on the Grassmann manifold: A geometric approach to the noncoherent multiple-antenna channel," *IEEE Trans. Inf. Theory*, vol. 48, no. 2, pp. 359–383, Feb. 2002.
- [22] B. Hassibi and B. M. Hochwald, "How much training is needed in multiple-antenna wireless links?" *IEEE Trans. Inf. Theory*, vol. 49, no. 4, pp. 951–963, Apr. 2003.
- [23] A. Lozano, R. Heath, and J. Andrews, "Fundamental limits of cooperation," *IEEE Trans. Inf. Theory*, vol. 59, no. 9, pp. 5213–5226, Sept. 2013.
- [24] V. M. Donald, "The cellular concept," *Bell System Technical Journal*, vol. 58, no. 15–41, pp. 113–381, 1979.
- [25] D. Cox, "Cochannel interference considerations in frequency reuse small-coverage-area radio systems," *IEEE Trans. Commun.*, vol. 30, no. 1, pp. 135–142, Jan. 1982.