Cooperative Multi-Cell Zero-Forcing Beamforming in Cellular Downlink Channels

Oren Somekh*, Osvaldo Simeone†, Yeheskel Bar-Ness†,

Alexander M. Haimovich †, and Shlomo Shamai (Shitz)‡

* Department of Electrical Engineering, Princeton University, Princeton, NJ 08544, USA, Email: orens@princeton.edu

† Department of Electrical and Computer Engineering, NJIT, Newark, NJ 07102, USA, Email: {simeone, haimovich, barness}@njit.edu

‡ Department of Electrical Engineering, Technion, Haifa 32000, Israel, Email: sshlomo@ee.technion.ac.il

Abstract

For a multiple-input single-output (MISO) downlink channel with M transmit antennas, it has been recently proved that zero-forcing beamforming (ZFBF) to a subset of (at most) M "semi-orthogonal" users is optimal in terms of sum-rate, asymptotically with the number of users. However, determining the subset of users for transmission is a complex optimization problem. Adopting the ZFBF scheme in a cooperative multi-cell scenario renders the selection process even more difficult since more users are likely to be involved. In this paper, we consider a multi-cell cooperative ZFBF scheme combined with a simple sub-optimal users selection procedure for the Wyner downlink channel setup. According to this sub-optimal procedure, the user with the "best" local channel is selected for transmission in each cell. The performance of this sub-optimal scheme is investigated in terms of both, the conventional scaling law of the sum-rate with the number of users, and a sum-rate offset. We term this characterization of the sum-rate for large number of users as *high-load regime* characterization, and point out the similarity of this approach with the standard affine approximation used in the high-SNR regime. It is shown that under an overall power constraint, the sub-optimal cooperative multi-cell ZFBF scheme achieves the same sum-rate growth rate and slightly degraded offset law, when compared to an optimal scheme deploying joint multi-cell dirty-paper coding (DPC) techniques, asymptotically with the number of users

per cell. Moreover, the overall power constraint is shown to ensure in probability equal per-cell power constraints when the number of users per-cell increases.

I. INTRODUCTION

The growing demand for ubiquitous access to high-data rate services, has produced a huge amount of research analyzing the performance of wireless communications systems. Cellular systems are of major interest as the most common method for providing continuous services to mobile users, in both indoor and outdoor environments. In particular, the use of joint multi-cell processing has been identified as a key tool for enhancing system performance (see [1][2] and references therein for surveys of recent results on multi-cell processing).

Most of the works on the downlink channel of cellular systems deal with a single-cell setup. References that consider multi-cell scenarios (e.g. [3]-[5]) tend to adopt complex multi-cell system models which render analytical treatment extremely hard (if not, impossible). Indeed, most of the results reported in these works are derived via intensive numerical calculations which provide little insight into the behavior of the system performance as a function of various key parameters. The main goal of this paper is to present and analyze efficient, sub-optimal scheduling schemes for the downlink channel of multi-cell systems. An emphasis is put on deriving analytical results which provide insight into the role of key parameters on system performance. To achieve this goal a simple cellular model based on a model presented by Wyner in [6] is considered. According to this model (depicted in Fig. 1 with four cells) the cells are placed on a circle and each users "sees" only three cell-site antennas. In addition, the path loss is modelled by a single parameter $\alpha \in [0,1]$. Although this model is hardly realistic it encompasses the essence of real-life system parameters such as fading and inter-cell interference.

The downlink channel of a similar model was first adopted in [7] were LQ factorization (forcing an arbitrary sub-optimal encoding order) combined with joint multi-cell dirty-paper coding (DPC) is deployed. The attainable per-cell sum-rates under an *overall power constraint* and in the presence of Rayleigh flat fading, are shown via numerical calculations to approach those of the optimal DPC scheme (with optimal encoding order) in the high-SNR region. Recently, bounds on the per-cell sum-rate capacity supported by the downlink of this model have been reported in [8] under *equal per-cell power constraints* in the presence of Rayleigh flat fading. To achieve theses rates, DPC techniques are deployed [9]. Unfortunately, DPC is

difficult to implement in practical systems due to the high computational burden of the successive encoding involved, in particular when the number of users is large. It is evident that for multi-cell processing, where more users are typically involved, this problem aggravates. Therefore, a search for sub-optimal broadcast schemes is the focal point of many works. In particular, linear precoding schemes which offer a tradeoff between complexity and performance have been intensively investigated in recent years [10]-[12]. A simple linear precoding scheme which projects the multi-user channel into multiple independent single-user channels, and reduces the design into scheduling and power allocation problem, is the zero-forcing beamforming (ZFBF) scheme [13]-[15]. The ZFBF scheme is asymptotically optimal with increasing SNR, and it is easily generalized to incorporate DPC techniques [13][16]. Recently, ZFBF scheme has been considered in [17] for a single cell M antennas MISO downlink setup under sum power constraint. In this sub-optimal scheme, a set of (at most) M "semi-orthogonal" users to be served is selected, so as to maximize the sum-rate, and independent coding is employed for each selected user. This strategy is proved to provide optimal rates (as DPC) asymptotically with the number of users K. However, determining the subset of users scheduled for transmission is a complex optimization problem especially when K is large.

Most works dealing with asymptotic analysis of channels with increasing number of users (referred to as the *high-load regime*), are focused on the scaling law. In recent work [18], the authors also considered the rate offset in addition to the rate scaling law of a MIMO broadcast channel with random beamforming. In this work, we formalize this approach by defining the high-load regime characterization. In particular, following the methodology of [19] regarding the high-SNR regime, we define the high-load regime slope and offset. The definitions are then used for asymptotic analysis of the various schemes of interest.

In this paper, we consider cooperative multi-cell ZFBF for the downlink of a Wyner circular setup, with simple scheduling. According to this scheme, in each cell the user with the "best" local channel (the channel from the local cell-site) is scheduled for transmission by means of cooperative multi-cell beamforming. The main results reported in this work include a closed form expression for the per-cell sum-rate of the proposed scheme in the absence of fading. It is proved that this rate is achieved under both overall, and equal per-cell power constraints. In addition, it is shown that ZFBF scheme is superior to a simple inter-cell time sharing (ICTS) scheme when the SNR is above a certain threshold, which increases with the inter-cell interference α. Introducing

Rayleigh fading, the per-cell sum-rate of ZFBF is proved to experience the same growth rate and slightly degraded offset law, when compared to an optimal DPC scheme asymptotically with the number of users per-cell K under sum-power constraint. Moreover, it is verified that the scheme satisfies in probability the more suitable *equal per-cell power constraints* asymptotically with increasing K. Numerical results show that the asymptotic expression derived for this setup already hold for a modest number of users per-cell. Next, we consider a different procedure according to which the user with the "best" total receive power from the three base-stations in sight, is selected for transmission. It is shown that this user selection scheme does not provide higher rates in the high-load regime.

The system model and the ZFBF scheme are described in Section II. Additional background and previous results are elaborated in Section III. The high-load regime characterization is defined in Section IV. Sum-rate analysis for the non-fading and Rayleigh fading setups is presented in Section V. In Section VI an alternative user selection procedure is discussed. Numerical results and concluding remarks are presented in Sections VII and VIII respectively. Various proofs and derivations are included in the Appendix.

II. System Model

Consider a circular variant of the infinite linear Wyner model [6] depicted in Fig. 1, in which $M > 2$ cells with K users each, are arranged on a circle. Assuming a synchronous intra-cell TDMA scheme, according to which only one user is selected for transmission per-cell, the $M \times 1$ vector baseband representation of the signals received by the *selected* users is given for an arbitrary time index by

$$\boldsymbol{y} = \boldsymbol{H}\boldsymbol{B}\boldsymbol{u} + \boldsymbol{z} \ , \tag{1}$$

where $\boldsymbol{u}$ is the $M \times 1$ complex Gaussian symbols vector $\boldsymbol{u} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_M)$, $\boldsymbol{B}$ is the beamforming $M \times M$ matrix, $\boldsymbol{z}$ is the $M \times 1$ complex Gaussian additive noise vector $\boldsymbol{z} \sim \mathcal{CN}(\boldsymbol{0}, \boldsymbol{I}_M)$, and

$\boldsymbol{H}$ is the $M \times M$ channel transfer matrix, given by

$$\boldsymbol{H} = \begin{pmatrix} a_0 & \alpha c_0 & 0 & \cdots & 0 & \alpha b_0 \\ \alpha b_1 & a_1 & \alpha c_1 & 0 & \cdots & 0 \\ 0 & \alpha b_2 & a_2 & \alpha c_2 & \ddots & \vdots \\ \vdots & 0 & \alpha b_3 & \ddots & \ddots & 0 \\ 0 & \vdots & \ddots & \ddots & a_{M-2} & \alpha c_{M-2} \\ \alpha c_{M-1} & 0 & \cdots & 0 & \alpha b_{M-1} & a_{M-1} \end{pmatrix} , \tag{2}$$

where $\alpha \in [0,1]$ is the inter-cell interference factor, representing the geometrical path losses. In addition, a_m, b_m and c_m are the flat fading coefficients of the signals transmitted by the m'th, $(\widehat{m-1})$'th[1] and $(\widehat{m+1})$'th cell-sites respectively, and received by the *selected* user of the m'th cell. It is noted that the fading coefficient might be statistically dependent, depending on the users selection procedure. In addition, ergodic block fading processes are assumed where the fade values remain constant during the TDMA slot duration. Each of the MK users, perfectly measures its *own* fade coefficients $\{a_{m,k}, b_{m,k}, c_{m,k}\}$, which are fed back to the multi-cell transmitter via an ideal delayless feedback channel. Moreover, no user cooperation is allowed.

A joint multi-cell ZFBF scheme is utilized, whose beamforming matrix for an arbitrary TDMA slot is given by

$$\boldsymbol{B} = \sqrt{\frac{MP}{\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right)}}\ \boldsymbol{H}^{-1} , \tag{3}$$

where MP is the overall average transmit power constraint, which is ensured by definition[2]. Substituting (3) into (1), the received signal vector reduces to

$$\boldsymbol{y} = \sqrt{\frac{MP}{\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right)}}\ \boldsymbol{u} + \boldsymbol{z} , \tag{4}$$

and single user encoding-decoding schemes with long code words lasting over many symbols (and many fading blocks) are used. Since (4) can be interpreted as a set of M identical independent parallel single user channels, its achievable per-channel ergodic sum-rate (or cell)

[1] $\widehat{n} \triangleq [n \bmod M]$.

[2] Later on it is argued that under certain conditions this scheme satisfies an equal per-cell average power constraints as well.

is given by[3]

$$R_{\text{zfbf}} = E\left\{\log\left(1+\frac{MP}{\text{tr}\left((\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right)}\right)\right\} \ , \tag{5}$$

where the expectation is taken over the entries of $\boldsymbol{H}$. In the sequel the following Proposition is helpful.

Proposition 1 *The achievable per-cell ergodic sum-rate of the ZFBF scheme with an arbitrary user selection procedure and total sum power constraint MP, is upper bounded by*

$$R_{\text{zfbf}} \leq \log\left(1+\frac{P}{M}E\{\text{tr}\left(\boldsymbol{H}\boldsymbol{H}^{\dagger}\right)\}\right) \ . \tag{6}$$

Proof: See Appendix A. ■

Although a sum-power constraint is assumed, a more natural choice for a cellular system is to maintain per-cell power constraints. Hence, we are interested in the transmitted power of an arbitrary cell, which is averaged over the TDMA time slot duration (many symbols) and is a function of the realization of $\boldsymbol{H}$,

$$P_m = [\boldsymbol{B}\boldsymbol{B}^{\dagger}]_{m,m} = \frac{MP\left[(\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right]_{m,m}}{\text{tr}\left((\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right)} \quad , \quad m=0,\ 1,\ldots,\ M-1 \ . \tag{7}$$

The following Proposition will be useful latter on.

Proposition 2 *The average transmitted power per-cell P_m of the ZFBF scheme with an arbitrary user selection procedure (expression* (7)*), satisfies the following inequality*

$$P\ \frac{\min_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^{\dagger})}{\max_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^{\dagger})} \leq P_m \leq P\ \frac{\max_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^{\dagger})}{\min_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^{\dagger})} \quad , \quad \forall m \ , \tag{8}$$

where $\{\lambda_n(\boldsymbol{H}\boldsymbol{H}^{\dagger})\}_{n=0}^{M-1}$ are the eigenvalues of $\boldsymbol{H}\boldsymbol{H}^{\dagger}$.

Proof: See Appendix B. ■

The above discussion holds for any ZFBF scheme with: sum-power constraint, no further power allocation via "waterfilling", and an arbitrary selection of M users (one in each cell). In Section V, simple scheduling for both non-fading and Rayleigh fading setups are presented and analyzed. The next section provides some additional background and related results derived for the Wyner downlink channel. Theses results are used as a reference for evaluating the performance of the proposed ZFBF scheme.

[3]A natural logarithmic base is used throughout this work.

III. BACKGROUND

For a similar model but with capacity achieving joint multi-cell DPC scheme, the downlink per-cell ergodic sum-rate capacity with an equal per-cell power constraint P, was recently considered in [8] (see also [20]) for the the *soft-handoff model* in which each user "sees" only two cell sites. These results can be extended to include the current circular Wyner model in a straightforward manner. Accordingly, the per-cell ergodic sum-rate capacity with an equal per-cell power constraint, for a non fading setup, is given by

$$C_{\mathrm{opt-nf}} \underset{M\to\infty}{=} \int_0^1 \log\left(1+P(1+2\alpha\cos(2\pi\theta))^2\right) d\theta \ , \tag{9}$$

and for a Rayleigh fading setup with many users per-cell ($K \gg 1$), to be bounded by

$$\log\left(1+P\left((1-\epsilon)\log K+1+2\alpha^2\right)\right) \quad \leq \quad C_{\mathrm{opt}} \quad \leq \quad \log\left(1+(1+2\alpha^2)P\log K\right) \quad , \tag{10}$$

for some $\epsilon \underset{K\to\infty}{\to} 0$.

As a reference, and assuming the system includes an even number of cells, an inter-cell time sharing (ICTS) scheduling, according to which odd and even cells are transmitting alternately in time, is used. This simple scheme (presented in [21] for the uplink channel) requires only limited cooperation between cells, and deploys single-user encoding decoding schemes. Since for each time slot only odd or even indexed cells are transmitting, and the model assumes interference from the two adjacent cells only, inter-cell interference is avoided and the scheme demonstrates a non interference limited behavior. It is easily verified that the achievable per-cell ergodic sum-rate for a non-fading setup is given by

$$R_{\mathrm{icts-nf}} = \frac{1}{2}\log(1+2P) \ . \tag{11}$$

and for a Raleigh fading setup is well approximated (for a large number of users per cell $K \gg 1$) by

$$R_{\mathrm{icts}} \cong \frac{1}{2}\log(1+2P\log K) \ . \tag{12}$$

The latter rate is achieved by scheduling in each active cell, the user with the "best" channel for transmission.

IV. High-Load Regime Characterization Definition

Most of the works dealing with asymptotic analysis of downlink channels with large number of users, are focused on the sum-rate scaling law. This traditional characterization is unable to assess the impact of other channel features since many considered channels demonstrate the same scaling law. Moreover, a characterization based only on the scaling law, dose not reveal much of the actual number of users and power required to achieve a finite rate. Here, we present a more refine characterization whereby as function of $\log\log K$ the rate in a scenario where large number of users are involved (referred to as the *high-load regime*), is expanded as an affine function. The resulting zero-order term or *high-load offset* captures the impact of other channel features such as scheduling and coding.

The high-load refinement is inspired by the seminal works of Shamai et-al [22] and Lozano et-al [19], dealing (among other things) with the power offset in the high-SNR regime. In fact, the definitions of the high-load slope and offset are similar to the respective high-SNR slope and offset where the term $\log P$ in the latter is played by the term $\log\log K$.

In the the high-load regime the sum-rate per receive element of "good" scheduling schemes in the presence of Rayleigh fading behaves as

$$R = \mathsf{S}_\infty \left(\log\log K + \mathsf{L}_\infty\right) + o(1) \ , \tag{13}$$

while S_∞ denotes the high-load slope in bits/sec/Hz,

$$\mathsf{S}_\infty \triangleq \lim_{K\to\infty} \frac{R}{\log\log K} \ , \tag{14}$$

and L_∞ defined by

$$\mathsf{L}_\infty \triangleq \lim_{K\to\infty} \left(\frac{R}{\mathsf{S}_\infty} - \log\log K\right) \ , \tag{15}$$

represent the high-load offset, with respect to some reference channel having the same high-load slope. It is noted that since we are interested in the asymptotic behavior of the rate in terms of the number of users a natural logarithm is used instead of the 10 base logarithm used for the high-SNR regime resulting the use of $3[dB]$ units for the power offset.

Applying these definitions, we get that the high-load characterization of the optimal joint multi-cell DPC scheme in the presence of Rayleigh fading (expression (10)) is given by

$$\mathsf{S}_\infty^{\mathrm{opt}} = 1 \quad ; \quad \log P \leq \mathsf{L}_\infty^{\mathrm{opt}} \leq \log P + \log(1+2\alpha^2) \ , \tag{16}$$

While the high-load characterization of the ICTS scheme in the presence of Rayleigh fading (expression (12)) is given by

$$\mathsf{S}_{\infty}^{\text{icts}} = \frac{1}{2} \quad ; \quad \mathsf{L}_{\infty}^{\text{icts}} = \log P \ . \tag{17}$$

To conclude this section we note that the baseline scaling law of $\log\log K$ is matched to Rayleigh fading channels and is motivated from the fact that the maximum of K i.i.d. χ^2_{2n} distributed r.v.'s roughly behaves like $\log K$ [23] (Example 1, Appendix A). Other fading distribution might lead to other baseline scaling laws (see for example [24] where a law of $\sqrt{\log K}$ is demonstrated for lognormal fading). It is also noted that although the high-load characterization is similar to the high-SNR characterization (where $\log\log K$ takes the role of $\log P$) there is a main difference between the two asymptotic regimes. The difference lie in the fact that while the SNR is related to the input signals and additive noise vector statistics, the number of users is related to the resulting fading statistics via the user selection scheme. This fact makes the high-load parameters hard to calculate and much more setting-dependent than the general expression derived for the high-SNR parameters [19].

V. Sum-Rate Analysis

In this section the achievable per-cell ergodic sum-rate of the ZFBF is analyzed for Gaussian and Rayleigh block flat-fading channels for specific user scheduling procedures. For the fading setup we also address the high-load characterization, and the implications of increasing number of users per-cell over the transmitted power of an arbitrary cell-site.

A. Non-Fading Setup

For non-fading channels, a *round-robin* scheduling is deployed and there is no need to feed back the channel coefficients since $a_{m,k} = b_{m,k} = c_{m,k} = 1 \ , \forall m, \ k$. Hence, for each time slot, the channel transfer matrix (2) becomes circulant with $(1, \ \alpha, \ 0, \ \ldots \ 0, \ \alpha)$ as first row, and the following proposition holds.

Proposition 3 *The per-cell average sum-rate of the ZFBF scheme is given for,* $\alpha < 1/2$*, by*

$$R_{\text{zfbf}-\text{nf}} \underset{M\to\infty}{=} \log\left(1 + \left(1 - 4\alpha^2\right)^{\frac{3}{2}} \ P\right) \ . \tag{18}$$

This rate holds for an overall power constraint MP, and for an equal per-cell power constraints P.

Proof: See Appendix C. ■

Evidently, $R_{\mathrm{zfbf-nf}}$ is a decreasing function of the interference factor α. Comparing (11) to (18), it is clear that the ZFBF scheme is superior to the ICTS scheme when the SNR P is above a certain threshold

$$P_{\mathrm{t}}(\alpha) = \frac{2\left(1-(1-4\alpha^2)^{\frac{3}{2}}\right)}{(1-4\alpha^2)^3} \ , \tag{19}$$

which is an increasing function of α. It is noted that for $\alpha = 1/2$ the circulant channel transfer matrix $\boldsymbol{H}$ is singular and channel inversion methods such as ZFBF are not applicable. Moreover, $\boldsymbol{H}$ is not guaranteed to be non-singular for $\alpha > 0.5$ and any finite number of cells M.

B. Rayleigh Fading Setup

For the Rayleigh fading setup, for each fading block (or TDMA slot) the multi-cell processor selects the user with the *"best" local received power* (BLP) for transmission in each cell. In other words, the selected user in the m'th cell is

$$\tilde{k}(m) = \underset{k}{\mathrm{argmax}}\{|a_{m,k}|^2\} \ , \tag{20}$$

where $\{a_{m,k}\}_{k=1}^{K}$ are the fading coefficients of the m'th cell transmitted signals as they are received by the m'th cell users.

The resulting channel transfer matrix of this sub-optimal scheduling $\boldsymbol{H}$ defined in (1), consists of diagonal entries $a_m = a_{m,\tilde{k}(m)}$ whose amplitudes are the *maximum* of K i.i.d. chi-square distributed random variables with two degrees of freedom. The other two diagonals entries of $\boldsymbol{H}$ are chi-square distributed random variables with two degrees of freedom times α.

In case $\boldsymbol{H}$ is ill conditioned, the joint beamformer can start replacing the "best" users by their second "best" users until the resulting $\boldsymbol{H}$ is well behaved. Since we assume that $K \gg 1$, the overall statistics is not expected to change by this user replacing procedure.

The special structure of the channel transfer matrix $\boldsymbol{H}$ resulting from the setup topology and the scheduling procedure, plays a key role in understanding the asymptotic high-load characterization of the scheme's per-cell sum-rate R_{zfbf} (expression (5)), which is stated in the following proposition.

Proposition 4 *The high-load characterization of the ZFBF-BLP scheme is given by*

$$\mathsf{S}_{\infty}^{\mathrm{blp}} = 1 \quad ; \quad \mathsf{L}_{\infty}^{\mathrm{blp}} = \log P \ . \tag{21}$$

Proof: See Appendix D. ■

This results, can be intuitively explained by the fact that due to the scheduling process, $(\boldsymbol{H}\boldsymbol{H}^{\dagger})$ "becomes" diagonal $(\log K \boldsymbol{I}_M)$ when K increases. Accordingly, for large K, $(\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}$ "behaves" like $(\boldsymbol{I}_M/\log K)$, and R_{zfbf} (expression (5)) is well approximated for the BLP selecting scheme by

$$R_{\mathrm{blp}} \underset{K\gg 1}{\cong} \log(1 + P \log K) \ . \tag{22}$$

It is concluded that the per-cell sum-rate of the ZFBF-BLP scheme scales as $\log\log K$ which is the same scaling law of the optimal multi-cell joint DPC scheme. However, it offers a smaller high-load offset than that of the offset predicted by the upper bound of (10)

$$0 \leq \mathsf{L}_{\infty}^{\mathrm{opt}} - \mathsf{L}_{\infty}^{\mathrm{blp}} \leq \log(1 + 2\alpha^2) \ . \tag{23}$$

It is also concluded that the ZFBF-BLP scheme provides a two fold scaling law than that of the ICTS scheme (12), in the presence of Rayleigh fading. Moreover, by definition, the sum-rate of the ZFBF scheme ensures a non-interference limited behavior for any number of users K (not necessarily large).

Finally, we consider the power constraint issue asymptotically with increasing number of users per-cell.

Proposition 5 *The considered ZFBF-BLP scheme, that maintains an overall power constraint of MP, ensures in probability an equal per-cell power constraint of P, asymptotically with increasing number of users per-cell. Hence,*

$$P_m \underset{K\to\infty}{\xrightarrow{p}} P \quad ; \quad m = 0,\ 1, \ldots,\ M-1 \ . \tag{24}$$

Proof: See Appendix E. ■

As mentioned earlier, for cellular systems an individual per-cell power constraint is a more reasonable choice than a sum-power constraint which is more suitable for compact antenna arrays.

VI. DISCUSSION

In this section we consider a more intuitive scheduling scheme for the multi-cell setup and show that it provides no better rates than the BLP scheme in the high-load regime.

One of the reasons the BLP user selection scheme was chosen to be the focal point of this work is that it enables analytical treatment, especially in deriving a lower bound to the per-cell sum-rate in the presence of Rayleigh fading. In this section we present a more intuitive user selection procedure referred to as *best total received power* (BRP), and we show that it does not provide higher rates in the high-load regime.

According to this scheme, for each fading block (or TDMA slot) the multi-cell processor selects the user with the best over-all received power, for transmission in each cell. In other words, the selected user in the m'th cell is

$$\tilde{k}(m) = \underset{k}{\operatorname{argmax}} \left\{ |a_{m,k}|^2 + \alpha^2 (|b_{m,k}|^2 + |c_{m,k}|^2) \right\} \ , \tag{25}$$

where $\{a_{m,k}\}_{k=1}^{K}$, $\{b_{m,k}\}_{k=1}^{K}$, and $\{c_{m,k}\}_{k=1}^{K}$ are the fading coefficients of the m'th, $\widehat{m-1}$'th, and $\widehat{m+1}$'th cell transmitted signals respectively, as they are received by the m'th cell users. In case $\boldsymbol{H}$ is ill conditioned, the same procedure of replacing "best" users with second "best" users id deployed.

Proposition 6 *The per-cell sum-rate of the ZFBF-BRP scheme is upper bounded for any number of users per-cell K by*

$$R_{\mathrm{brp}} \leq R_{\mathrm{brp-ub}} \triangleq \log \left(1 + P \ \frac{\log(2K+1) + 3 \log\log K}{1 - \frac{1}{\log K}} \right) \ . \tag{26}$$

Proof: See Appendix H. ■

Applying the definitions of the high-load slope and offset (expressions (14) and (15)) directly to the upper bound of (26) yields the following

$$\mathsf{S}_{\infty}^{\mathrm{brp-ub}} = 1 \quad ; \quad \mathsf{L}_{\infty}^{\mathrm{brp-ub}} = \log P \ . \tag{27}$$

Since the high-load slope and offset predicted by the upper bound equals to ones of the ZFBF-BPL scheme, it is concluded that the ZFBF-BRP scheme does not provide higher rates in the high-load regime, than the ZFBF-BLP scheme. Although it is conjectured that Proposition 5 also hold for the ZFBF-BRP scheme, the issue of the individual cells transmitted power does not change the latter conclusion.

VII. Numerical Results

First, the non-fading setup is considered under the assumption that the number of cells is large $M \gg 1$ and $\alpha < 1/2$. In Fig. 2 the spectral efficiencies[4] per-cell of the optimal, ICTS, and ZFBF schemes (expressions (9), (11), and (18) respectively) are plotted as a function of the transmitted E_b/N_0, for $\alpha = 0.4$. It is observed that the ZFBF scheme outperforms the ICTS scheme above a certain SNR threshold. The threshold $P_t(\alpha)$ (expression (19)) is shown in Fig. 3 as a function of the inter-cell interference factor α ; the ICTS scheme is superior in the region below this curve (which is a monotonically increasing function of the α), while the ZFBF scheme prevails in the region above the curve.

Turning to the Rayleigh fading setup, the spectral efficiencies per-cell (calculated by Monte-Carlo simulations) of the ICTS and ZFBF-BLP (expression (5)) schemes, and the asymptotic upper bound of the optimal scheme (expression (10)) are plotted in Fig. 4 as a function of the transmitted E_b/N_0, for $\alpha = 0.4$, $K = 100$, and finite dimensional system of $M = 30$ cells[5]. It is observed that for this set of parameters the ZFBF-BLP scheme loses only a fraction of a bit/sec/Hz when compared to the upper bound of the optimal scheme already for a modest number of users per-cell (it is noted that the upper bound is valid for $K \gg 1$ and it might not be accurate for small values of K). The gap between the ZFBF-BLP curve and the per-cell sum-rate capacity upper bound is clearly explained by the fact that the ZFBF-BLP scheme does not use the distributed antenna array to enhance the reception power, but to eliminate inter-cell interferences. Hence, the additional array power gain of $(1+2\alpha^2)$ predicted by the upper bound cannot be achieved. Moreover, for large values of E_b/N_0, the ZFBF provides approximately twice bits/sec/Hz than the ICTS scheme, which can be explained by the 0.5 high-load slope of the ICTS per-cell sum-rate expression (17).

In Fig. 5 the sum-rates per-cell of the ICTS and ZFBF-BLP schemes (Monte-Carlo simulations, and asymptotic expressions (22) and (12)), the ZFBF-BRP (Monte-Carlo simulation) and the upper bound of the optimal scheme, are plotted as a function of the number of users per-cell for $P = 10$ [dB], $\alpha = 0.4$ and $M = 30$. Examining the curves, the observations made for Fig.

[4]The spectral efficiency $\mathsf{C}(E_b/N_0)$ is defined through the following relations: $\mathsf{C}(E_b/N_0) = C(\mathrm{SNR})$ and $\mathrm{SNR} = C(\mathrm{SNR})E_b/N_0$.

[5]It is noted that a circular setup of $M = 30$, may be considered for any practical purpose as an infinite array [8].

4 are strengthened, since the small loss suffered by the ZFBF scheme when compared to the upper bound is demonstrated to hold over a wide range of K values. A good match between the ZFBF-BLP Monte-Carlo simulation results to its asymptotic curve is observed as well, already for a modest number of users per cell. Another observation is that the BRP scheduling does not provide higher rates than the BLP scheduling, demonstrating that claim of Proposition 4 is also valid for a modest number of users per-cell.

In Figure 6 the empiric cumulative distribution function (CDF) of of an arbitrary cell normalized transmit power P_m/P is plotted for several values of the number of users per-cell K and $P = 10$ [dB]. The curves, derived by Monte-Carlo simulation, demonstrate the convergence of P_m to P as predicted by Proposition 5 ; the probability of P_m to be in the rage of $P \pm 1$ [dB] increases from 0.4 to 0.84, while the number of users increases from $K = 10$ to $K = 1000$.

VIII. Concluding Remarks

In this work a cooperative multi-cell ZFBF scheme for the downlink of the circular Wyner model is considered for Gaussian and flat Rayleigh fading channels. For the non fading setup and round-robin scheduling, a closed form expression for the per-cell sum-rate (under both, overall and equal per-cell power constraints), is derived. The latter demonstrates superior performance over the ICTS scheme when the SNR crosses a certain threshold, which increases with the inter-cell interference power level (increasing α).

To address the asymptotic analysis for the Rayleigh fading setup, we introduce the high-load characterization through the slope and offset parameters which accompanied the $\log\log K$ law. Next, the per-cell sum-rate of the ZFBF-BLP scheme is proved to demonstrate the same growth rate of $\log\log K$ and a slightly degraded offset when compared to the optimal DPC scheme, asymptotically with increasing number of users per-cell K, while satisfying (in probability) equal per-cell power constraints. We also show that the alternative ZFBF-BRP scheme does not provide better rates in the high-load regime. Furthermore, numerical results derived by Monte-Carlo simulation show a good match to the results predicted by the various analyses included, already for a modest number of users. It is noted that since ZFBF is utilized, non-interference behavior is guaranteed for any number of users per-cell K (not necessarily large). Moreover, extending the results presented here to a two dimensional planner Wyner model can be done in a straightforward (though tedious) manner.

It is concluded that the simple cooperative multi-cell ZFBF-BLP scheme presented here provides near optimal performance already for a moderate number of users per-cell. Moreover, the scheduling mechanism facilitates one active user per-cell on one hand (a natural choice for cellular systems), while treating the whole system as a distributed antenna array for inter-cell interference cancellation by ZFBF (allowing the use of single-user encoding-decoding schemes), on the other. Combined with the inherent non-interference limited behavior, the ZFBF-BLP scheme provides a fair alternative to the optimal complex joint multi-cell DPC scheme.

APPENDIX

A. Proof of Proposition 1

The claim is proved by the following set of inequalities

$$\begin{aligned} R_{\text{zfbf}} &= E\left\{\log\left(1+\frac{MP}{\text{tr}\left((\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right)}\right)\right\} \\ &= E\left\{\log\left(1+\frac{P}{\frac{1}{M}\sum_m \frac{1}{\lambda_m(\boldsymbol{H}\boldsymbol{H}^{\dagger})}}\right)\right\} \\ &\overset{(a)}{\leq} E\left\{\log\left(1+\frac{P}{M}\sum_m \lambda_m(\boldsymbol{H}\boldsymbol{H}^{\dagger})\right)\right\} \\ &\leq E\left\{\log\left(1+\frac{P}{M}\,\text{tr}\left(\boldsymbol{H}\boldsymbol{H}^{\dagger}\right)\right)\right\} \\ &\overset{(b)}{\leq} \log\left(1+\frac{P}{M}E\left\{\text{tr}\left(\boldsymbol{H}\boldsymbol{H}^{\dagger}\right)\right\}\right)\ , \end{aligned} \tag{28}$$

where (a) is achieved since the arithmetic mean of a non-negative set is larger than its harmonic mean, and (b) is achieved by Jensen's inequality.

B. Proof of Proposition 2

To prove the RHS of (8) we note that P_m (expression (7)) satisfies the following set of inequalities

$$\begin{aligned} P_m &= \frac{MP\left[(\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right]_{m,m}}{\text{tr}\left((\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right)} \\ &\leq P\left(\max_n\left[(\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right]_{n,n}\right)\left(\max_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^{\dagger})\right) \\ &\overset{(a)}{\leq} P\left(\max_n \lambda_n((\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1})\right)\left(\max_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^{\dagger})\right) \\ &= P\frac{\max_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^{\dagger})}{\min_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^{\dagger})}\ , \end{aligned} \tag{29}$$

where (a) is achieved by recalling that the eigenvalues of an Hermitian matrix majorize its diagonal entries (Horn's Theorem [25]).

Next, we denote the non-increasing ordered eigenvalues and non-increasing ordered diagonal entries of a semi-positive definite (SPD) Hermitian $M \times M$ matrix $\boldsymbol{A}$ by $\{\lambda_n^{\text{o}}(\boldsymbol{A})\}_{n=0}^{M-1}$ and $\{d_n^{\text{o}}(\boldsymbol{A})\}_{n=0}^{M-1}$ respectively. Now to complete the proof, the RHS of (8) is justified by the following

set of inequalities

$$\begin{aligned} P_m &= \frac{MP\left[(\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right]_{m,m}}{\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right)} \\ &\geq \left(\min_n \left[(\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right]_{n,n}\right)\left(\min_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^\dagger)\right) \\ &= P\left(\lambda^{\mathrm{o}}_{M-1}((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}) + \sum_{n=0}^{M-2}\left(\lambda^{o}_n((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}) - d^{\mathrm{o}}_n((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1})\right)\right)\left(\min_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^\dagger)\right) \\ &\overset{(a)}{\geq} P\left(\lambda^{\mathrm{o}}_{M-1}((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1})\right)\left(\min_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^\dagger)\right) \\ &= P\frac{\min_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^\dagger)}{\max_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^\dagger)}\ , \end{aligned} \tag{30}$$

where (a) is achieved by recalling that the eigenvalues of an Hermitian matrix majorize its diagonal entries (Horn's Theorem [25]), hence the summation portion is non-negative.

C. Proof of Proposition 3

Since for the non-fading setup $\boldsymbol{H}$ is deterministic and (5) reduces to

$$R_{\mathrm{zfbf-nf}} = \log\left(1 + \frac{MP}{\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right)}\right)\ . \tag{31}$$

To evaluate the inner log term of (31), the following set of equalities are useful

$$\begin{aligned} \frac{1}{M}\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right) &\overset{(a)}{=} \frac{1}{M}\sum_{m=0}^{M-1}\frac{1}{\lambda_m\left(\boldsymbol{H}\boldsymbol{H}^\dagger\right)} \\ &\overset{(b)}{=} \frac{1}{M}\sum_{m=0}^{M-1}\frac{1}{\lambda^2_m(\boldsymbol{H})} \\ &= \frac{1}{M}\sum_{m=0}^{M-1}\left(\frac{1}{1+2\alpha\cos\left(\frac{2\pi m}{M}\right)}\right)^2\ , \end{aligned} \tag{32}$$

where (a) and (b) are achieved since $\boldsymbol{H}$ is an Hermitian non-singular $M \times M$ matrix for $0 \leq \alpha < 1/2$, and the last equality is due to the fact that $\boldsymbol{H}$ is also circulant [26]. Taking M to infinity yields

$$\begin{aligned} \frac{1}{M}\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right) &\underset{M\to\infty}{=} \int_0^1 \left(1+2\alpha\cos(2\pi\theta)\right)^{-2} d\theta \\ &= \left(1-4\alpha^2\right)^{\frac{3}{2}}\ , \end{aligned} \tag{33}$$

where the last equality is achieved by setting $a = 1$, $b = 2\alpha$, $n = 2$ in equation 3.661.4 of [27, pp. 399] and some algebra. Substituting (33) into (31) yields (18).

Since, for the non-fading setup, $\boldsymbol{H}$ is a circulant matrix, then according to [26], $(\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}$ is also circulant, and by definition its diagonal entries are equal. Hence, the average transmit power of the m'th cell site antenna is given by

$$P_m = [\boldsymbol{B}\boldsymbol{B}^{\dagger}]_{m,m} = \frac{MP\left[(\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right]_{m,m}}{\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right)} = P \ , \tag{34}$$

where $\boldsymbol{B}$ is the beamforming matrix defined in (3). It is concluded that the overall power constraint of MP ensures an equal per-cell power constraints of P.

D. Proof of Proposition 4

In order to prove the claim we need to provide a lower bound to the the sum-rate. Towards this end we start by showing that the channel transfer matrix $\boldsymbol{H}$, resulting from the "best" local channel selection procedure, satisfies the following Propositions.

Proposition 7 *The Frobenius norm of the matrix $(\boldsymbol{H}\boldsymbol{H}^{\dagger}/\log K - \boldsymbol{I}_M)$ converges in probability to 0. Hence,*

$$\left\|\boldsymbol{H}\boldsymbol{H}^{\dagger}/\log K - \boldsymbol{I}_M\right\|_{\mathrm{F}} \xrightarrow[K\to\infty]{p} 0 \ , \tag{35}$$

where $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm of a matrix.

Proof: See Appendix F. ■

Proposition 8 *The eigenvalues of the matrix $(\boldsymbol{H}\boldsymbol{H}^{\dagger}/\log K)$ converge in probability to 1. Hence,*

$$\lambda_m(\boldsymbol{H}\boldsymbol{H}^{\dagger})/\log K \xrightarrow[K\to\infty]{p} 1 \quad , \quad \forall m \ . \tag{36}$$

Proof: See Appendix G. ■

Now, let us define the following event

$$\mathcal{A} = \{|\lambda_m/\log K - 1| < \epsilon \ , \forall m\} \ , \tag{37}$$

where λ_m is the m'th eigenvalue of $(\boldsymbol{H}\boldsymbol{H}^{\dagger})$. Next, we apply the high-load slope definition (14) to the per-cell sum-rate (5) and write the following set of inequalities

$$\begin{aligned}\frac{R_{\rm blp}}{\log\log K} &= E\left\{\log\left(1+\frac{MP}{\sum_m \frac{1}{\lambda_m}}\right)\right\}/\log\log K \\ &\geq E\left\{\mathbf{1}_{\mathcal{A}}\log\left(1+\frac{MP}{\sum_m \frac{1}{\lambda_m}}\right)\right\}/\log\log K \\ &\stackrel{(a)}{\geq} Pr(\mathcal{A})\frac{\log\left(1+(1-\epsilon)P\log K\right)}{\log\log K} \\ &\stackrel{(b)}{\geq} \left(\sum_m Pr\left(\left|\frac{\lambda_m}{\log K}-1\right|<\epsilon\right)-(M-1)\right) \\ &\qquad \frac{\log\left(1+(1-\epsilon)P\log K\right)}{\log\log K} \underset{K\to\infty}{\longrightarrow} 1\ ,\end{aligned} \tag{38}$$

where $\mathbf{1}_{\mathcal{A}}$ is an indicator function, $\epsilon>0$ is an arbitrary small constant, (a) is achieved by noting that the definition of $\mathcal{A}$ implies that $\lambda_m > (1-\epsilon)\log K$, and (b) is achieved by using the following inequality

$$Pr\left(\bigcap_{n=1}^{N}\mathcal{B}_n\right) \geq \sum_{n=1}^{N} Pr(\mathcal{B}_n) - (N-1)\ , \tag{39}$$

where $\{\mathcal{B}_n\}_{n=1}^{N}$ is a set of arbitrary events. In addition, the final limit of (38) is achieved by invoking Proposition 8 and taking K to infinity.

Since the scheme achieves the optimal high-load slope $\mathsf{S}_{\infty}^{\rm blp}\geq 1$ it is evident that $\mathsf{S}_{\infty}^{\rm blp}=1$. Nevertheless, we show that $\mathsf{S}_{\infty}^{\rm blp}\leq 1$ for the sake of completeness. Towards this end, and for the later calculation of the high-load offset $\mathsf{L}_{\infty}^{\rm blp}$, the following bound is useful

$$\begin{aligned}R_{\rm blp} &= E\left\{\log\left(1+\frac{MP}{\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1}\right)}\right)\right\} \\ &\stackrel{(a)}{\leq} \log\left(1+\frac{P}{M}E\left\{\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{H}^{\dagger}\right)\right\}\right) \\ &= \log\left(1+\frac{P}{M}\sum_m E\left\{|a_m|^2+\alpha^2(|b_m|^2+|c_m|^2)\right\}\right) \\ &\stackrel{(b)}{=} \log\left(1+\frac{P}{M}\sum_m\left(E\{|a_m|^2\}+2\alpha^2\right)\right) \\ &\stackrel{(c)}{\leq} \log\left(1+P\left(\log(2K+1)+2\alpha^2\right)\right)\ ,\end{aligned} \tag{40}$$

where (a) is due to Proposition 1, (b) is achieved due to the fact that according to the BLP selection procedure $|b_m|^2,\ |c_m|^2 \sim \chi_2^2$ with $E\{|b_m|^2\} = E\{|c_m|^2\} = 1$, and (c) is achieved by utilizing a known result regarding the mean value of the r'th order statistic of n independent exponential (or central χ_2^2 with $\sigma^2 = 1/2$) distributed r.v's (see [28, pp. 62]) stating that

$$E\{X_{r:n}\} \leq \log \frac{n+\frac{1}{2}}{n-r+\frac{1}{2}} \ . \tag{41}$$

Applying the upper bound of (40), it is easily verified that the high-load slope of the ZFBF-BLP scheme is upper bounded by

$$\mathsf{S}_\infty^{\text{blp}} \leq \lim_{K\to\infty} \frac{\log\left(1+P\left(\log(2K+1)+2\alpha^2\right)\right)}{\log\log K} = 1 \ . \tag{42}$$

Turning to the high-load offset we apply the high-load slope definition (15) to the per-cell sum-rate (5) and write the following set of inequalities

$$\begin{aligned}
\frac{R_{\text{blp}}}{\mathsf{S}_\infty^{\text{blp}}} - \log\log K &= E\left\{\log\left(\frac{1}{\log K} + \frac{MP}{\sum_m \frac{\log K}{\lambda_m}}\right)\right\} \\
&\geq E\left\{\mathbf{1}_{\mathcal{A}}\log\left(\frac{1}{\log K} + \frac{MP}{\sum_m \frac{\log K}{\lambda_m}}\right)\right\} \\
&\overset{(a)}{\geq} Pr(\mathcal{A})\log\left(\frac{1}{\log K} + (1-\epsilon)P\right) \\
&\overset{(b)}{\geq} \left(\sum_m Pr\left(\left|\frac{\lambda_m}{\log K} - 1\right| < \epsilon\right) - (M-1)\right)\log\left(\frac{1}{\log K} + (1-\epsilon)P\right) \\
&\underset{K\to\infty}{\longrightarrow} \log\left((1-\epsilon)P\right) \ ,
\end{aligned} \tag{43}$$

where (a) is achieved due to the definition of $\mathcal{A}$ which implies that $\lambda_m > (1-\epsilon)\log K$, and (b) is due to 39. The final limit of (43) is achieved by invoking Proposition 8 and taking K to infinity.

Since the high-load lower bound of (43) holds for arbitrary small ϵ it is concluded that

$$\mathsf{L}_\infty^{\text{blp}} \geq \log P \ . \tag{44}$$

On the other hand, applying the upper bound of (40) we get that

$$\mathsf{L}_\infty^{\text{blp}} \leq \lim_{K\to\infty} \log\left(\frac{1}{\log K} + P\frac{\left(\log(2K+1)+2\alpha^2\right)}{\log K}\right) = \log P \ . \tag{45}$$

Combining (44) and (45) the proof is completed.

E. Proof of Proposition 5

To prove the claim, it is enough to show that the random variable P_m satisfies

$$Pr\left(|P_m - P| \leq \epsilon\right) \underset{K\to\infty}{\longrightarrow} 1 \ , \tag{46}$$

for any arbitrarily small $\epsilon > 0$.

Now, Let us define $\bar{\lambda} \triangleq \max_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^\dagger)/\log K$, $\underline{\lambda} \triangleq \min_n \lambda_n(\boldsymbol{H}\boldsymbol{H}^\dagger)/\log K$, and rewrite (46) as follows

$$\begin{aligned}
Pr\left(|P_m - P| \leq \epsilon\right) &= Pr\left(\frac{P_m}{P} \leq 1 + \frac{\epsilon}{P} \bigcap \frac{P_m}{P} \geq 1 - \frac{\epsilon}{P}\right) \\
&\overset{(a)}{\geq} Pr\left(\frac{\bar{\lambda}}{\underline{\lambda}} \leq 1 + \frac{\epsilon}{P} \bigcap \frac{\underline{\lambda}}{\bar{\lambda}} \leq 1 + \frac{\epsilon}{P}\right) \\
&\overset{(b)}{\geq} Pr\left(\frac{\bar{\lambda}}{\underline{\lambda}} \leq 1 + \frac{\epsilon}{P} \bigcap \frac{\underline{\lambda}}{\bar{\lambda}} \geq 1 - \frac{\epsilon}{P} \bigcap \left|\bar{\lambda} - 1\right| < \epsilon_1 \bigcap |\underline{\lambda} - 1| < \epsilon_2\right) \\
&\overset{(c)}{\geq} Pr\left(\frac{1+\epsilon_1}{1-\epsilon_1} \leq 1 + \frac{\epsilon}{P} \bigcap \frac{1-\epsilon_2}{1+\epsilon_2} \geq 1 - \frac{\epsilon}{P} \bigcap \left|\bar{\lambda} - 1\right| < \epsilon_1 \bigcap |\underline{\lambda} - 1| < \epsilon_2\right) \\
&\overset{(d)}{=} Pr\left(\left|\bar{\lambda} - 1\right| < \epsilon_1 \bigcap |\underline{\lambda} - 1| < \epsilon_2\right) \\
&\overset{(e)}{\geq} Pr\left(\left|\bar{\lambda} - 1\right| < \epsilon_1\right) + Pr\left(|\underline{\lambda} - 1| < \epsilon_2\right) - 1 \underset{K\to\infty}{\longrightarrow} 1 \ ,
\end{aligned} \tag{47}$$

where, (a) is due to (8), (b) is due to the fact that $Pr(\mathcal{A}) \geq Pr(\mathcal{A} \cap \mathcal{B})$ where $\mathcal{A}$, $\mathcal{B}$ are arbitrary events, (c) is achieved by increasing $\bar{\lambda}/\underline{\lambda}$ and decreasing $\underline{\lambda}/\bar{\lambda}$, (d) is achieved by setting $\epsilon_1 < \epsilon/(2P + \epsilon)$ and $\epsilon_2 < \epsilon/(2P - \epsilon)$, hence, ensuring that the first two events have probability 1, and (e) is achieved by invoking (39) (setting $N = 2$). Finally, the last limit is due to Proposition 8.

F. Proof of Proposition 7

To prove the claim we have to show that

$$Pr\left(\left\|\boldsymbol{H}\boldsymbol{H}^\dagger/\log K - \boldsymbol{I}_M\right\|_{\mathrm{F}} > \epsilon\right) \underset{K\to\infty}{\longrightarrow} 0 \ . \tag{48}$$

Using the definition of the Frobenius norm [26], the LHS of (48) is rewritten as

$$\begin{aligned}
&Pr\left(\sum_{m=0}^{M-1}\sum_{n=0}^{M-1}\left|\left[\boldsymbol{H}\boldsymbol{H}^\dagger/\log K - \boldsymbol{I}_M\right]_{m,n}\right|^2 > \epsilon^2\right) \\
&\qquad\qquad \leq \sum_{m=0}^{M-1}\sum_{n=0}^{M-1} Pr\left(\left|\left[\boldsymbol{H}\boldsymbol{H}^\dagger/\log K - \boldsymbol{I}_M\right]_{m,n}\right|^2 > \frac{\epsilon^2}{5M}\right) \ ,
\end{aligned} \tag{49}$$

where the last inequality is achieved by recalling that for any set of non-negative r.v.'s, $\{x_n\}_{n=1}^N$, the following inequality holds

$$Pr\left(\sum_{n=1}^{N} x_n > \delta\right) \leq \sum_{n=1}^{N} Pr(x_n > \delta/N) \ . \tag{50}$$

and that the matrix $\boldsymbol{H}\boldsymbol{H}^{\dagger}$ has $5M$ non-zero entries.

Examining the five-diagonal matrix $\boldsymbol{H}\boldsymbol{H}^{\dagger}$, resulting from the BLP scheduling, reveals that its non-zero entries can be divided into three groups of identical distributed r.v.'s: (1) main diagonal entries indexed (m,m), (2) the entries of the first diagonals below and above the main diagonal, indexed $(m,\widehat{m\pm 1})$, and (3) the entries of the second diagonals below and above the main diagonal, indexed $(m,\widehat{m\pm 2})$. Hence, the RHS of (49) boils down, for an arbitrary m, to

$$\begin{aligned} &MPr\left(\left|\frac{|a_m|^2+\alpha^2|b_m|^2+\alpha^2|c_m|^2}{\log K}-1\right|^2>\frac{\epsilon^2}{5M}\right)+\\ &2MPr\left(\left|\frac{\alpha b_m a^*_{\widehat{m-1}}+\alpha a_m c^*_{\widehat{m-1}}}{\log K}\right|^2>\frac{\epsilon^2}{5M}\right)+\\ &2MPr\left(\left|\frac{\alpha^2 b_m c^*_{\widehat{m-2}}}{\log K}\right|^2>\frac{\epsilon^2}{5M}\right) \ , \end{aligned} \tag{51}$$

where for Rayleigh fading $b_m, c_m, c_{\widehat{m-2}}, c_{\widehat{m-1}} \sim \mathcal{CN}(0,1)$, and $a_m, a_{\widehat{m-1}}$ each has the maximum amplitude of K i.i.d. χ^2_2 distributed r.v.'s.

In the sequel the following order statistic result is useful. According to [23] (Example 1, Appendix A), the maximum of N χ^2_2 distributed r.v.'s, x, behaves like $\log N$ with high probability. In particular, while neglecting little orders of $\log\log N$, x satisfies

$$\Pr\left(|x-\log N|\leq \log\log N\right) > 1-O\left(\frac{1}{\log N}\right) \ . \tag{52}$$

Next, we denote $d \triangleq \alpha^2(|b_m|^2+|c_m|^2)$ and rewrite the first summand of (51) as follows

$$\begin{aligned} & Pr\left(\left|\frac{|a_m|^2+d}{\log K}-1\right|^2>\frac{\epsilon^2}{5M}\right) \\ & \overset{(a)}{\leq} Pr\left(\left||a_m|^2-\log K\right|+d>\frac{\epsilon\log K}{\sqrt{5M}}\right) \\ & \overset{(b)}{\leq} Pr\left(\left(\left||a_m|^2-\log K\right|+d\right)>\frac{\epsilon\log K}{\sqrt{5M}}\bigcap\left||a_m|^2-\log K\right|\leq\log\log K\right)+O\left(\frac{1}{\log K}\right) \\ & \leq Pr\left(\log\log K+d>\frac{\epsilon\log K}{\sqrt{5M}}\right)+O\left(\frac{1}{\log K}\right) \\ & \overset{(c)}{=} e^{-g_1(\epsilon,K,M,\alpha)}(1+g_1(\epsilon,K,M,\alpha))+O\left(\frac{1}{\log K}\right)\underset{K\to\infty}{\longrightarrow}0\ , \end{aligned} \tag{53}$$

where

$$g_1(\epsilon,K,M,\alpha)\triangleq\frac{1}{\alpha^2}\left(\frac{\epsilon\log K}{\sqrt{5M}}-\log\log K\right)\underset{K\to\infty}{\longrightarrow}\infty\ , \tag{54}$$

(a) is achieved by algebraic manipulation and by invoking the triangular inequality, (b) is due to the fact that $Pr(\mathcal{A})\leq Pr(\mathcal{A}\cap\mathcal{B})+Pr(\mathcal{B}^c)$ for any events $\mathcal{A},\mathcal{B}$, and by noting that (52) implies that $Pr(|\,|a_m|^2-\log K|>\log\log K)<O(1/\log K)$, and (c) is due to the fact that d/α^2 is a χ^2_4 distributed r.v. .

Turning to the second summand of (51) we have

$$\begin{aligned} & Pr\left(\left|\frac{\alpha b_m a^*_{\widehat{m-1}}+\alpha a_m c^*_{\widehat{m-1}}}{\log K}\right|^2>\frac{\epsilon^2}{5M}\right) \\ & \qquad\leq Pr\left(\max(|a_{\widehat{m-1}}|^2,|a_m|^2)\frac{(|b_m|+|c_{\widehat{m-1}}|)^2}{(\log K)^2}>\frac{\epsilon^2}{5M\alpha^2}\right)\ , \end{aligned} \tag{55}$$

where the last inequality is achieved by invoking the triangular inequality and some algebraic manipulations. Let us define $d\triangleq\max(|a_{\widehat{m-1}}|^2,|a_m|^2)$ and $f\triangleq(|b_m|+|c_{\widehat{m-1}}|)^2$ and rewrite the

RHS of (55) as follows

$$\begin{aligned}
&Pr\left(d\ f > \frac{\epsilon^2(\log K)^2}{5M\alpha^2}\right) \\
&\overset{(a)}{\leq} Pr\left(d\ f > \frac{\epsilon^2(\log K)^2}{5M\alpha^2}\bigcap |d-\log(2K)| < \log\log(2K)\right) + O\left(\frac{1}{\log(2K)}\right) \\
&\leq Pr\left(f > \frac{\epsilon^2(\log K)^2}{5M\alpha^2(\log(2K)+\log\log(2K))}\right) + O\left(\frac{1}{\log(2K)}\right) \\
&\overset{(c)}{\leq} Pr\left(|c_{\widehat{m-1}}|^2 > g_2(\epsilon,K,M,\alpha)\right) + Pr\left(|b_m|^2 > g_2(\epsilon,K,M,\alpha)\right) + O\left(\frac{1}{\log(2K)}\right) \\
&\overset{(d)}{=} 2e^{-g_2(\epsilon,K,M,\alpha)} + O\left(\frac{1}{\log(2K)}\right) \underset{K\to\infty}{\longrightarrow} 0\ ,
\end{aligned} \tag{56}$$

where

$$g_2(\epsilon,K,M,\alpha) \triangleq \frac{\epsilon^2(\log K)^2}{20M\alpha^2(\log(2K)+\log\log(2K))} \underset{K\to\infty}{\longrightarrow} \infty\ , \tag{57}$$

(a) is achieved by noting that d is the maximum of $2K$ i.i.d. χ_2^2 distributed r.v.'s, and adhering to similar argumentation used in step (b) of (53), (c) is achieved by using (50) and some algebra, and (d) is due to the fact that $|c_{\widehat{m-1}}|^2$ and $|b_m|^2$ are χ_2^2 distributed r.v.'s.

Considering the third summand of (51) we get

$$\begin{aligned}
Pr\left(\left|\frac{\alpha^2 b_m c^*_{\widehat{m-2}}}{\log K}\right|^2 > \frac{\epsilon^2}{5M}\right) &\overset{(a)}{\leq} Pr\left(|b_m|^2 + |c_{\widehat{m-2}}|^2 > \frac{2\epsilon\log K}{\sqrt{5M}\alpha^2}\right) \\
&\overset{(b)}{\leq} e^{-g_3(\epsilon,K,M,\alpha)}(1+g_3(\epsilon,K,M,\alpha)) \underset{K\to\infty}{\longrightarrow} 0\ ,
\end{aligned} \tag{58}$$

where

$$g_3(\epsilon,K,M,\alpha) \triangleq \frac{2\epsilon\log K}{\sqrt{5M}\alpha^2} \underset{K\to\infty}{\longrightarrow} \infty\ , \tag{59}$$

(a) is achieved by recalling that for any two non-negative r.v.'s $x,\ y$, the following inequality holds $Pr(x\ y > \delta) \leq Pr(x+y > 2\sqrt{\delta})$, and (b) is due to the fact that $|b_m|^2 + |c_{\widehat{m-2}}|^2$ is a χ_4^2 distributed r.v. .

Remark: Examining (51), (53), (56), and (58) it is evident that in order for the Frobenius norm to converge we need that $M = o(\log K)$. This requirement that connects the number of cells M to the the number of users per-cell K is most unwanted, since it means that a huge number of users per-cell is needed already for a small number of cells. Fortunately, this requirement evolves from the lower bounding technique being used, and it is not an inherent genuine requirement.

In fact numerical results derived by Monte-Carlo simulation, reveal that the per-cell sum-rate demonstrates almost perfect match to its asymptotic expression and also a weak dependency on the number of cells M, already for a modest number of users per-cell K.

G. Proof of Proposition 8

Since the Frobenius norm of an arbitrary rectangular $M \times M$ Hermitian matrix $\boldsymbol{A}$ may be expressed as

$$\|\boldsymbol{A}\|_{\mathrm{F}} = \sqrt{\mathrm{tr}\left(\boldsymbol{A}^{\dagger}\boldsymbol{A}\right)} = \sqrt{\sum_{m=0}^{M-1} \lambda_m^2(\boldsymbol{A})} \ , \tag{60}$$

then according to Proposition 7, for any finite M we get

$$\begin{aligned} \left\| \frac{\boldsymbol{H}\boldsymbol{H}^{\dagger}}{\log K} - \boldsymbol{I}_M \right\|_{\mathrm{F}} &= \sqrt{\sum_{m=0}^{M-1} \lambda_m^2 \left(\frac{\boldsymbol{H}\boldsymbol{H}^{\dagger}}{\log K} - \boldsymbol{I}_M \right)} \\ &= \sqrt{\sum_{m=0}^{M-1} \left(\frac{\lambda_m(\boldsymbol{H}\boldsymbol{H}^{\dagger})}{\log K} - 1 \right)^2} \xrightarrow[K\to\infty]{p} 0 \ , \end{aligned} \tag{61}$$

and the proof is completed by noting that the last equality of (61) holds if and only if (36) holds.

H. Proof of Proposition 6

To assess the high-load characterization parameters of the ZFBF-BRP scheme, the following bound is useful

$$\begin{aligned} R_{\mathrm{brp}} &= E\left\{ \log\left(1 + \frac{MP}{\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^{\dagger})^{-1} \right)} \right) \right\} \\ &\overset{(a)}{\leq} \log\left(1 + \frac{P}{M} E\{\mathrm{tr}\left(\boldsymbol{H}\boldsymbol{H}^{\dagger}\right)\} \right) \\ &\leq \log\left(1 + \frac{P}{M} \sum_m E\left\{ \left[\boldsymbol{H}\boldsymbol{H}^{\dagger}\right]_{m,m}\Big|_{\alpha=1} \right\} \right) \\ &= \log\left(1 + \frac{P}{M} \sum_m E\left\{ |d_m|^2 \right\} \right) \ , \end{aligned} \tag{62}$$

where (a) is due to Proposition 1, and $|d_m|^2 \triangleq |a_m|^2 + |b_m|^2 + |c_m|^2$. It is noted that $|d_m|^2$ is the maximum of K i.i.d. χ_6^2 distributed r.v's with $\sigma^2 = 1/2$. Next we upper bound $E\left\{|d_m|^2\right\}$, utilizing order statistics argumentations.

Proposition 9 *The mean value of the r'th ($1 \leq r \leq n$) order statistics of a central χ^2_{2l} distribution satisfies*

$$E\{X_{r:n}\} \leq 2\sigma^2 \left(\frac{\log\left(\frac{2n+1}{2n-2r+1}\right) + l\log\frac{1}{\beta} + \log l - l}{1-l\beta} \right) , \tag{63}$$

where $0 < \beta < 1/l$.

Proof: See Appendix I. ■

For the current case of interest i.e. $l = 3$, $\sigma^2 = 1/2$, $r = n$, $n = K$, and $\beta = 1/(3\log K)$, we get that (63) boils down to

$$E\left\{|d_m|^2\right\} \leq \frac{\log(2K+1) + 3\log\log K + \log 3 + 3\log\log 3 - 3}{1 - \frac{1}{\log K}} \quad ; \quad \forall m. \tag{64}$$

Finally, substituting (64) in (62), noting that $\log 3 + 3\log\log 3 - 3 < 0$, and applying the definitions of the high-load slope and offset, the proof is completed.

I. Proof of Proposition 9

In [28, pp. 62], it is shown that the average of the r'th order statistics $\bar{x} = E\{X_{r:n}\}$ of an arbitrary distribution $P(x)$ satisfies the following

$$P(\bar{x}) \leq \frac{r}{n+\frac{1}{2}} , \tag{65}$$

if its *hazard rate*, defined as

$$h(x) \triangleq \frac{\frac{dP(x)}{dx}}{1-P(x)} , \tag{66}$$

is an increasing function.

It is easily verified that the hazard rate of a central χ^2_{2l} distribution is an increasing function. Hence, the average of its r'th order statistics satisfies (65). On the other hand, it is easy to verify that for the χ^2_{2l} distribution function also satisfies

$$P(x) \geq 1 - le^{-\frac{x}{2\sigma^2}} \left(\frac{x}{2\sigma^2}\right)^l \quad , \quad x \geq 0. \tag{67}$$

Combining (65) and (67) and some algebra we get the following inequality

$$\frac{\bar{x}}{2\sigma^2} \leq \log\frac{2n+1}{2n-2r+1} + l\log\frac{\bar{x}}{2\sigma^2} + \log l \ . \tag{68}$$

Finally, by invoking the following inequality

$$\log y \leq \log\frac{1}{\beta} - 1 + \beta y \quad , \quad y \geq 0,\ \beta > 0 \ , \tag{69}$$

to the second summand of the LHS of (68) and some algebra the proof is completed.

REFERENCES

[1] S. Shamai (Shitz), O. Somekh, and B. M. Zaidel, "Multi-cell communications: An information theoretic perspective," in *Proceedings of the Joint Workshop on Communications and Coding (JWCC'04)*, (Donnini, Florence, Italy), Oct.14–17, 2004.

[2] O. Somekh, O. Simeone, Y. Bar-Ness, A. M. Haimovich, U. Spagnolini, and S. Shamai (Shitz), *Distributed Antenna Systems: Open Architecture for Future Wireless Communications*, ch. An Information Theoretic View of Distributed Antenna Processing in Cellular Systems. Auerbach Publications, CRC Press, May 2007.

[3] G. Foschini, H. C. Huang, K. Karakayali, R. A. Valenzuela, and S. Venkatesan, "The value of coherent base station coordination," in *Proceedings of the 2005 Conference on Information Sciences and Systems (CISS'05)*, (John Hopkins University, Baltimore, ML), Mar. 16 – 18, 2005.

[4] H. Zhang, H. Dai, and Q. Zhou, "Base station cooperation for multiuser MIMO: Joint transmission and BS selection," in *Proceedings of the 2004 Conference on Information Sciences and Systems (CISS'04)*, (Princeton University, Princeton, NJ), Mar. 17 – 19, 2004.

[5] A. Ekbal and J. M. Cioffi, "Distributed transmit beamforming in cellular networks," in *Proceedings of the ICC 2005 Wireless Communications Theory (ICC'05)*, (Seoul, Korea), May 16–20, 2005.

[6] A. D. Wyner, "Shannon-theoretic approach to a Gaussian cellular multiple-access channel," *IEEE Transactions on Information Theory*, vol. 40, pp. 1713–1727, Nov. 1994.

[7] S. Shamai (Shitz) and B. M. Zaidel, "Enhancing the cellular downlink capacity via co-processing at the transmitting end," in *Proceedings of the IEEE 53rd Vehicular Technology Conference (VTC 2001 Spring)*, vol. 3, (Rhodes, Greece), pp. 1745–1749, May 6–9, 2001.

[8] O. Somekh, B. M. Zaidel, and S. Shamai (Shitz), "Sum-rate characterization of multi-cell processing," in *Proceedings of the Canadian workshop on information theory (CWIT'05)*, (McGill University, Montreal, Quibec, Canada), Jun. 5–8, 2005.

[9] H. Weingarten, Y. Steinberg, and S. Shamai (Shitz), "The capacity region of the Gaussian MIMO broadcast channel," in *Proceedings of the 2004 IEEE International Symposium on Information Theory (ISIT'04)*, (Chicago, USA), p. 174, Jun. 27 – Jul. 2, 2004.

[10] A. Wiesel, Y. C. Eldar, and S. Shamai (Shitz), "Linear precoding via conic optimization for fixed MIMO receivers," *IEEE Transactions on Signal Processing*, vol. 54, pp. 161–176, Jan. 2006.

[11] M. Stojnic, H. Vikalo, and B. Hassibi, "Maximizing the sum-rate of multi-antenna broadcast channels using linear preprocessing," *IEEE Transactions on Wireless Communications*, vol. 5, pp. 2338–2342, Sep. 2006.

[12] W. Yu and T. Lan, "Transmitter optimization for the multi-antenna downlink with per-antenna power constraint." to appear in IEEE Transactions on Signal Processing, 2007.

[13] G. Caire and S. Shamai (Shitz), "On the achievable throughput of a multi-antenna Gaussian broadcast channel," *IEEE Transactions on Information Theory*, vol. 49, no. 7, pp. 1691–1706, 2003.

[14] F. Boccardi and H. Huang, "Optimum power allocation for the MIMO-BC zero-forcing precoder with per-antenna power constraints," in *Proceedings of the 2006 Conference on Information Sciences and Systems (CISS'06)*, (Princeton, NJ), Mar. 22–24 2006.

[15] C. B. Chae, R. Heath, and D. Mazzarese, "Achievable sum-rate bounds of zero-forcing based linear multi-user MIMO systems," in *Proceedings of the 44th Annual Allerton Conference on Communication, Control and Computing*, (Allerton House, Monticello, Illinois), Sep. 27-29 2006.

[16] J. Jiang, R. Buehrer, and W. Tranter, "Greedy scheduling performance for a zero-forcing dirty-paper coded system," *IEEE Transaction on Communications*, vol. 54, pp. 789–793, May 2006.

[17] T. Yoo and A. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero forcisng beam forming," *IEEE Journal on Selected Areas in Communications (JSAC), Special Issue on 4G Wireless Systems*, vol. 24, pp. 528–541, Mar. 2006.

[18] A. Vakili, A. Dana, M. Sharif, and B. Hasibi, "Differentiated rate sceduling for MIMO broadcast channels," in *Proc. of the Allerton conference on communication, control, and computing (Allerton'06)*, (Monticello, IL, USA), Sep. 2006.

[19] A. Lozano, A. Tulino, and S. Verdú, "High-SNR power offset in multi-antenna communications," *IEEE Transactions on Information Theory*, vol. 51, pp. 4134–4151, Dec. 2005.

[20] O. Somekh, B. M. Zaidel, and S. Shamai (Shitz), "Sum-rate characterization of joint multiple cell-site processing." Submitted to the IEEE Transactions on Information Theory, 2005.

[21] S. Shamai (Shitz) and A. D. Wyner, "Information-theoretic considerations for symmetric, cellular, multiple-access fading channels - Part I," *IEEE Transactions on Information Theory*, vol. 43, pp. 1877–1894, Nov. 1997.

[22] S. Shamai (Shitz) and S. Verdú, "The impact of frequency-flat fading on the spectral efficiency of CDMA," *IEEE Transactions on Information Theory*, vol. 47, pp. 1302–1327, May 2001.

[23] M. Sharif and B. Hassibi, "On the capacity of MIMO broadcast channel with partial side information," *IEEE Transactions on Information Theory*, vol. 51, pp. 506–522, Feb. 2005.

[24] W. Choi and J. G. Andrews, "The capacity gain from base station cooperative scheduling in a mimo dpc cellular system," in *Proceedings of the IEEE International Symposium on Information Theory (ISIT'06)*, (Seattle, WA), Jul. 2006.

[25] A. Horn, "Doubly stochastic matrices and the diagonal of a rotation matrix," *American Journal of Mathematics*, vol. 76, pp. 620–630, 1954.

[26] R. M. Gray, "On the asymptotic eigenvalue distribution of Toeplitz matrices," *IEEE Transactions on Information Theory*, vol. IT-18, pp. 725–730, Nov. 1972.

[27] I. S. Gradshteyn and I. M. Ryzhik, *Table of Integrals, Series, and Products*. Academic Press, 6 ed., 2000.

[28] H. A. David, *Order Statistics*. J. Wiley & Sons, 1970.

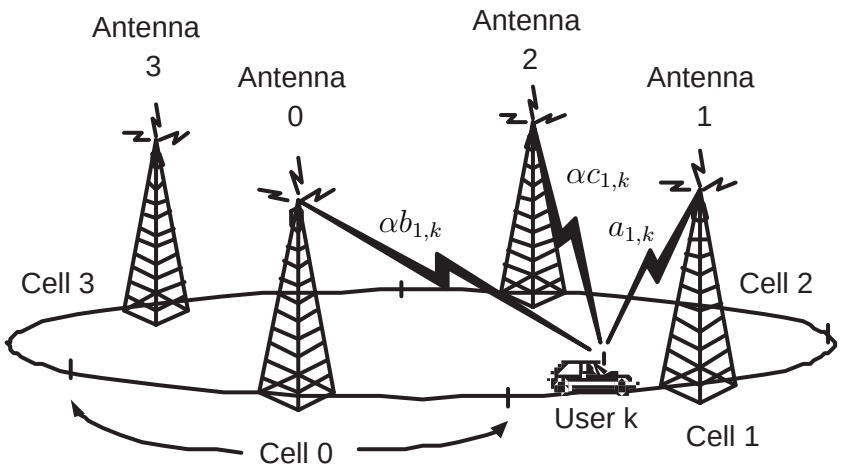

Fig. 1. Wyner's circular array system model $M = 4$.

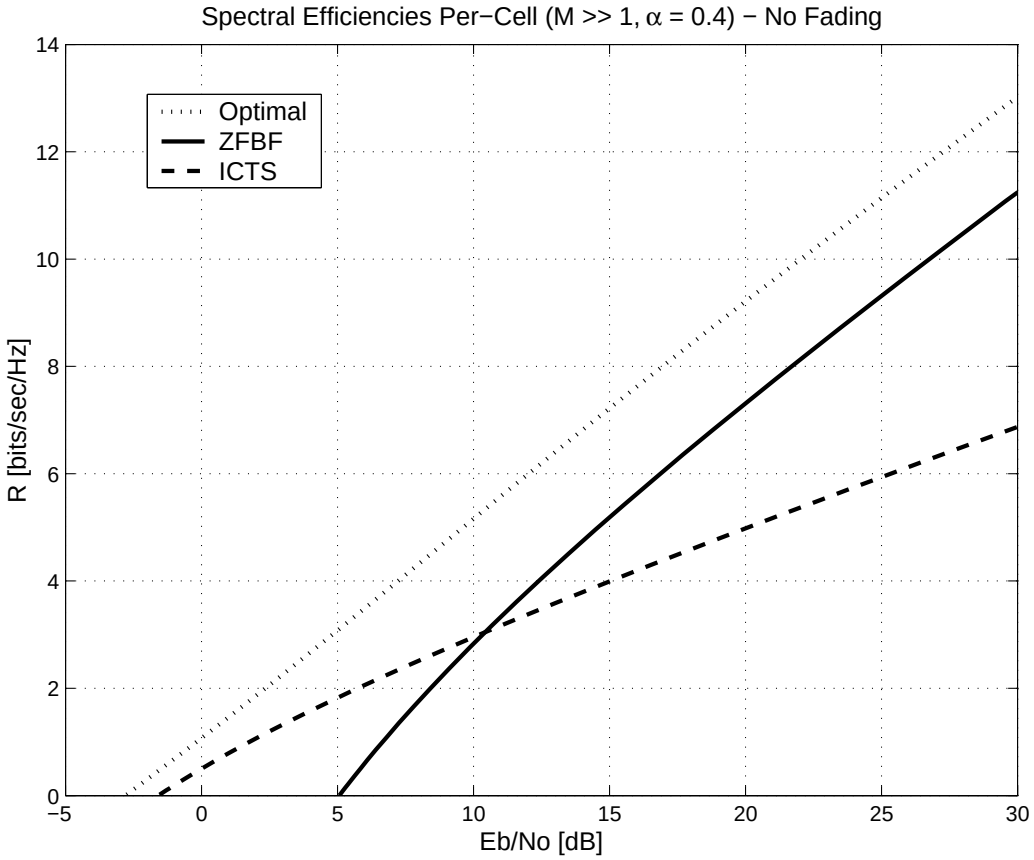

Fig. 2. Spectral efficiencies per-cell with no fading vs. E_b^t/N_0 for $\alpha = 0.4$.

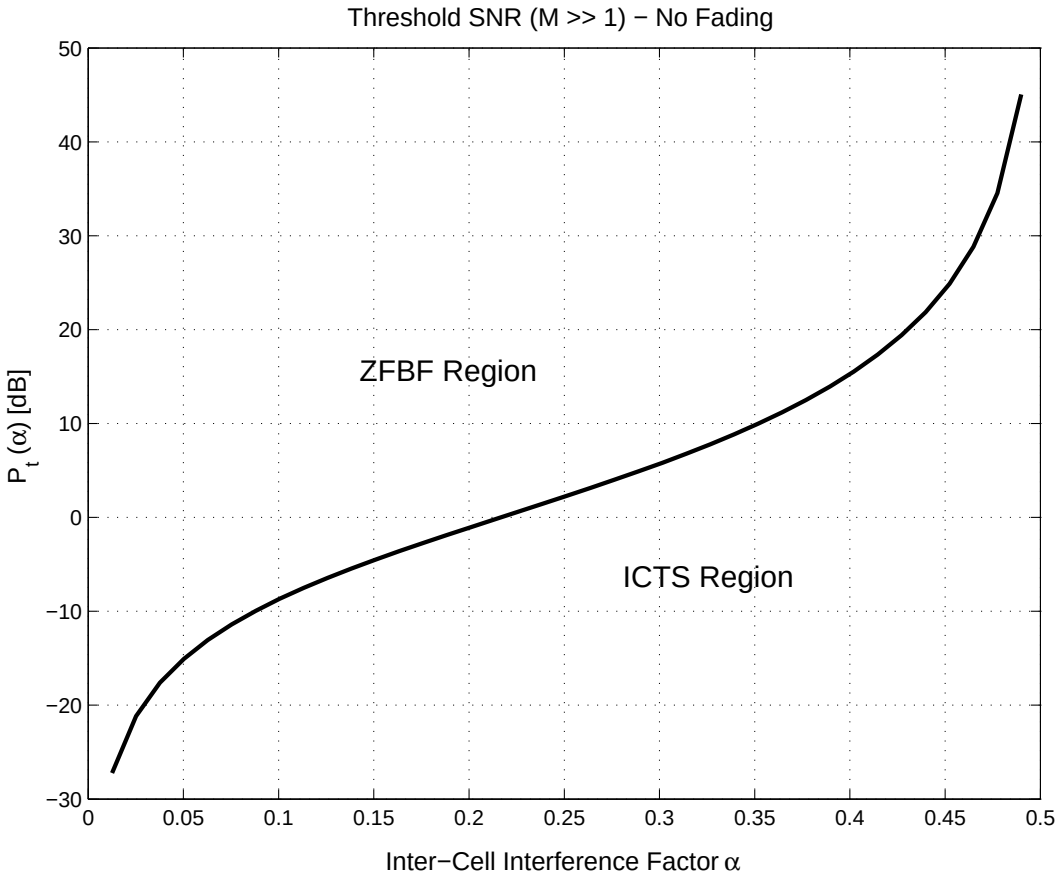

Fig. 3. The SNR threshold $P_t(\alpha)$ with no fading vs. the inter-cell interference factor α.

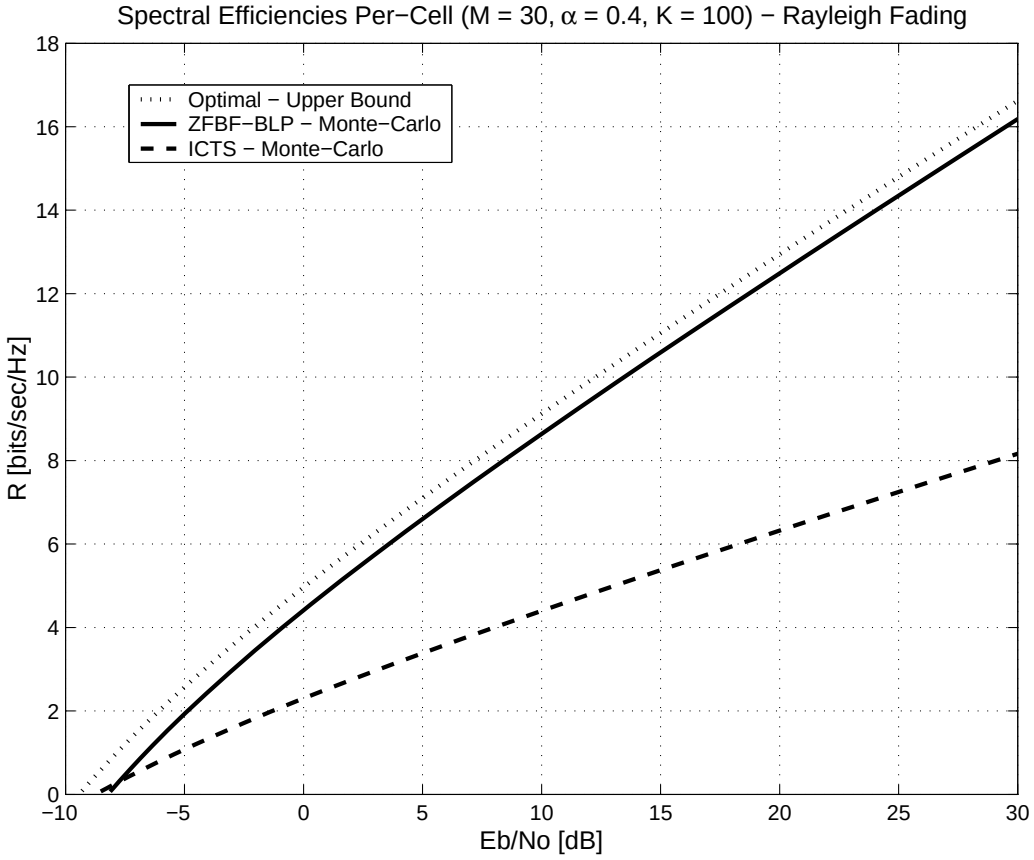

Fig. 4. Spectral efficiencies per-cell in the presence of Rayleigh fading vs. E_b^t/N_0 for $K = 100$, $\alpha = 0.4$ and $M = 30$.

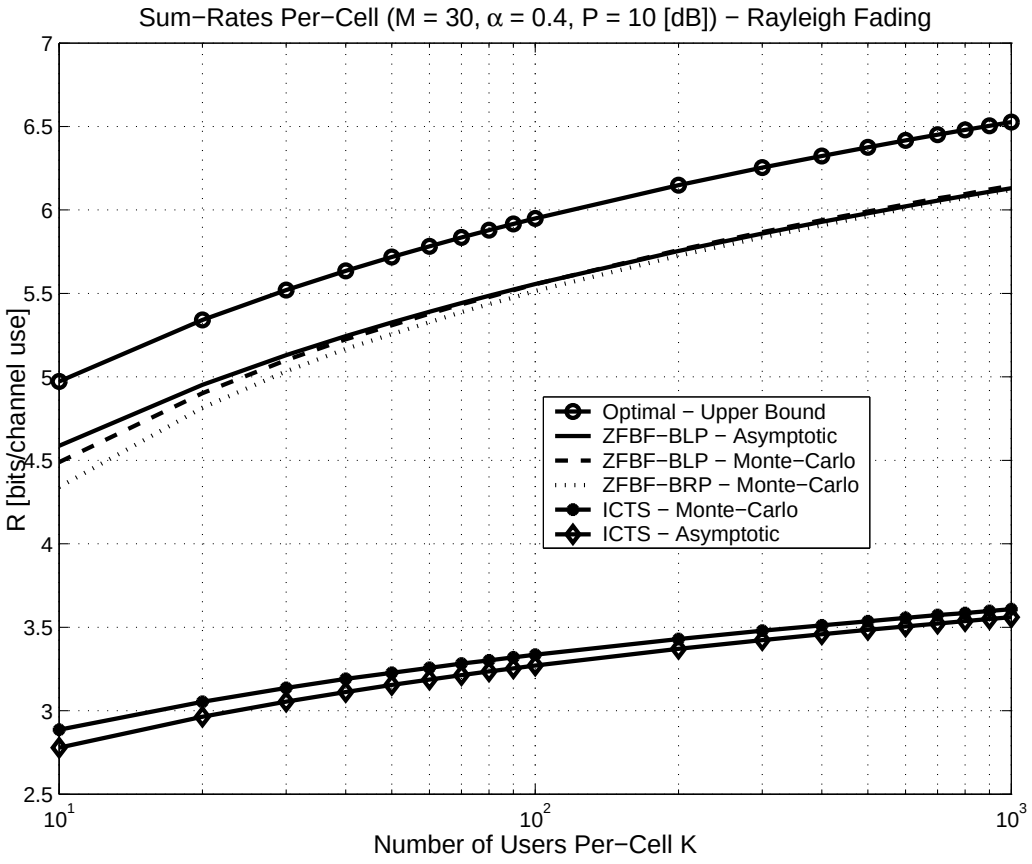

Fig. 5. Sum-rates per-cell in the presence of Rayleigh fading vs. the number of users per-cell K for $P = 10$ [dB], $\alpha = 0.4$, and $M = 30$.

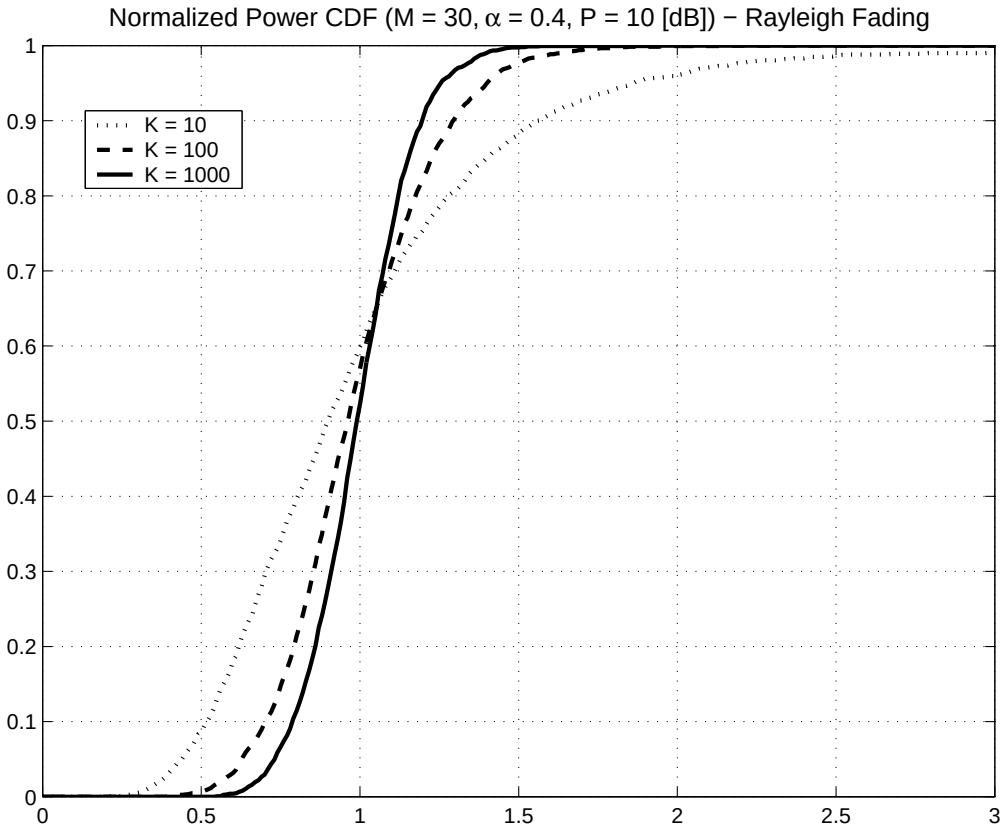

Fig. 6. Normalized transmitted power CDF of an arbitrary base-station in the presence of Rayleigh fading for $P = 10$ [dB], $\alpha = 0.4$, $M = 30$, and several values of K.