

Chalmers Publication Library



CHALMERS

Copyright Notice

©2010 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

(Article begins on next page)

AN EFFICIENT SIGNALING FOR MULTI-MODE TRANSMISSION IN MULTI-USER MIMO

Nima Seifi, Tony Ottosson, Mats Viberg

Department of Signals and Systems
Chalmers University of Technology
Göteborg, Sweden
{seifin,tony.ottosson,viberg}@chalmers.se

Mikael Coldrey

Ericsson Research
Ericsson AB
Göteborg, Sweden
mikael.coldrey@ericsson.com

Andreas Wolfgang

Qamcom Technology AB
Göteborg, Sweden
andreas.wolfgang@qamcom.se

ABSTRACT

In this paper the downlink of a multi-user MIMO (MU-MIMO) system with multi-mode transmission is considered. We propose a low-complexity algorithm for selecting users and the corresponding number of data streams to each user, denoted as user transmission mode (UTM). The selection is only based on the average received signal-to-noise ratio (SNR) from the base station (BS) for each user. This reduces the overall amount of feedback for scheduling, as opposed to techniques that assume perfect instantaneous channel state information (CSI) from all users. Analytical average throughput approximations are derived for each user at different UTMs. Simulation results demonstrate that the proposed algorithm provides performance close to dirty paper coding (DPC) with considerably reduced feedback.

Index Terms— MIMO, multi-mode transmission.

1. INTRODUCTION

A downlink multi-user MIMO (MU-MIMO) system, where a base station (BS) equipped with multiple antennas simultaneously communicates with several multiple antenna users, is considered. It has been shown that the optimal transmission strategy for MU-MIMO is dirty paper coding (DPC) [1]. The DPC provides a useful performance bound, but it has a very high level of complexity to be implemented in practice. Although sub-optimal, practical linear precoding techniques such as zero forcing (ZF) and block diagonalization (BD) can perform very close to DPC in many scenarios [2].

Using linear precoding techniques, a downlink channel with M transmit antennas is decomposed into $L \leq M$ interference-free subchannels, allowing L active data streams in the system, denoted as system transmission mode (STM). In a system with a large number of users, the throughput can be optimized by selecting the optimum subset of users and the optimum number of data streams for transmission to each user, denoted as user transmission mode (UTM).

In this paper, we obtain analytical sum throughput approximation based on the knowledge of average received signal-to-noise ratio (SNR) from all users, and propose a greedy algorithm to select the subset of users and their corresponding UTMs that maximizes the sum throughput. Since only average SNR is used, the amount of feedback is considerably reduced. After scheduling, the selected users are informed about their UTMs and are asked to feed back the corresponding singular values and right singular vectors to the BS for precoding design. This further reduces feedback overhead at the cost of computing the singular value decomposition (SVD) of the channel at each receiver. Simulation results show that the sum throughput of the proposed algorithm achieves a significant fraction of the DPC sum throughput while reducing the feedback by about 90%.

2. SYSTEM MODEL

We consider the downlink of a MU-MIMO system with a BS equipped with M transmit antennas and K users each equipped with N receive antennas. A narrowband frequency flat-fading channel is considered. The received signal vector $\mathbf{y}_k \in \mathbb{C}^{N \times 1}$ at user k is given by

$$\mathbf{y}_k = \sqrt{\Gamma_k} \mathbf{H}_k \mathbf{x} + \mathbf{n}_k, \quad (1)$$

where $\mathbf{x} \in \mathbb{C}^{M \times 1}$ is the transmitted signal vector from the BS, $\mathbf{n}_k \sim \mathcal{CN}(0, \mathbf{I}_N)$ is the additive white Gaussian noise vector and $\mathbf{H}_k \in \mathbb{C}^{N \times M}$ is the channel matrix between user k and the BS. Further, Γ_k represents the distance-based pathloss including shadow fading for user k and is given by:

$$\Gamma_k = (d_k/d_0)^{-\gamma} \chi_k, \quad (2)$$

where d_k is the distance between the k th user and the BS, d_0 is the reference distance, χ_k represent the log-normal shadowing between the user k and the BS with standard deviation of $\sigma_\chi = 8$ dB and $\gamma = 3.5$ denotes the pathloss exponent, which are the typical values in an urban cellular environment [3]. For simplicity of analysis, we assume an average total power constraint (TPC) P for all the M transmit antennas at the BS,

This work has been supported in part by VINNOVA within the VINN Excellence Center Chase; and in part by SSF within the Strategic Research Center Charmant.

i.e. $\mathbb{E}\{\|\mathbf{x}\|^2\} \leq P$, where $\mathbb{E}[\cdot]$ denotes the expectation operator and $\|\cdot\|_2$ is the 2-norm of a vector.

3. MULTI-MODE TRANSMISSION STRATEGY

In a system with $KN > M$ the optimum subset of users and UTMs must be selected as mentioned in Sec.1. Let \tilde{K} and ℓ_k denote the total number of selected users and the selected UTM for user k , respectively. We assign indices $k = 1, \dots, \tilde{K}$ to the selected users. It is also assumed that the selected UTM for user k are indexed from 1 to ℓ_k . We denote the set consisting of selected users and their corresponding UTMs, when the STM is L , as

$$\mathcal{S}_L = \left\{ (k, \ell_k) : 1 \leq k \leq \tilde{K}, 1 \leq \ell_k \leq \min(M, N), \sum_{j=1}^{\tilde{K}} \ell_j = L \right\}. \quad (3)$$

The linearly precoded transmit signal can now be written as

$$\mathbf{x} = \sum_{k=1}^{\tilde{K}} \mathbf{M}_k \mathbf{s}_k, \quad (4)$$

where $\mathbf{M}_k \in \mathbb{C}^{M \times \ell_k}$ is the precoding matrix and $\mathbf{s}_k \in \mathbb{C}^{\ell_k \times 1}$ denotes the transmit vector symbol for user k . In order to be able to simultaneously transmit multiple spatially multiplexed streams to multiple users, we adopt a technique called multiuser eigen-mode transmission (MET) that was proposed in [4] and was shown to achieve a performance near the optimum capacity-achieving DPC. Under the linear precoding framework in MET, the channel of the k th user is decomposed using singular value decomposition (SVD) as $\mathbf{H}_k = \mathbf{U}_k \boldsymbol{\Sigma}_k \mathbf{V}_k^H$. It is assumed that the singular values in $\boldsymbol{\Sigma}_k$ are arranged so that the ℓ_k selected ones appear in the leftmost columns. We denote the i th singular value and the corresponding left and right singular vectors as $\sigma_{k,i}$, $\mathbf{u}_{k,i}$ and $\mathbf{v}_{k,i}$, respectively. Let $[\mathbf{U}_k]_{:(m:n)}$ denote the matrix obtained by choosing $n - m + 1$ columns from \mathbf{U}_k starting from the m th column. Using (1) and (4), the received signal \mathbf{y}_k in (1) is post-processed as $\mathbf{r}_k = [\mathbf{U}_k]_{:(1:\ell_k)}^H \mathbf{y}_k$ [4], resulting in the received output

$$\mathbf{r}_k = \sqrt{\Gamma_k} \mathbf{F}_k \mathbf{M}_k \mathbf{s}_k + \sqrt{\Gamma_k} \mathbf{F}_k \sum_{j=1, j \neq k}^{\tilde{K}} \mathbf{M}_j \mathbf{s}_j + \mathbf{w}_k, \quad (5)$$

where $\mathbf{F}_k = [\sigma_{k,1} \mathbf{v}_{k,1} \dots \sigma_{k,\ell_k} \mathbf{v}_{k,\ell_k}]^H$. Since $[\mathbf{U}_k]_{:(1:\ell_k)}^H$ is unitary, the processed noise $\mathbf{w}_k = [\mathbf{U}_k]_{:(1:\ell_k)}^H \mathbf{n}_k$ is also white. In the case of perfect knowledge of $\mathbf{F}_1, \dots, \mathbf{F}_{\tilde{K}}$, a zero-forcing criterion in [4] is used to design the precoders $\mathbf{M}_1, \dots, \mathbf{M}_{\tilde{K}}$ such that the inter-user interference is suppressed completely. In other words, the second term on the

right hand side of the equality in (5) will be zero. The average sum throughput for a given selection of users and UTMs \mathcal{S}_L can now be written as [5]

$$\bar{R}(\mathcal{S}_L) = \mathbb{E} \left[\max_{\mathbf{Q}_k} \sum_{k=1}^{\tilde{K}} \log \det (\mathbf{I}_{\ell_k} + \Gamma_k \mathbf{F}_k \mathbf{M}_k \mathbf{Q}_k \mathbf{M}_k^H \mathbf{F}_k^H) \right], \quad (6)$$

where

$$\sum_{k=1}^{\tilde{K}} \text{trace}(\mathbf{Q}_k) \leq P, \quad (7)$$

and the set of power allocation matrices $\{\mathbf{Q}_k\}_{k=1}^{\tilde{K}}$ are obtained through water-filling algorithm.

4. SUM THROUGHPUT ANALYSIS

In this section, we derive an analytical framework to calculate the average throughput for different number of users and UTMs. We assume equal power allocation among all the L active data streams for simplicity of the analysis. For a given selection of users and UTMs \mathcal{S}_L , the average throughput of a user k with assigned UTM ℓ_k , denoted as $I_k(\ell_k, L)$, can be written as [4]

$$\begin{aligned} I_k(\ell_k, L) &= \mathbb{E} \left[\log \det \left(\mathbf{I}_{\ell_k} + \frac{\Gamma_k P}{L} \mathbf{F}_k \mathbf{M}_k \mathbf{M}_k^H \mathbf{F}_k^H \right) \right] \\ &= \mathbb{E} \left[\sum_{i=1}^{\ell_k} \log \left(1 + \frac{\Gamma_k P}{L} \sigma_i^2(\mathbf{F}_k \mathbf{M}_k) \right) \right]. \end{aligned} \quad (8)$$

Let $\tilde{\ell}_k = \sum_{j=1, j \neq k}^{\tilde{K}} \ell_j$ denote the total number of interfering data streams for user k from all the other $(\tilde{K} - 1)$ scheduled users. In [6] it was shown that the precoding matrix \mathbf{M}_k can be written as a cascade of two precoding matrices as $\mathbf{M}_k = \mathbf{B}_k \mathbf{D}_k$, where the $M \times (M - \tilde{\ell}_k)$ matrix \mathbf{B}_k removes the inter-user interference and the $(M - \tilde{\ell}_k) \times \ell_k$ matrix \mathbf{D}_k is used for parallelization and power allocation. It was also shown that \mathbf{D}_k is constructed from the right singular vectors of $\mathbf{F}_k \mathbf{B}_k$. This results in $\sigma_i(\mathbf{F}_k \mathbf{M}_k) = \sigma_i(\mathbf{F}_k \mathbf{B}_k)$ for $i = 1, \dots, \ell_k$. Therefore, the average throughput of a user k depends on the distribution of $\sigma_i^2(\mathbf{F}_k \mathbf{B}_k)$ for $i = 1, \dots, \ell_k$. In order to obtain the distribution of $\sigma_i^2(\mathbf{F}_k \mathbf{B}_k)$, the following definition and lemmas prove useful.

Definition Let \mathbf{Y} denote a $q \times p$ matrix with i.i.d zero-mean and unit variance complex Gaussian elements and $q \leq p$, then the Hermitian matrix $\mathbf{Y}\mathbf{Y}^H$ is an uncorrelated central Wishart matrix, denoted as $\mathbf{Y}\mathbf{Y}^H \sim \mathcal{W}_q(p, \mathbf{I}_q)$.

Lemma 4.1 (Chen, Heath and Andrews [7]) *If the $m \times n$ MIMO channel \mathbf{H}_k follows i.i.d complex Gaussian distribution $\mathcal{CN}(0, 1)$, then the effective MIMO channel after unitary precoding with $\mathbf{M}_k \in \mathbb{U}(n \times p)$, i.e. $\mathbf{H}_k \mathbf{M}_k$, is also i.i.d Gaussian distributed $\mathcal{CN}(0, 1)$, if \mathbf{M}_k is independent of \mathbf{H}_k .*

Lemma 4.2 Let \mathbf{X} be an $n \times m$ matrix, $r \leq \min\{m, n\}$. Let $\mathbf{A}_r = [\mathbf{a}_1 \cdots \mathbf{a}_r] \in \mathbb{U}(n, r)$ be an arbitrary unitary matrix and $\tilde{\mathbf{X}} = \mathbf{A}_r^H \mathbf{X} \in \mathbb{C}^{r \times m}$. Arranging the singular values of \mathbf{X} and $\tilde{\mathbf{X}}$ in decreasing order yields

$$\sigma_{i+n-r}(\mathbf{X}) \leq \sigma_i(\tilde{\mathbf{X}}) \leq \sigma_i(\mathbf{X}), \quad i = 1, 2, \dots, r \quad (9)$$

Proof Since $r \leq n$, we choose $n - r$ additional vectors $\mathbf{a}_{r+1}, \dots, \mathbf{a}_n$ such that $\mathbf{A} = [\mathbf{a}_1 \cdots \mathbf{a}_r, \mathbf{a}_{r+1}, \dots, \mathbf{a}_n] \in \mathbb{U}(n, n)$. We know that $\mathbf{A}\mathbf{X}$ has the same singular values as \mathbf{X} , and the matrix $\tilde{\mathbf{X}}$ is obtained by deleting the last $n - r$ rows of $\mathbf{A}\mathbf{X}$. The assertion now follows the corollary 3.1.3 in [8].

From Lemma 4.1, it can be concluded that $\mathbf{H}_k \mathbf{B}_k \in \mathbb{C}^{N \times (M - \ell_k)}$ is an i.i.d complex Gaussian matrix $\mathcal{CN}(0, 1)$. Let $\lambda_i(\mathbf{F}_k \mathbf{B}_k) = \sigma_i^2(\mathbf{F}_k \mathbf{B}_k)$, where we use $\lambda_i(\Phi)$ to denote the i th ordered eigenvalue of $\Phi \Phi^H$. Using the fact that $\mathbf{F}_k \mathbf{B}_k = [\mathbf{U}_k]_{(1:\ell_k)}^H \mathbf{H}_k \mathbf{B}_k$, $\ell_k \leq (M - \ell_k)$ and $\ell_k \leq N$ [4] it can be easily concluded from Lemma 4.2 that

$$\lambda_{i+N-\ell_k}(\mathbf{H}_k \mathbf{B}_k) \leq \lambda_i(\mathbf{F}_k \mathbf{B}_k) \leq \lambda_i(\mathbf{H}_k \mathbf{B}_k), \quad (10)$$

for $i = 1, 2, \dots, \ell_k$. The exact marginal distribution of both the upper bound and the lower bound of $\lambda_i(\mathbf{F}_k \mathbf{B}_k)$ in (10) is found in [9] and can be used for scheduling the users and UTMs, however, simulations results indicate that the upper bound is very tight and therefore is used to approximate the average throughput for each user. In the following, for the ease of notation, we represent $\lambda_i(\mathbf{H}_k \mathbf{B}_k)$ by λ_i . Using the upper bound of (10) in (8) and assuming that the ℓ_k received data streams by user k are decoded without any inter-stream interference, the approximate average throughput for user k with UTM of ℓ_k when the STM is L , is given by

$$\hat{I}_k(\ell_k, L) = \sum_{i=1}^{\ell_k} \int_0^\infty \log \left(1 + \frac{\Gamma_k P}{L} \lambda_i \right) f_{\lambda_i}(\lambda_i) d\lambda_i, \quad (11)$$

where $f_{\lambda_i}(\lambda_i)$ for $1 \leq i \leq N$ is found in [9]. As a result, the approximate average sum throughput for a given \mathcal{S}_L can be written as

$$\hat{R}(\mathcal{S}_L) = \mathbb{E} \left[\sum_{k=1}^{\tilde{K}} \hat{I}_k(\ell_k, L) \right]. \quad (12)$$

5. USER AND MODE SELECTION

Selection of the optimum set $\mathcal{S} = \arg \max_{\mathcal{S}_L} \bar{R}(\mathcal{S}_L)$ can be done in brute-force manner by computing (6) for all possible combination of users and UTMs with STM up to M , which is very complex, especially when the number of users increase. We propose an algorithm that finds the suboptimum set $\hat{\mathcal{S}} = \arg \max_{\mathcal{S}_L} \hat{R}(\mathcal{S}_L)$, by greedily increasing the STM L from 1 to M . For any given L , the average throughput approximations for all users at different UTMs $i = 1, 2, \dots, N$

```

1: Initialization:  $L = 1, \hat{\mathcal{S}} = \emptyset, \hat{R}(\emptyset) = 0$ 
2: while  $L \leq \min(KN, M)$  do
3:    $r = 0, \nu = 0, \mathcal{S}_L = \emptyset, \hat{I}_k(i, L) = 0, \ell_k = 0, u_k = 0$  for
      $k = 1, \dots, K$  and  $i = 1, \dots, N$ 
4:   Compute  $\hat{I}_k(i, L)$  from (8) for  $k = 1, \dots, K, i =$ 
      $1, \dots, \min\{N, L\}$ 
5:   while  $\nu \leq L$  do
6:     for  $k = 1$  to  $K$  do
7:        $s_k \leftarrow r + \hat{I}_k(\ell_k + 1, L) - \hat{I}_k(\ell_k, L)$ 
8:     end for
9:      $kmax \leftarrow \arg \max_k s_k, \ell_{kmax} \leftarrow \ell_{kmax} + 1$ 
10:     $r \leftarrow s_{kmax}, u_{kmax} \leftarrow 1, \nu \leftarrow \nu + 1$ 
11:   end while
12:   for  $k = 1$  to  $K$  do
13:     if  $u_k \neq 0$  then
14:        $\mathcal{S}_L = \mathcal{S}_L \cup \{(k, \ell_k)\}$ 
15:     end if
16:   end for
17:    $\hat{R}(\mathcal{S}_L) \leftarrow r$ 
18:   if  $\hat{R}(\mathcal{S}_L) < \hat{R}(\mathcal{S}_{L-1})$  then
19:      $\hat{\mathcal{S}} = \mathcal{S}_{L-1}$ , break
20:   end if
21:    $L \leftarrow L + 1$ 
22: end while

```

Table 1. Pseudo-code for the proposed algorithm

are computed using (11) and then L data streams which maximize (12) are selected.

Specifically in the L th iteration of the algorithm, the approximate average throughput of user k with UTM i , i.e. $\hat{I}_k(i, L)$, is computed using the fact that the BS should suppress the interference from $(L - i)$ data streams transmitted to the other scheduled users and at the same time serve the user k with i data streams. It results that $\mathbf{H}_k \mathbf{B}_k$ must be a $N \times (M - (L - i))$ matrix with i.i.d complex Gaussian elements. This dimension is used to compute the distribution of $\lambda_i(\mathbf{H}_k \mathbf{B}_k)$. The algorithm continues until the average sum throughput starts to decrease by further increase of L . The resulting algorithm is summarized in Table 1. Although the proposed greedy algorithm does not achieve the global optimum necessarily, but it achieves a good balance between performance and complexity.

6. SIMULATION RESULTS

For any given realization of user locations, the set $\hat{\mathcal{S}}$ is obtained using the algorithm in Table 1 and $\hat{R}(\hat{\mathcal{S}})$ is stored. For each obtained $\hat{\mathcal{S}}$ at each realization, 1000 iterations are simulated with independent channel states and $\bar{R}(\hat{\mathcal{S}})$ is also stored. In Fig. 1 we compare the average sum throughput versus the number of users K , for $M = 4, N = 2, 4$ and $P = 20$ dB. It is observed that the proposed scheme achieves about 90% of the DPC sum throughput. Furthermore, it can be seen the average sum throughput approximation matches very well with

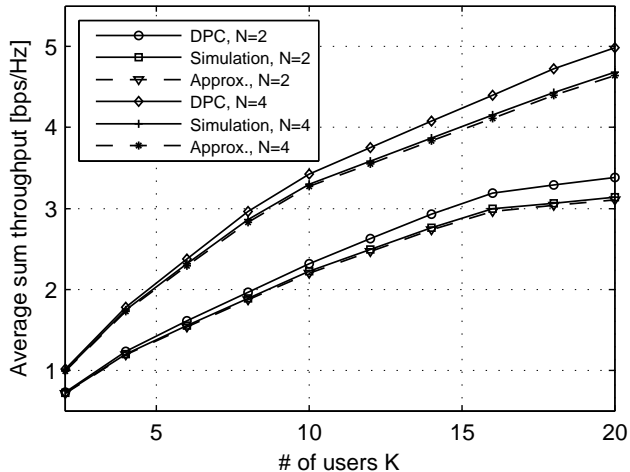


Fig. 1. Comparison of average sum throughput versus number of users K , for $M = 4$ transmit antennas, $N = 2, 4$ receive antennas and $P = 20$ dB

the actually achieved average sum throughput. In Fig. 2, the average feedback load (AFL) versus the number of users K is compared, where the AFL is defined as the average number of real coefficients that are fed back to the BS at each transmission interval. For the proposed scheme, all K user will feed back their average SNR for scheduling. Furthermore, each selected user k at UTM of ℓ_k should feed back ℓ_k real-valued singular values and ℓ_k complex-valued singular vectors of size M for the transmission. Therefore the AFL at each transmission interval would be

$$AFL = K + \mathbb{E}_{\tilde{K}, \ell_k} \left[\sum_{k=1}^{\tilde{K}} \ell_k (1 + 2M) \right], \quad (13)$$

while for the DPC all K users should feed back their $N \times M$ complex channel matrix. It is observed in Fig. 2 that the proposed scheme reduces the feedback by up to 90% for the case of $K = 20$, while the average sum throughput loss is about 6%.

7. CONCLUSION

In this paper, we propose a greedy algorithm for selecting users and the corresponding number of data streams to each user. The algorithm is only based on the knowledge of average received SNR for each user and thus reduces the amount of feedback for scheduling significantly. Simulation results indicate that the provided average sum throughput is very close to that of DPC. Therefore the proposed scheme is suitable for systems in which the performance measure is average sum throughput and the total feedback bits are limited.

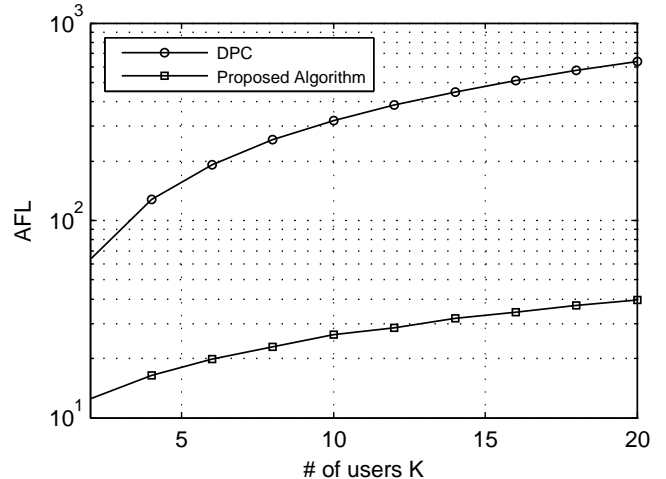


Fig. 2. Comparison of AFL versus number of users K , for $M = 4$ transmit antennas, $N = 4$ receive antennas and $P = 20$ dB

8. REFERENCES

- [1] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the gaussian MIMO broadcast channel," in *In Proc. IEEE Int. Sym. on Information Theory*, Jun. 2004, p. 174.
- [2] Juyul Lee and N. Jindal, "High SNR analysis for MIMO broadcast channels: Dirty paper coding versus linear precoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4787–4792, Dec. 2007.
- [3] H. Huang, M. Trivellato, A. Hottinen, M. Shafi, P. Smith, and R. Valenzuela, "Increasing downlink cellular throughput with limited network MIMO coordination," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2983–2989, June 2009.
- [4] F. Boccardi and H. Huang, "A near-optimum technique using linear precoding for the MIMO broadcast channel," in *IEEE Int. Conf. on Acoustics, Speech and Signal Proc. (ICASSP'07)*, 2007, Honolulu, Hawaii, USA, Apr. 2007, vol. 3, pp. III–17–III–20.
- [5] Runhua Chen, Zukang Shen, J.G. Andrews, and R.W. Heath, "Multimode transmission for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3294–3302, July 2008.
- [6] Seijoon Shim, Jin Kwak, R. Heath, and J. Andrews, "Block diagonalization for multi-user MIMO with other-cell interference," *IEEE Trans. Wireless Commun.*, vol. 7, no. 7, pp. 2671–2681, July 2008.
- [7] Runhua Chen, Robert W. Heath, and Jeffrey G. Andrews, "Transmit selection diversity for unitary precoded multiuser spatial multiplexing systems with linear receivers," *IEEE Trans. Signal Process.*, vol. 55, no. 3, pp. 1159–1171, March 2007.
- [8] Roger A. Horn and Charles R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1991.
- [9] A. Zanella, M. Chiani, and M.Z. Win, "On the marginal distribution of the eigenvalues of Wishart matrices," *IEEE Trans. Commun.*, vol. 57, no. 4, pp. 1050–1060, April 2009.