# Interference Mitigation and Multiuser Multiplexing with Beam-Steering Antennas

Martin Kurras, Lars Thiele
Fraunhofer - Heinrich Hertz Institute
Einsteinufer 37, 10587 Berlin, Germany
Email: First.Last@hhi.fraunhofer.de

Giuseppe Caire
Technical University of Berlin
Einsteinufer 25, 10587 Berlin, Germany
Email: caire@tu-berlin.de

***Abstract*—We revisit a recently proposed framework for "massive MIMO" named Joint Space Division and Multiplexing (JSDM), that explicitly distinguishes between long-term channel statistics and instantaneous channel state information. In JSDM, the beamforming function is split into two components: *pre-beamforming* and *multiuser precoding*. The pre-beamforming stage depends only on the channel covariance information, which is (locally) time-invariant and can be acquired without significant training overhead by averaging over time. Based on this information, the pre-beamforming stage separates the users into groups without significant mutual interference, irrespectively of their instantaneous random channel realization. Then, inside each user group, a standard multiuser MIMO precoding scheme based on instantaneous channel state information can be used in order to achieve further multiuser multiplexing gain. JSDM is naturally suited to a Hybrid Digital Analog (HDA) implementation, where the pre-beamforming function can be accomplished through a reconfigurable RF front-end. In this paper, we consider a low-complexity HDA implementation of JSDM, where the canonical massive MIMO antenna array is replaced by a smaller number of programmable *beam-steering antennas*. We show by simulation that with this architecture it is possible to achieve both multiplexing gain and interference mitigation across different user groups.**

## I. Introduction

Massive MIMO relies on the principle that with TDD reciprocity it is possible to train an arbitrarily large number of antennas from a limited number of uplink mutually orthogonal pilot signals [1]. Therefore, the number of users $S$ to be served on each time-frequency slot is essentially fixed by the pilot overhead and complexity of the channel estimation and precoder computation scheme, while the number of antennas at the base station $M$ can be much larger, i.e., $M \gg S$. This has non-trivial advantages in terms of system simplicity (the high dimensional fading channel converges towards a deterministic behavior, such that multiuser scheduling and rate adaptation become trivial), power efficiency, and inter-cell interference mitigation, due to the very large beamforming gain [2].

A recently proposed approach to massive MIMO, nicknamed *Joint Space Division and Multiplexing* (JSDM), takes explicit advantage of the channel slowly-varying statistical information (second-order statistics) in order to decrease the computation burden and the training overhead [3]. JSDM is also is naturally suited to a Hybrid Digital Analog (HDA) implementation, where the component that depends on the statistical channel information can be implemented in the RF domain.

In this paper we validated that it is possible to achieve multiuser multiplexing gain and interference mitigation at the same time applying a simplified form of JSDM with HDA beamforming, where instead of using a very large array of isotropic radiating elements (as in conventional massive MIMO), we can use one steerable antenna per RF chain, with a number of RF chains equal to the maximum number of simultaneously served users $S$. In principles, if the $S$ directional antennas have a total effective aperture similar to the $M$ array elements, the same type of massive MIMO performance (i.e., an effective channel of rank $S$ to the desired users, with a required level of interference rejection to the undesired users) should be achievable. Nevertheless, the steerable antennas architecture puts some structural constraint on the form of the beamformign matrices, such that their optimization is not straightforward any longer.

Here, we obtain a first encouraging preliminary answer to the question of the effectiveness of this architecture by an accurate statistical channel model and extensive system simulations. A still open question for future work is how to design the constrained beamforming vectors for the general case of multiple groups of ousers per base station subject to interference mitigation with respect to the users of other base stations.

## II. Joint Spatial Division and Multiplexing

Consider the downlink (DL) of a single-cell system, with a base station equipped with $M$ antennas and $K$ single-antenna users. One channel use of the canonical discrete-time complex baseband model is given by

$$y_k = \sqrt{\beta_k}\mathbf{h}_k^{\mathsf{H}}\mathbf{x} + z_k, \quad k = 1, \dots, K, \tag{1}$$

where $\mathbf{h}_k \in \mathbb{C}^M$ is the vector of channel coefficients between the base station antennas and user $k$ single antenna, $\mathbf{x} \in \mathbb{C}^M$ is the transmitted signal vector, and $z_k \sim \mathcal{CN}(0,1)$ is AWGN at user $k$'s receiver. Without loss of generality, we let the channel vector $\mathbf{h}_k$ with normalized energy, such that $\mathrm{tr}(\mathbb{E}[\mathbf{h}_k\mathbf{h}_k^{\mathsf{H}}]) = M$, and collect the effect of the propagation pathloss (distance, shadowing, absorption) into the coefficient $\beta_k$, common to all the antennas. The transmit signal vector satisfies the sum power constraint $\mathrm{tr}(\mathbb{E}[\mathbf{x}\mathbf{x}^{\mathsf{H}}]) \leq \mathsf{SNR}$. We assume the standard block-fading model, such that the channel vectors $\mathbf{h}_k$ for all $k = 1, \dots, K$ are random, but constant in the time-frequency

ISBN 978-3-8007-3662-1

domain for blocks of $T \approx W_c T_c/2$ channel uses, where $W_c$ is channel coherence bandwidth and $T_c$ is channel coherence time. For example, the LTE system has been designed such that the channel is constant over a transmission resource block formed by 12 OFDM subcarriers in frequency and 14 OFDM symbols in time, such that $T$ is at least equal to 168 channel uses.[1]

In general, the channel vectors $\mathbf{h}_k$ are realizations of vector random processes. For the time being, we assume that the second-order statistics of such processes are known, i.e., for each $k$ the base station knows its mean value $\boldsymbol{\mu}_k = \mathbb{E}[\mathbf{h}_k]$ and its covariance matrix $\mathbf{R}_k = \mathbb{E}[(\mathbf{h}_k - \boldsymbol{\mu}_k)(\mathbf{h}_k - \boldsymbol{\mu}_k)^{\mathsf{H}}]$. According to the well-known Karhunen-Loeve decomposition, it follows that $\mathbf{h}_k$ can be expressed in the form

$$\mathbf{h}_k = \mathbf{U}_k \Lambda_k^{1/2} \mathbf{w}_k + \boldsymbol{\mu}_k \tag{2}$$

where $\mathbf{R}_k = \mathbf{U}_k \boldsymbol{\Lambda}_k \mathbf{U}_k^{\mathsf{H}}$, with $\mathbf{U}_k \in \mathbb{C}^{M \times r_k}$ unitary, $\boldsymbol{\Lambda}_k$ diagonal $r_k \times r_k$ with positive diagonal elements, and where $\mathbf{w}_k \in \mathbb{C}^{r_k}$ is a zero-mean vector random process with uncorrelated unit-variance elements. In particular, when $\boldsymbol{\mu}_k$ is not zero, we have a LOS propagation component and $\boldsymbol{\mu}_k$ describes the array response vector for the phase shifts of the paths between the base stations antennas and user $k$ antenna. It is customary to define the Rician parameter $\mathcal{K}$ such that

$$\frac{1}{M}\text{tr}(\boldsymbol{\Lambda}_k) = \frac{1}{1+\mathcal{K}}, \quad \frac{1}{M}\|\boldsymbol{\mu}_k\|^2 = \frac{\mathcal{K}}{1+\mathcal{K}}.$$

Hence, the case of no LOS corresponds to $\mathcal{K} = 0$ (when $\mathbf{h}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{R}_k)$ this is the classical case of correlated Rayleigh fading) and the case of pure LOS corresponds to $\mathcal{K} \to \infty$.

Notice that according to (2), the channel vector $\mathbf{h}_k$ can be represented as the linear combination of the vectors $\boldsymbol{\mu}_k$ and $\mathbf{u}_k^{(1)}, \ldots, \mathbf{u}_k^{(r_k)}$, i.e., $\mathbf{h}_k \in \text{Span}(\boldsymbol{\mu}_k, \mathbf{u}_k^{(1)}, \ldots, \mathbf{u}_k^{(r_k)})$.

**Remark 1.** *In conventional beamforming based on LOS propagation, the channel vector $\mathbf{h}_k = \boldsymbol{\mu}_k$ spans a subspace of dimension 1. In the general case of a random channel vector in the form (2), the channel vector is contained in a subspace of dimension at most $r_k + 1 \ll M$. Hence, the knowledge of the channel second-order statistics $\{\boldsymbol{\mu}_k, \mathbf{R}_k\}$ generalizes the direction of arrival information intuitively connected with conventional beamforming.* ◊

With DL linear precoding, the base station transmits $\mathbf{x} = \mathbf{V}\mathbf{d}$, $\mathbf{d} \in \mathbb{C}^S$ is a vector of user data symbols, and $\mathbf{V} \in \mathbb{C}^{M \times S}$ is the precoding matrix. In the case of single-antenna users (considered here for simplicity of exposition), the number of DL data streams $S \le \min\{M, K\}$ coincides with the number of users scheduled to be served by spatial multiplexing in the same DL transmission slot.

### A. JSDM two-stage precoding

In general, $\mathbf{V}$ is a function of the Channel State Information at the Transmitter (CSIT). With TDD reciprocity, a (noisy) estimate $\widehat{\mathbf{H}}$ of the DL channel matrix $\mathbf{H}$ is obtained at the base station, and $\mathbf{V}$ is computed as a function of $\widehat{\mathbf{H}}$ [1], [2]. In contrast, in the JSDM approach [3], $\mathbf{V}$ is obtained as the concatenation of two linear transformations: a "pre-beamforming" matrix $\mathbf{B} \in \mathbb{C}^{M \times b}$ and a precoding matrix $\mathbf{P} \in \mathbb{C}^{b \times S}$, such that $\mathbf{V} = \mathbf{BP}$. The intermediate dimension $b$, which satisfies $S \le b \le M$, is a design parameter. The pre-beamforming matrix $\mathbf{B}$ depends only on the channel second-order statistics, i.e., it is a function of $\{(\boldsymbol{\mu}_k, \mathbf{R}_k) : k = 1, \ldots, K\}$. In contrast, $\mathbf{P}$ is a function of the *effective* channel matrix $\underline{\mathbf{H}} = \mathbf{B}^{\mathsf{H}}\mathbf{H}$ (in practice, a noisy estimate thereof must be obtained either by TDD reciprocity or by explicit DL training and UL feedback, for FDD systems).

**Remark 2.** *According to the well-known* Wide-Sense Stationary *(WSS) channel model, the second-order statistics $\{(\boldsymbol{\mu}_k, \mathbf{R}_k) : k = 1, \ldots, K\}$ are time-invariant. This holds true under the assumption that the users move for distances significantly smaller than the range between the users and the base station, such that the "scattering geometry" is unchanged while the phases with which the different multipath components superimpose change significantly and cause the small-scale fading effect. In other words, while the channel vectors are random, they are confined to vary in the small dimensional subspace determined by $\boldsymbol{\mu}_k$ and by the columns of $\mathbf{U}_k$, which remains invariant over time for intervals much larger than the fading coherence block $T$. In practice, this means that fast per-block training/feedback is not needed to learn the channel covariance information, which can be obtained on a much slower time-scale, by averaging over many fading blocks. Therefore, the training overhead cost of learning such long-term statistical information is negligible with respect to the instantaneous (per-block) CSIT training/feedback implemented in schemes such as [1], [4], [5].* ◊

### B. User grouping by subspace similarity

In the originally proposed JSDM idea [3], the $K$ users are partitioned into $G$ groups based on the similarity of their channel covariances. Denoting by $K_g$ and $S_g$ the number of users and the number of data streams in group $g$, we have $\sum_{g=1}^{G} K_g = K$ and $\sum_{g=1}^{G} S_g = S$, with the index $g_k$ used to denote the $k^{\text{th}}$ user in group $g$. Assuming zero-mean channels (Rayleigh fading) for simplicity of exposition and following (2), we let $\mathbf{h}_{g_k} = \mathbf{U}_g \boldsymbol{\Lambda}_g^{1/2} \mathbf{w}_{g_k}$, where $\mathbf{U}_g$ and $\boldsymbol{\Lambda}_g$ are group-specific (they are common to all users in group $g$), while $\mathbf{w}_{g_k}$ is independent across different users. Denoting by $\mathbf{H}_g = [\mathbf{h}_{g_1} \ldots \mathbf{h}_{K_g}]$ the aggregate channel matrix of users in group $g$, we have the overall $M \times K$ system channel matrix $\mathbf{H} = [\mathbf{H}_1 \ldots \mathbf{H}_G]$. We denote the pre-beamforming matrix for group $g$ as $\mathbf{B}_g$ of dimensions $M \times b_g$, such that $\sum_{g=1}^{G} b_g = b$, we have $\mathbf{B} = [\mathbf{B}_1, \ldots, \mathbf{B}_G]$. As a result, the collectively received signal by the users can be compactly written in vector form as:

$$\mathbf{y} = \underline{\mathbf{H}}^{\mathsf{H}}\mathbf{P}\mathbf{d} + \mathbf{z} \tag{3}$$

[1] Of course, this statement holds true in the range of user mobility for which the system has been designed.

where

$$\underline{\mathbf{H}}^{\mathsf{H}} = \begin{pmatrix} \mathbf{H}_1^{\mathsf{H}}\mathbf{B}_1 & \mathbf{H}_1^{\mathsf{H}}\mathbf{B}_2 & \dots & \mathbf{H}_1^{\mathsf{H}}\mathbf{B}_G \\ \mathbf{H}_2^{\mathsf{H}}\mathbf{B}_1 & \mathbf{H}_2^{\mathsf{H}}\mathbf{B}_2 & \dots & \mathbf{H}_2^{\mathsf{H}}\mathbf{B}_G \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{H}_G^{\mathsf{H}}\mathbf{B}_1 & \mathbf{H}_G^{\mathsf{H}}\mathbf{B}_2 & \dots & \mathbf{H}_G^{\mathsf{H}}\mathbf{B}_G \end{pmatrix} \tag{4}$$

and $\mathbf{H}_g^{\mathsf{H}}\mathbf{B}_{g'}$ denotes the effective channel matrix of coupling channel coefficients between the users of group $g$ and the pre-bamforming vectors of group $g'$ (columns of $\mathbf{B}_{g'}$).

A particularly interesting choice for its good complexity-performance tradeoff consists of the Per-Group Processing (PGP) approach, where $\mathbf{P}$ is constrained to have the block diagonal form $\mathbf{P} = \mathrm{diag}(\mathbf{P}_1, \dots, \mathbf{P}_G)$, with $\mathbf{P}_g$ of dimensions $b_g \times S_g$. This approach is attractive since it requires only the knowledge of the "direct" effective channels $\underline{\mathbf{H}}_g = \mathbf{B}_g^{\mathsf{H}}\mathbf{H}_g$ instead of the whole matrix $\underline{\mathbf{H}}$. Focusing only on the received signal for users in group $g$, we have

$$\mathbf{y}_g = \underline{\mathbf{H}}_g^{\mathsf{H}}\mathbf{P}_g\mathbf{d}_g + \left( \sum_{g'=1, g' \neq g}^{G} \mathbf{H}_g^{\mathsf{H}}\mathbf{B}_{g'}\mathbf{P}_{g'}\mathbf{d}_{g'} \right) + \mathbf{z}_g \tag{5}$$

where the bracketed term denotes the *inter-group interference*. When $\underline{\mathbf{U}} = [\mathbf{U}_1, \cdots, \mathbf{U}_G]$ is tall unitary, then the inter-group interference can be completely eliminated by letting $\mathbf{B}_g = \mathbf{U}_g$. Otherwise, inter-group interference can be eliminated by using block-diagonalization, such that $\mathbf{B}_g^{\mathsf{H}}\mathbf{U}_{g'} = \mathbf{0}$ for all $g' \neq g$ (see [3] for details).

**Remark 3.** *Intuitively, the PGP approach for JSDM consists of forming "fat beams", based only on the channel covariance information which, as discussed before, is stable in time and should be easy to estimate. The pre-beamforming projection creates smaller dimensional and either exactly or approximately decoupled multiuser MIMO channels, one per "fat beam". Inside each of such multiuser MIMO channels, it is still possible to achieve some multiplexing gain, depending on the rank of the resulting effective channel matrix* $\underline{\mathbf{H}}_g$. *This can be obtained by using standard instantaneous CSIT and multiuser MIMO precoding, independently and simultaneously, for each "fat beam". A pictorial representation of this idea is given in Fig. 1.* ◊

In the special case of uniform linear arrays (ULAs), the channel covariances are Toeplitz. Owing to Szego's asymptotic theory [6], the eigenvectors $\mathbf{U}_g$ of $\mathbf{R}_g$ are well-approximated by the columns of an $M \times M$ unitary DFT matrix whose indices corresponds to the support of the Fourier transform of the channel correlation function. Also, [3] shows that two groups with disjoint Angle of Departure (AoD) support have disjoint supports of their correlation Fourier transform, such that in the limit of very large $M$ (massive MIMO) the tall unitary condition holds when the groups are chosen with disjoint AoD support. Also, in this case the (optimal) eigen-beamforming matrices $\mathbf{U}_g$ reduce to simple "DFT beamforming", that can be implemented in the RF analog domain by simple phase shifts and therefore are suited to hybrid beamforming implementations. For more general AoA distributions and/or array configurations, [7] present algorithms for user grouping in order to efficiently partition the user population into groups for which JSDM with PGP precoding can be successfully applied. These algorithm can be seen as a quantization of the Grassmannian manyfold, in order to group users according to the similarity of their channel subspaces.

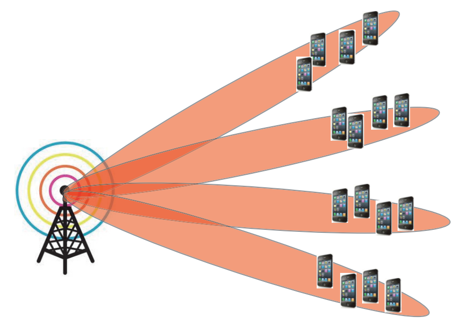

Fig. 1: A pictorial representation of JSDM with per-group processing, where the separation between the user groups is obtained through the pre-beamformign matrix $\mathbf{B}$, and the multiplexing gain inside each group is obtained through a block-diagonal precoding matrix $\mathbf{P}$.

### C. Hybrid digital-analog implementation

The two-stage structure of the JSDM precoder is suitable to HDA beamforming implementation, where the pre-beamforming matrix $\mathbf{B}$ is implemented in the analog RF domain (see Fig. 2). In this way, any standard training/feedback scheme to gather instantaneous CSIT can be applied on a per-group basis, to estimate directly the projected effective channel $\underline{\mathbf{H}}$ in order to determine the baseband precoding matrix $\mathbf{P}$. In short, the effect of $\mathbf{B}$ (RF front-end) is seen as part of the channel by the baseband training/feedback and estimation process. The HDA approach has also the non-trivial advantage of limiting the number of RF chains (especially A/D and D/A conversion) to $b \ll M$, while still allowing for a very large number $M$ of antennas.

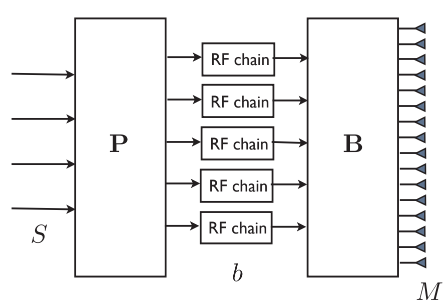


Fig. 2: The concept of a hybrid digital analog implementation of the JSDM two-stage precoding.

While a fully reconfigurable analog RF front-end with $b$ inputs and $M$ outputs is definitely desirable, this is not easy to be implemented in practice, especially within the form factor and cost constraints of small cell base stations. On the other hand, beam-steering directional antennas [8] are easily implemented and used in today's technology (e.g., see [9]).

Here we are not concerned on the particular implementation of reconfigurable beam-steering antennas. In contrast, we consider the fundamental structural constraint that such antennas impose on the general JSDM approach. From a conceptual viewpoint, a beam-steering antenna can be seen as an antenna array driven by a single RF input. For the same number of total array elements $M$, the beam-steering architecture consists of partitioning the array into groups of $m = M/S$ elements, and using such elements to create $S$ beams, as shown in Fig. 3.

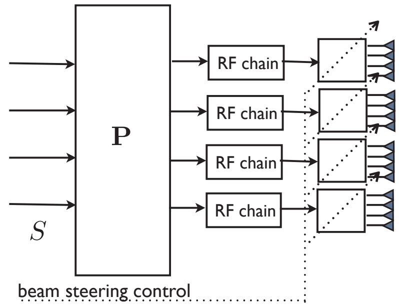


Fig. 3: The concept of a hybrid digital analog implementation of the JSDM two-stage precoding.

Mathematically, this means that the $M \times S$ matrix $\mathbf{B}$ is constrained to take on a block-diagonal form with $S$ diagonal blocks of dimension $m \times 1$, each (column) block being the steering vector of the corresponding beam-steering antenna. At this point, a number of possibilities exist to associate users to beams. In fact, we need not associate a single data stream to a specific beam. This corresponds to restricting the baseband precoding matrix $\mathbf{P}$ to be a permutation matrix. More generally, we can send linear combinations of data streams to each beam. Consider for example the case where we wish to serve $S$ users belonging to the same group $g$, with channel eigenspace $\text{Span}(\mathbf{U}_g)$, of rank $r_g \geq S$, and at the same time create small interference to users in group $g'$, with channel eigenspace $\text{Span}(\mathbf{U}_{g'})$. This is possible if we can find an $M \times S$ block-diagonal matrix $\mathbf{B}_g$ in the form said above such that $\mathbf{B}_g^{\mathsf{H}}\mathbf{U}_g$ has rank $S$ and $\mathbf{B}_g^{\mathsf{H}}\mathbf{U}_{g'} \approx \mathbf{0}$. However, while for unconstrained pre-beamformign matrix this is easily achieved by block-diagonalization (see above comments and details in [3]), in the case of the low-complexity HDA implementation with beam-steering antennas, imposing a block-diagonal constraint on $\mathbf{B}_g$, the optimization of the pre-beamforming matrix is far from obvious. Driven by the LOS intuition exposed in Remark 1, this should be possible if the matrix $\mathbf{B}_g$ defines beams that point in the direction of the desired users (group $g$) and put a null in the direction of the undesired users (group $g'$). However, whether this is possible and to what extent we can actually achieve multiuser multiplexing gain for a group of users, while mitigating interference to another group of users, in the presence of multipath fading which involves channel subspaces of dimensions larger than 1, is to be verified through an accurate statistical channel model and extensive system simulation. In the next sections, we shall describe such model and report some numerical experiments that indicate that it is indeed possible to achieve multiuser multiplexing and interference mitigation at the same time, by using multiple beam-steering antennas at the base station side.

## III. Simulation Model

### A. *Quasi Deterministic Radio Channel Generator*

The *Quasi Deterministic Radio Channel Generator* (QUADRIGA) is a stochastic channel model. It is an extension of the popular Wireless World Initiative for New Radio (WINNER) model [10] with several new features to achieve higher accuracy compared to measurements. The new model supports 3-D propagation, 3-D antenna patterns, time evolving channel traces of arbitrary length, scenario transitions and variable terminal speeds. Further details can be found in [11] and references therein. A reference implementation in MATLAB is available under Lesser General Public License (LGPL) [12].

To evaluate and verify the JSDM approach with HDA beamforming the following deployment is considered depicted in Fig. 4. We placed 2 base stations with orientation to each other as transmitter and 2 user groups of size 8. The focus of this paper is to demonstrate the potential of the scheme so the deployment is kept intentionally simple. We also force user group 1 and 2 to be served by transmitter 1 and 2 respectively. This leads to the following parameter per transmitter $G = 1$, $K_g = 8$. Note, that users within one group have the same “scattering geometry”.

The transmitter antenna is a ULA consisting of 4 blocks separated by $2\lambda$ where each block is formed by 4 antennas with $\frac{\lambda}{2}$ spacing given in Fig. 3. The single antenna of the users is modeled as isotropic radiator.

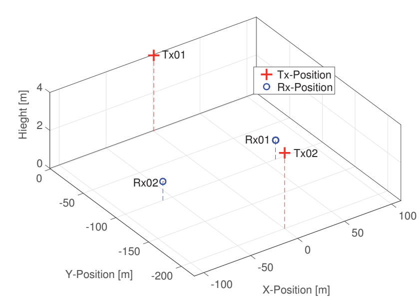


Fig. 4: Deployment of user groups and base stations.

### B. *User Grouping and Precoding*

To find a suitable set of users for multiplexing we apply semi-orthogonal user selection (SUS) [13] per group with projection based rate approximation according to [14]. The objective is to maximize the sum-rate per group. From the number of users selected by the grouping algorithm $S_g$ is determined. The pre-beamforming vectors $\mathbf{b}_g$ are obtained from a normalized DFT matrix.

For the reasons mentioned above we focus on PGP where $\mathbf{P}$ has a block diagonal form $\mathbf{P} = \text{diag}(\mathbf{P}_1, \ldots, \mathbf{P}_G)$ with $\mathbf{P}_g$ of

ISBN 978-3-8007-3662-1  

dimensions $b_g \times S_g$. $\mathbf{P}_g$ is obtained by regularized zero-forcing as in [3], [15] based on the idea from [16].

## IV. Results

In order to utilize the multiplexing gain for a user group each of the antenna blocks generates the same "fat beam" with the same pre-beamforming vector $\mathbf{b}_1$ selected from a DFT matrix. The effective channel $\mathbf{H}_g^{\mathrm{H}}\mathbf{B}_{g'}$ of such a beam for a single antenna-block is shown in Fig. 5. We see that the power is steered to user group 1 and at the same time interference towards group 2 is mitigated by a null. In this case the design parameter $b_g$ is set equal to 4. The Rician parameter $\mathcal{K} = 3$ and propagation conditions are set according to the "urban micro" scenario from [10], where base stations antennas are considered below rooftop. The center frequency is set to 2.4 GHz for city deployed Wireless Local Area Network (WLAN) as a possible application. Due to the stochastic channel model approach numerical simulation results are obtained by Monte-Carlo simulations with 500 independent channel realizations (from one run to the next the scattering geometry is changed at random according to the underlying statistical model).

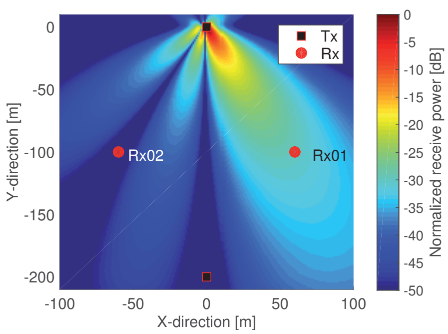


Fig. 5: Receive power of stream 1 in LOS scenario for antenna block 1 with pre-beamforming vector $\mathbf{b}_1$

Performance results from system simulations are shown in Fig. 6. For comparison we include the performance of a [1,4] ULA first with $\frac{\lambda}{2}$ and second with $2\lambda$ antenna spacing. To have the optimal performance in this scenario we consider the interference free single transmitter case with a single user group labeled as 1Tx. The case of 2 transmitter and 2 user groups as in Fig. 4 is labeled with 2Tx. Looking at the sold lines with one transmitter and user group we observe a beamforming gain from $\frac{\lambda}{2}$ to $2\lambda$ spacing and to the geometry described in Sec. III-A. Without pre-beamforming and interference mitigation the sum spectral efficiency of two transmitter is less than the single transmitter. In contrast, the sum spectral efficiency with the pre-beamforming in Fig. 5 is increased from 13 to 20 bit/s/Hz. Dividing the sum performance by the number of transmitters this results in 10 bit/s/Hz per base station which is only a loss of 3 bit/s/Hz compared to the optimal scenario of two transmitters serving simultaneously the two user groups without any mutual interference. This loss is due to the residual interference from NLOS components of the channel, which cannot be fully eliminated by the pre-beamforming projection due to its constrained block-diagonal form.

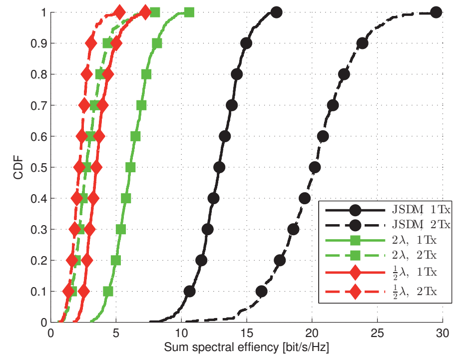


Fig. 6: Sum spectral efficiency of one and two transmitters for 4 respective 16 antennas using 4 RF-chains as in Fig. 3. In case of 4 antennas the spacing of the ULA is $\frac{\lambda}{2}$ or $2\lambda$ and for 16 antenna pre-beamforming from JSDM is used as shown in Fig. 5.

## References

[1] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, 2010.

[2] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprashad, "Achieving massive MIMO spectral efficiency with a not-so-large number of antennas," *IEEE Transactions on Wireless Communications*, vol. 11, no. 9, pp. 3226–3239, Sep. 2012.

[3] A. Adhikary, J. Nam, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: The large-scale array regime," *IEEE Transactions on Information Theory*, vol. 59, no. 10, pp. 6441–6463, 2013.

[4] R. Van Nee, "Breaking the gigabit-per-second barrier with 802.11 ac," *IEEE Wireless Communications*, vol. 18, no. 2, pp. 4–4, 2011.

[5] H. Huh, A. M. Tulino, and G. Caire, "Network MIMO With Linear Zero-Forcing Beamforming: Large System Analysis, Impact of Channel Estimation, and Reduced-Complexity Scheduling," *IEEE Transactions on Information Theory*, vol. 58, no. 5, pp. 2911 – 2934, 2012.

[6] U. Grenander and G. Szegő, *Toeplitz forms and their applications*. Chelsea Pub Co, 1984.

[7] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint Spatial Division and Multiplexing: Opportunistic Beamforming, User Grouping and Simplified Downlink Scheduling," *IEEE J. on Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 878 – 890, 2014.

[8] R. F. Harrington, "Reactively controlled directive arrays," *IEEE Transactions on Antennas and Propagation*, vol. 26, no. 3, pp. 390–395, 1978.

[9] [Online]. Available: http://www.ruckuswireless.com

[10] K. Pekka, M. Juha, H. Lassi, Z. Xiongwen, J. Tommi, S. Christian, N. Milan, M. Marko, H. Aihua, Y. Juha, H. Veli-Matti, A. Mikko, B. Robert, d. J. Yvo, and R. Terhi, "WINNER II Channel Models," EBITG, TUI, UOULU, CU/CRC, NOKIA, Tech. Rep. D112, 2007.

[11] S. Jaeckel, L. Raschkowski, K. Borner, and L. Thiele, "Quadriga: A 3-d multi-cell channel model with time evolution for enabling virtual field trials," *Antennas and Propagation, IEEE Transactions on*, vol. 62, no. 6, pp. 3242–3256, June 2014.

[12] [Online]. Available: http://www.quadriga-channel-model.de

[13] T. Yoo and A. Goldsmith, "On the Optimality of Multiantenna Broadcast Scheduling Using Zero-Forcing Beamforming," *Selected Areas in Communications, IEEE Journal on*, vol. 24, no. 3, pp. 528 – 541, Mar 2006.

[14] L. Thiele, M. Kurras, K. Börner, and T. Haustein, "User-Aided Sub-Clustering for CoMP Transmission: Feedback Overhead vs. Data Rate Trade-off," in *Signals, Systems and Computers (ASILOMAR), Conference Record of the Forty Sixth Asilomar Conference on*, Nov 2012.

[15] L. Thiele and M. Kurras, "Hierarchical precoding for ultra-dense heterogeneous networks," in *Signals, Systems and Computers, 2014 Asilomar Conference on*, Nov 2014.

[16] T. Haustein, C. von Helmolt, E. Jorswieck, V. Jungnickel, and V. Pohl, "Performance of mimo systems with channel inversion," in *Vehicular Technology Conference, 2002. VTC Spring 2002. IEEE 55th*, vol. 1, 2002, pp. 35–39 vol.1.

ISBN 978-3-8007-3662-1