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

Oren Somekh, Osvaldo Simeone, Yeheskel Bar-Ness, and Alexander M. Haimovich
CWCSPR, ECE Departement, NJIT, Newark, NJ 07102, USA
Email: {somekh, simeone, barness, haimovich}@njit.edu

***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 the 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 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. It is shown that under an overall power constraint, a distributed multi-cell ZFBF to this sub-optimal subset of users achieves the same sum-rate growth rate as 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] and references therein for a short survey 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. [2][3][4]) 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 [5] 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.

This work was supported in part by NSF grant ANI-0338788

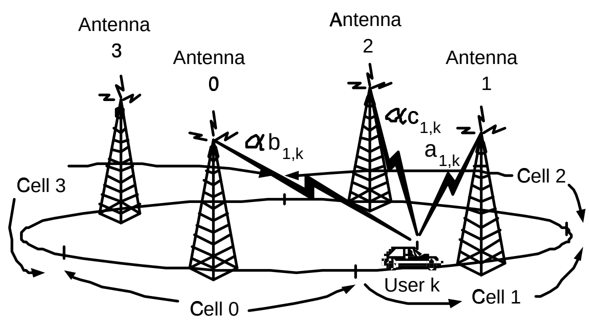

Fig. 1. Wyner's circular array system model.

The downlink channel of a similar model was first adopted in [6] were LQ factorization (forcing an arbitrary sub-optimal encoding order) combined with joint multi-cell dirty-paper coding (DPC) is deployed. The attainable 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) at the high SNR region. Recently, bounds to the sum-rate capacity supported by the downlink of this model has been reported in [7] under equal per-cell power constraints in the presence of Rayleigh flat fading. To achieve theses rates, DPC techniques are deployed [8]. 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. Therefore, a search for sub-optimal broadcast schemes is the focal point of many works. It is evident that for multi-cell processing when more users are involved this problem aggravates. Recently a zero-forcing beamforming (ZFBF) scheme has been considered in [9] for an M antennas MISO downlink setup under sum power constraint (see also [10]). In this sub-optimal scheme, the 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. However, determining the subset of users for transmission is a complex optimization problem.

In this paper, we consider 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 decreases with the inter-cell interference α. Introducing Rayleigh fading, the per-cell sum-rate of ZFBF is proved to experience the same growth rate as the optimal DPC scheme asymptotically with the number of users per-cell K under sum-power constraint. Finally, it is verified that the scheme satisfies in probability the more suitable *equal per-cell power constraints* asymptotically with increasing K.

II. System Model

Consider a circular variant of the infinite linear Wyner model [5] 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 independent 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. 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.

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

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 ergodic achievable sum-rate per-channel (or cell) is given by[3]

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

where the expectation is taken over the entries of $\boldsymbol{H}$.

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 $\boldsymbol{H}$,

$$P_m = [\boldsymbol{B}\boldsymbol{B}^\dagger]_{m,m} = \frac{MP\left[(\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right]_{m,m}}{\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right)} \ . \tag{6}$$

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

III. Background

For a similar model but with capacity achieving joint multi-cell DPC scheme, the downlink ergodic sum-rate capacity per-cell with a per-cell equal power constraint P, was recently proved in [7] for a non fading setup, to be

$$C_{\mathrm{opt-nf}} \underset{M\to\infty}{=} \int_0^1 \log\left(1 + P(1 + 2\alpha\cos(2\pi\theta))^2\right) d\theta \ , \tag{7}$$

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

$$\log\left(1 + \frac{(1+2\alpha^2)}{3} P\left((1-\epsilon)\log K + 3\right)\right) \leq C_{\mathrm{opt}} \leq \log\left(1 + (1+2\alpha^2)P\log K\right) \ , \tag{8}$$

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

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

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

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 [11] 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 ergodic sum-rate per-cell for a non-fading setup is given by

$$R_{\mathrm{icts-nf}} = \frac{1}{2}\log(1+2P) \ . \qquad (9)$$

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

$$R_{\mathrm{icts}} \cong \frac{1}{2}\log(1+2P\log K) \ . \qquad (10)$$

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

IV. Sum-Rate Analysis

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

$$R_{\mathrm{zfbf-nf}} \underset{M\to\infty}{=} \log\left(1+\mathcal{F}(\alpha)\ P\right) \qquad (11)$$

where

$$\mathcal{F}(\alpha) \triangleq \frac{1}{\int_0^1 \left(1+2\alpha\cos(2\pi\theta)\right)^{-2} d\theta} \ . \qquad (12)$$

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

Proof: See Appendix A. ■

It is easily verified that $0 < \mathcal{F}(\alpha) \leq 1$ and that it is a decreasing function of the interference factor α. Comparing (9) to (11), it is clear that the ZFBF scheme is superior to the ICTS scheme when the SNR P is above a certain threshold

$$P_{\mathrm{t}}(\alpha) = \frac{2(1-\mathcal{F}(\alpha))}{(\mathcal{F}(\alpha))^2} \qquad (13)$$

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. The case of finite and infinite M for $\alpha > 0.5$ is treated in [12].

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 channel for transmission in each cell. In other words, the selected user in the m'th cell is

$$\tilde{k}(m) = \underset{k}{\mathrm{argmax}}\{|a_{m,k}|^2\} \ , \qquad (14)$$

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 scaling law of the scheme's per-cell sum-rate R_{zfbf} (expression (5)), which is stated in the following proposition.

Proposition 2 *The scaling law of R_{zfbf} is asymptotically optimal with increasing number of users per-cell. Hence,*

$$\frac{R_{\mathrm{zfbf}}}{C_{\mathrm{opt}}} \underset{K\to\infty}{\longrightarrow} 1 \ . \qquad (15)$$

Proof: See Appendix B. ■

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 by

$$R_{\mathrm{zfbf}} \cong \log(1+P\log K) \ . \qquad (16)$$

It is also concluded that the ZFBF scheme provides a two fold scaling law than that of the ICTS scheme (10), 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 3 *The considered ZFBF 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.*

Proof: See Appendix C. ■

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.

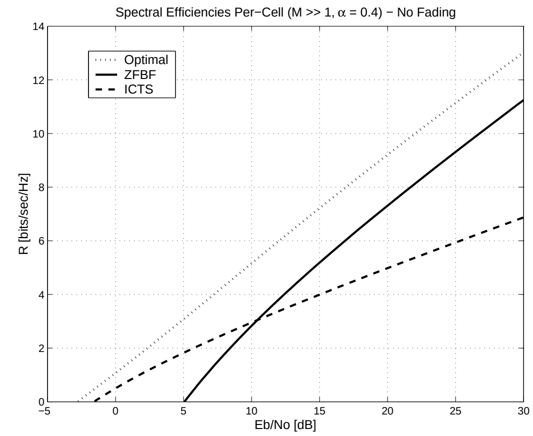

Fig. 2. Spectral efficiencies per-cell with no fading vs. E_b^t/N_0 for $\alpha = 0.4$.

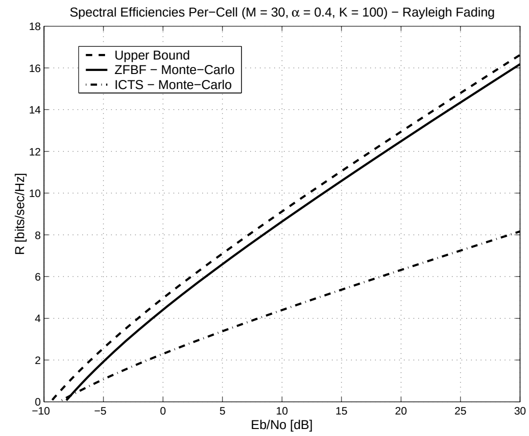

Fig. 4. Spectral efficiencies per-cell in the presence of Rayleigh fading vs. E_b^t/N_0 for $\alpha = 0.4$ and $M = 30$.

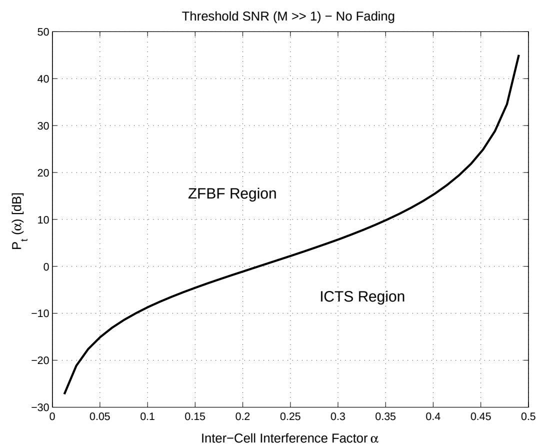

Fig. 3. The SNR threshold $P_t(\alpha)$ with no fading vs. the inter-cell interference factor α.

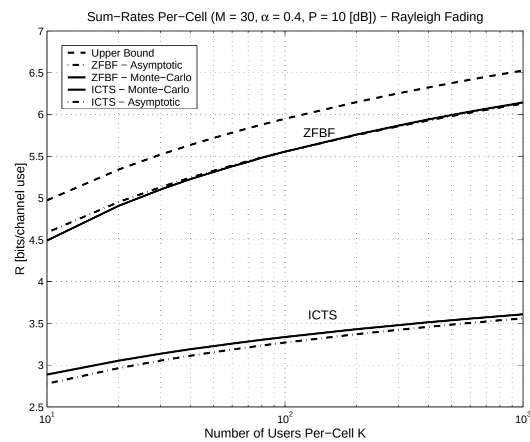

Fig. 5. Sum-rates per-cell in the presence of Rayleigh fading vs. the number of users per-cell K for $P = 10$ [dB], $\alpha = 0.4$ and $M = 30$.

V. Numerical Results

At first, the non-fading setup is considered under the assumption that the number of cells is large $M \gg 1$. In Fig. 2 the spectral efficiencies[4] per-cell of the optimal, ICTS, and ZFBF schemes (expressions (7), (9), and (11) 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 power threshold. The threshold $P_t(\alpha)$ (13) 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 (expression (5)) schemes, and the asymptotic upper bound of the optimal scheme (expression (8)) 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 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 curve and the sum-rate capacity upper bound is clearly explained by the fact that the ZFBF scheme does not use the 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 pre-log term of the ICTS sum-rate expression (10).

In Fig. 5 the sum-rates per-cell of the ICTS and ZFBF schemes (Monte-Carlo simulations, and asymptotic expressions (16) and (10)), and the upper bound of the optimal scheme, are plotted as a function of the number of users per-

[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 [7].

cell for $P = 10$ [dB], $\alpha = 0.4$ and $M = 30$. Examining the curves, the observations made for Fig. 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 Monte-Carlo simulation results to its asymptotic curve is observed as well, already for a modest number of users per cell.

Additional numerical results (not presented here due to space limitations, see [12]) show that for a fixed K the gap between the per-cell sum-rate curves of the optimal and ZFBF schemes increases with α. Moreover, by increasing the number of users per-cell, the standard deviation of an arbitrary cell transmit power decreases, as anticipated by Proposition 3; the normalized standard deviation reduces by 6 [dB] while the number of users increases from $K = 10$ to $K = 1000$.

VI. Concluding Remarks

In this work a ZFBF scheme for the downlink of the circular Wyner model is considered in the absence and presence of Rayleigh fading. For the no fading setup, a closed form expression for the per-cell sum-rate and round-robin scheduling (under both, overall and equal per-cell power constraints) demonstrates superior performance over the ICTS scheme when the SNR crosses a certain threshold which is an increasing function of α. Introducing Rayleigh fading, and utilizing a simple scheduling based on “best” local channel user selection, the per-cell sum-rate of the scheme demonstrates the same growth rate of $\log\log K$ as the optimal DPC scheme, asymptotically with increasing number of users per-cell K, while satisfying (in probability) equal per-cell power constraints. Furthermore, numerical results derived by Monte-Carlo simulations 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 necessary large).

Finally, it is noted that the results presented here, can be expanded to include other setups such as planner cellular models, and MIMO downlink channels. Other issues such as extreme SNR analysis, and detailed proofs are reported in [12].

Appendix

A. Proof of Proposition 1

The sum-rate per-cell of the non-fading setup is immediately derived from (5) by omitting the expectation since H is deterministic. Hence,

$$R_{\mathrm{zfbf-nf}} = \log\left(1 + \frac{MP}{\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right)}\right) \ . \tag{17}$$

To evaluate the inner log term of (17), the following set of equalities are useful

$$\begin{aligned}\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right) &= \sum_{m=0}^{M-1}\frac{1}{\lambda_m\left(\boldsymbol{H}\boldsymbol{H}^\dagger\right)} = \sum_{m=0}^{M-1}\frac{1}{\lambda_m^2(\boldsymbol{H})} \\ &= \sum_{m=0}^{M-1}\left(\frac{1}{1+2\alpha\cos\left(\frac{2\pi m}{M}\right)}\right)^2 \ ,\end{aligned} \tag{18}$$

where $\lambda_m(\cdot)$ is the m'th eigenvalue of an arbitrary matrix, and the last equality is derived following [13]. Substituting (18) into (17), and taking M to infinity yields (11).

Since, for the non-fading setup, H is a circulant matrix, then according to [13], $(\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}$ 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

$$P_m = [\boldsymbol{B}\boldsymbol{B}^\dagger]_{m,m} = \frac{MP\left[(\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right]_{m,m}}{\mathrm{tr}\left((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right)} = P \ , \tag{19}$$

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.

B. Proof of Proposition 2

Examining (8) it is evident that the optimal sum-rate scales like $\log\log K$ asymptotically with K. Hence, it is sufficient to show that

$$\frac{R_{\mathrm{zfbf}}}{\log\log K} \underset{K\to\infty}{\longrightarrow} 1 \ . \tag{20}$$

Next, the channel transfer matrix $\boldsymbol{H}$, resulting from the “best” local channel selection procedure, is shown to satisfy the following Propositions (Proposition 4 is not proved here due to space limitations and can be found in [12]).

Proposition 4 *The Frobenius norm of the matrix $(\boldsymbol{H}\boldsymbol{H}^\dagger/\log K - \boldsymbol{I}_M)$ converges in probability to 0. Hence,*

$$\left\|\boldsymbol{H}\boldsymbol{H}^\dagger/\log K - \boldsymbol{I}_M\right\|_{\mathrm{F}} \underset{K\to\infty}{\overset{p}{\longrightarrow}} 0 \ , \tag{21}$$

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

Proposition 5 *The eigenvalues of the matrix $(\boldsymbol{H}\boldsymbol{H}^\dagger/\log K)$ converge in probability to 1. Hence,*

$$\lambda_m(\boldsymbol{H}\boldsymbol{H}^\dagger)/\log K \underset{K\to\infty}{\overset{p}{\longrightarrow}} 1 \ , \forall m \ . \tag{22}$$

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

$$\|\boldsymbol{A}\|_{\mathrm{F}} = \sqrt{\frac{1}{M}\,\mathrm{tr}\left(\boldsymbol{A}^\dagger\boldsymbol{A}\right)} = \sqrt{\frac{1}{M}\sum_{m=0}^{M-1}\lambda_m^2(\boldsymbol{A})} \ , \tag{23}$$

then according to Proposition 4, for any finite M we get

$$\begin{aligned}\left\|\frac{\boldsymbol{H}\boldsymbol{H}^\dagger}{\log K} - \boldsymbol{I}_M\right\|_{\mathrm{F}} &= \sqrt{\frac{1}{M}\sum_{m=0}^{M-1}\lambda_m^2\left(\frac{\boldsymbol{H}\boldsymbol{H}^\dagger}{\log K} - \boldsymbol{I}_M\right)} \\ &= \sqrt{\frac{1}{M}\sum_{m=0}^{M-1}\left(\frac{\lambda_m(\boldsymbol{H}\boldsymbol{H}^\dagger)}{\log K} - 1\right)^2} \underset{K\to\infty}{\overset{p}{\longrightarrow}} 0 \ ,\end{aligned} \tag{24}$$

and the proof is completed by noting that the last equality of (24) holds if and only if (22) holds. ■

Now, let us define the following event

$$\mathcal{A} = \{w : |\lambda_m/\log K - 1| < \epsilon \ , \forall m\} \ , \tag{25}$$

where λ_m is the m'th eigenvalue of $(\boldsymbol{H}\boldsymbol{H}^\dagger)$. Next, we rewrite the LHS of (20) as

$$\begin{aligned}\frac{R_{\text{zfbf}}}{\log\log K} &= E\left\{\log\left(1+\frac{MP}{\sum_m \frac{1}{\lambda_m}}\right)\right\}/\log\log K\\ &\geq E\left\{\mathbf{1}_{\mathcal{A}}\log\left(1+\frac{MP}{\sum_m \frac{1}{\lambda_m}}\right)\right\}/\log\log K\\ &\overset{(a)}{\geq} Pr(\mathcal{A})\frac{\log\left(1+(1-\epsilon)P\log K\right)}{\log\log K}\\ &\overset{(b)}{\geq}\left(\sum_m Pr\left(\left|\frac{\lambda_m}{\log K}-1\right|<\epsilon\right)-(M-1)\right)\\ &\qquad\frac{\log\left(1+(1-\epsilon)P\log K\right)}{\log\log K}\underset{K\to\infty}{\longrightarrow} 1\ ,\end{aligned} \tag{26}$$

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

$$Pr\left(\bigcap_{n=1}^{N}\mathcal{B}_n\right)\geq\sum_{n=1}^{N}Pr(\mathcal{B}_n)-(N-1)\ , \tag{27}$$

where $\{\mathcal{B}_n\}_{n=1}^N$ is a set of arbitrary events. In addition, the final limit of (26) is achieved by invoking Proposition 5 and taking K to infinity.

C. Proof of Proposition 3

The average transmit power of the m'th cell site antenna conditioned on the channel transfer matrix $\boldsymbol{H}$ (expression (6)) satisfies the following set of inequalities

$$\begin{aligned}P_m &= \frac{MP\left[(\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right]_{m,m}}{\operatorname{tr}\left((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right)}\\ &\leq P\left(\max_m\left[(\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1}\right]_{m,m}\right)\left(\max_m\lambda_m(\boldsymbol{H}\boldsymbol{H}^\dagger)\right)\\ &\overset{(a)}{\leq} P\left(\max_m\lambda_m((\boldsymbol{H}\boldsymbol{H}^\dagger)^{-1})\right)\left(\max_m\lambda_m(\boldsymbol{H}\boldsymbol{H}^\dagger)\right)\\ &= P\frac{\max_m\lambda_m(\boldsymbol{H}\boldsymbol{H}^\dagger)}{\min_m\lambda_m(\boldsymbol{H}\boldsymbol{H}^\dagger)}\ ,\end{aligned} \tag{28}$$

where (a) is achieved by recalling that the eigenvalues of an Hermitian matrix majorize its diagonal entries (Horn's Theorem [14]).

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

$$Pr\left(P_m\leq P+\epsilon\right)\underset{K\to\infty}{\longrightarrow}1\ , \tag{29}$$

for any arbitrarily small $\epsilon>0$.

Now, Let us define $\bar{\lambda}\triangleq\max_m\lambda_m(\boldsymbol{H}\boldsymbol{H}^\dagger)/\log K$, $\underline{\lambda}\triangleq\min_m\lambda_m(\boldsymbol{H}\boldsymbol{H}^\dagger)/\log K$, and rewrite the LHS of (29) as follows

$$\begin{aligned}&Pr\left(P_m\leq P+\epsilon\right)\overset{(a)}{\geq}Pr\left(\frac{\bar{\lambda}}{\underline{\lambda}}\leq 1+\frac{\epsilon}{P}\right)\\ &\overset{(b)}{\geq}Pr\left(\frac{\bar{\lambda}}{\underline{\lambda}}\leq 1+\frac{\epsilon}{P}\bigcap\left|\bar{\lambda}-1\right|<\epsilon_1\bigcap\left|\underline{\lambda}-1\right|<\epsilon_1\right)\\ &\overset{(c)}{\geq}Pr\left(\frac{1+\epsilon_1}{1-\epsilon_1}\leq 1+\frac{\epsilon}{P}\bigcap\left|\bar{\lambda}-1\right|<\epsilon_1\bigcap\left|\underline{\lambda}-1\right|<\epsilon_1\right)\\ &\overset{(d)}{=}Pr\left(\left|\bar{\lambda}-1\right|<\epsilon_1\bigcap\left|\underline{\lambda}-1\right|<\epsilon_1\right)\\ &\overset{(e)}{\geq}Pr\left(\left|\bar{\lambda}-1\right|<\epsilon_1\right)+Pr\left(\left|\underline{\lambda}-1\right|<\epsilon_1\right)-1\underset{K\to\infty}{\longrightarrow}1\ ,\end{aligned} \tag{30}$$

where, (a) is achieved using (28), (b) is due to the fact that $Pr(A)\geq Pr(A\cap B)$ where $A,\ B$ are arbitrary events, (c) is achieved by increasing the eigenvalue ratio, (d) is achieved by setting $\epsilon_1<\frac{1}{2}\epsilon/(2P+\epsilon)$, hence, ensuring that the first event has probability 1, and (e) is achieved by invoking (27). Finally, the last limit is due to Proposition 5.

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] 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.

[3] 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.

[4] 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.

[5] 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.

[6] 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.

[7] 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.

[8] 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.

[9] T. Yoo and A. Goldsmith, "Optimality of zero-forcing beam forming with multiuser diversity," in *Proceedings of the ICC 2005 Wireless Communications Theory (ICC2005)*, (Seoul, Korea), May 16–20, 2005.

[10] 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.

[11] 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.

[12] O. Somekh, O. Simeone, Y. Bar-Ness, A. M. Haimovich, and S. Shamai (Shitz), "Distributed multi-cell zero-forcing beamforming in celullar downlink channels." in preparation.

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

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