



<http://www.diva-portal.org>

## Postprint

This is the accepted version of a paper published in *IEEE Transactions on Signal Processing*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Björnson, E., Kountouris, M., Bengtsson, M., Ottersten, B. (2013)  
Receive Combining vs. Multi-Stream Multiplexing in Downlink Systems With Multi-Antenna Users.  
*IEEE Transactions on Signal Processing*, 61(13): 3431-3446  
<http://dx.doi.org/10.1109/TSP.2013.2260331>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:kth:diva-125858>

# Receive Combining vs. Multi-Stream Multiplexing in Downlink Systems with Multi-Antenna Users

Emil Björnson, *Member, IEEE*, Marios Kountouris, *Member, IEEE*, Mats Bengtsson, *Senior Member, IEEE*, and Björn Ottersten, *Fellow, IEEE*

**Abstract**—In downlink multi-antenna systems with many users, the multiplexing gain is strictly limited by the number of transmit antennas  $N$  and the use of these antennas. Assuming that the total number of receive antennas at the multi-antenna users is much larger than  $N$ , the maximal multiplexing gain can be achieved with many different transmission/reception strategies. For example, the excess number of receive antennas can be utilized to schedule users with effective channels that are near-orthogonal, for multi-stream multiplexing to users with well-conditioned channels, and/or to enable interference-aware receive combining. In this paper, we try to answer the question if the  $N$  data streams should be divided among few users (many streams per user) or many users (few streams per user, enabling receive combining). Analytic results are derived to show how user selection, spatial correlation, heterogeneous user conditions, and imperfect channel acquisition (quantization or estimation errors) affect the performance when sending the maximal number of streams or one stream per scheduled user—the two extremes in data stream allocation.

While contradicting observations on this topic have been reported in prior works, we show that selecting many users and allocating one stream per user (i.e., exploiting receive combining) is the best candidate under realistic conditions. This is explained by the provably stronger resilience towards spatial correlation and the larger benefit from multi-user diversity. This fundamental result has positive implications for the design of downlink systems as it reduces the hardware requirements at the user devices and simplifies the throughput optimization.

**Index Terms**—Multi-user MIMO, channel estimation, limited feedback, block-diagonalization, zero-forcing, receive combining.

## I. INTRODUCTION

The performance of downlink wireless communication systems can be improved by multi-antenna techniques, which

enable efficient utilization of spatial dimensions. Depending on the available channel state information (CSI), these dimensions can be used for enhanced reliability and/or spatial multiplexing of multiple data streams with controlled interference [1]. The downlink single-cell sum capacity (with perfect CSI) behaves as

$$\min(N, MK) \log_2(P) + \mathcal{O}(1) \quad (1)$$

where  $N$  is the number of base station antennas,  $K$  is the number of users, each user has  $M \geq 1$  antennas, and  $P$  is the signal-to-noise ratio (SNR) defined as the total transmit power divided by the noise power. The number of users is assumed to be large such that  $K \geq N$ , thus we have  $MK \geq N$  and the maximal *multiplexing gain* becomes  $\min(N, MK) = N$ . The multiplexing gain will have a major impact on the throughput of future cellular networks, where high SNRs can be achieved in an energy-efficient way by large-scale antenna arrays [2] and/or increased cell density [3].

The sum capacity in (1) is theoretically achieved by dirty-paper coding [4], but this non-linear scheme has impractical complexity and is very sensitive to CSI imperfections. Fortunately, the maximal multiplexing gain of  $N$  can be achieved by linear *spatial division multiple access* (SDMA) strategies [5], such as *block-diagonalization* (BD) [6], [7] and *zero-forcing with combining* (ZFC) [8], [9]. Such SDMA strategies transmit  $N$  simultaneous data streams, but can divide them among the users in different ways; the system can select between  $\lceil \frac{N}{M} \rceil$  and  $N$  users to be active and allocate from 1 to  $M$  streams to each of them. This raises a fundamental design question: *how should the receive antennas at each user be used to maximize the system throughput?*

Inter-user interference degrades user performance, while the mutual interference between users' own streams can be handled by receive processing. It thus seems beneficial to only have a few active users and multiplex many streams to each of them. However, one should keep in mind that every additional stream allocated to a user experiences a weaker channel gain than the previous streams. If fewer than  $M$  streams are allocated to a user, this user has degrees of freedom for interference-aware receive combining to achieve a strong effective channel and better spatial co-user compatibility. In other words, it is not clear whether receive antennas should be utilized for *multi-stream multiplexing* or *receive combining*, or perhaps something intermediate. The answer has a profound impact on wireless system design, including the CSI acquisition protocols, scheduling algorithms, and receiver architecture.

©2013 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

The research leading to these results has received funding from the European Research Council under the European Community's Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement number 228044. The work of E. Björnson is funded by the International Postdoc Grant 2012-228 from The Swedish Research Council. This work was presented in part at the IEEE Swedish Communication Technologies Workshop (Swe-CTW), Stockholm, Sweden, October 2011 [32].

E. Björnson, M. Bengtsson, and B. Ottersten are with the Signal Processing Laboratory, ACCESS Linnaeus Center, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden (e-mail: emil.bjornson@ee.kth.se; mats.bengtsson@ee.kth.se; bjorn.ottersten@ee.kth.se). B. Ottersten is also with Interdisciplinary Centre for Security, Reliability and Trust (SnT), University of Luxembourg, L-1359 Luxembourg-Kirchberg, Luxembourg (email: bjorn.ottersten@uni.lu). M. Kountouris and E. Björnson are with SUPELEC (Ecole Supérieure d'Electricité), Gif-sur-Yvette, France (e-mail: marios.kountouris@supelec.fr; emil.bjornson@supelec.fr).

### A. Related Work

The sum-rate maximization problem is nonconvex and combinatorial [10], thus only suboptimal strategies are feasible in practice. Such low-complexity algorithms have been proposed in [11]–[14], among others, by successively allocating data streams to users in a greedy manner. Simulations have indicated that fewer than  $N$  streams should be used when  $P$  and  $K$  are small, and that spatial correlation makes it beneficial to divide the streams among many users. Simulations in [12] indicates that the probability of allocating more than one stream per user is small when  $K$  grows large, but [12] only considers users with homogeneous channel conditions and all the aforementioned papers assume perfect CSI.

The authors of [9] claim that transmitting at most one stream per user is desirable when there are many users in the system. They justify this statement by using asymptotic results from [15] where  $K \rightarrow \infty$ . This argumentation ignores some important issues: 1) asymptotic optimality can also be proven with multiple streams per user;<sup>1</sup> 2) the performance at practical values on  $K$  is unknown; and 3) the analysis implies an unbounded asymptotic multi-user diversity gain, which is a modeling artifact of fading channels [17].

The authors of [7], [8] arrive at a different conclusion when they compare BD (which selects  $\frac{N}{M}$  users and sends  $M$  streams/user) and ZFC (which selects  $N$  users and sends one stream/user) under quantized CSI. Their simulations reveal a distinct advantage of BD (i.e., multi-stream multiplexing), but are limited to uncorrelated channels and neither include user selection nor interference rejection. We show that their results are misleading, because single-user transmission greatly outperforms both BD and ZFC in the scenario that they simulate.

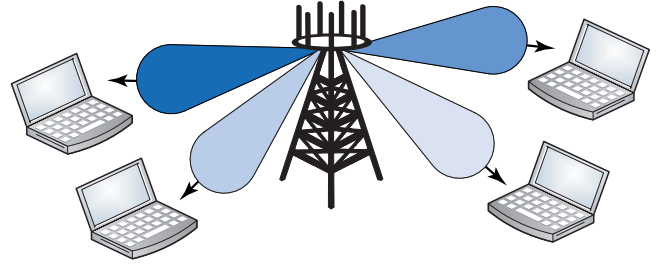
Despite the similar terminology, our problem is fundamentally different from the classic works on the diversity-spatial multiplexing tradeoff (DMT) in [18], [19]. The DMT brings insight on how many streams should be transmitted in the high-SNR regime, while we consider how a fixed number of streams should be divided among the users.

### B. Main Contributions

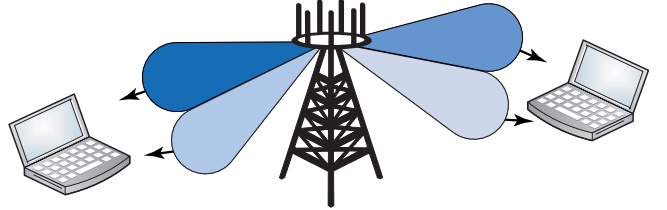
This paper provides a comprehensive answer to how multi-antenna users should utilize their antennas in downlink transmissions, or similarly how many data streams that should be allocated per active user under different system conditions; see Fig. 1. The main contributions are:

- New analytic results for analyzing the problem under spatial correlation, user selection, heterogeneous user channel conditions, and realistic CSI acquisition. These enable asymptotic comparison of the two extremes: allocating  $M$  streams per active user (called BD) and one stream per active user (called ZFC). We show that ZFC is more resilient to spatial correlation and well adapted to find near-orthogonal users, while BD is better at utilizing heterogeneous user conditions. Imperfect CSI acquisition is shown to have a similar impact on both strategies.

<sup>1</sup>The uplink analysis in [16] shows that a non-zero (but bounded) number of users can use multiple streams, and the well-established uplink-downlink duality makes this result applicable also in our downlink scenario.



(a) 1 stream per 2-antenna user: ZFC enables receive combining.



(b) 2 streams per 2-antenna user: BD exploits multi-stream multiplexing.

Fig. 1. Two ways of dividing four data streams among multi-antenna users, which also represents two ways of utilizing the receive antennas to reduce interference. (a) Receive one stream per user and linearly combine the antenna to achieve an effective channel that rejects interference. (b) Receive multiple streams and handle their mutual interference through receive processing.

- Numerical examples show that allocating one stream per active user is essentially optimal under realistic system conditions, and we explain how other conclusions may arise. The main conclusion is that utilizing receive combining is preferable over multi-stream multiplexing.

## II. SYSTEM MODEL

We consider a downlink multi-user MIMO system where a single base station with  $N$  antennas communicates with  $K \geq N$  users. Each user has  $M$  antennas. For analytical convenience<sup>2</sup> we assume that  $M < N$  and often also that  $\frac{N}{M}$  is an integer, but the precoding strategies considered herein can be applied for any  $M$ . The narrowband, flat-fading channel to user  $k$  is represented in the complex-baseband by  $\mathbf{H}_k \in \mathbb{C}^{M \times N}$ . The received signal at this user is

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k \quad (2)$$

where  $\mathbf{x} \in \mathbb{C}^{N \times 1}$  is the joint transmitted signal for all users and  $\mathbf{n}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$  is the (normalized) circularly-symmetric complex Gaussian noise vector. For analytic convenience, and motivated by measurements [20], [21], we employ the Kronecker model with  $\mathbf{H}_k = \mathbf{R}_{R,k}^{1/2} \tilde{\mathbf{H}}_k \mathbf{R}_{T,k}^{1/2}$ , where  $\mathbf{R}_{T,k}$  and  $\mathbf{R}_{R,k}$  are the positive-definite spatial correlation matrices at the transmitter and receiver side, respectively, and  $\tilde{\mathbf{H}}_k$  has independent  $\mathcal{CN}(0,1)$ -entries. We assume  $\mathbf{R}_{T,k} = \mathbf{I}_N$  (i.e., large antenna separation at the base station) throughout the analysis, because transmit correlation both creates complicated mathematical structures and requires limiting assumptions on

<sup>2</sup>The case  $M \geq N$  is analytically different because 1) Single-user transmission achieves the full multiplexing gain; 2) CSI acquisition is simplified since  $\mathbf{H}_k$  has full row rank, thus any effective channel  $\mathbf{C}_k^H \mathbf{H}_k$  can be achieved by selecting the receive combining  $\mathbf{C}_k$  properly. Since user devices are size-constrained, the case  $M < N$  is also reasonable in practice.

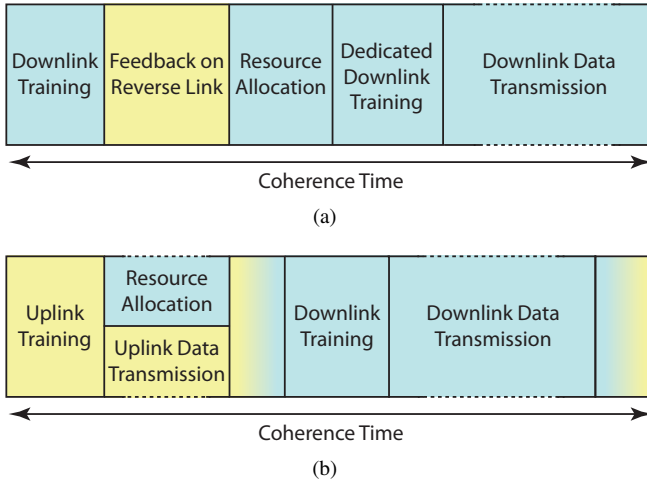


Fig. 2. Basic block-fading system operation of (a) FDD systems; and (b) TDD systems. The system operation is repeated in a cyclic manner.

the user distribution geometry and fading environment. Observe that  $\mathbf{R}_{R,k}$  generally is different for each users, describing different spatial properties.

#### A. Cyclic System Operation

We assume block fading where  $\mathbf{H}_k$  is static for a set of channel uses, called the *coherence time*, and then updated independently. We consider both frequency division duplex (FDD) and time division duplex (TDD); baselines of the respective cyclic system operations are illustrated in Fig. 2.

In FDD systems, the users acquire CSI through training signaling [22] and some users feed back quantized CSI. The base station then performs resource allocation (i.e., data stream allocation and precoding) and informs the scheduled users of their precoding through a second training stage. Data transmission follows until the end of the coherence time, when the cycle in Fig. 2(a) restarts.

In TDD systems, the system toggles between uplink and downlink transmission on the same channel, thus enabling training signaling in both directions. We assume perfect channel reciprocity<sup>3</sup> and that the coherence time makes CSI obtained in one block of Fig. 2(b) correct until the same block occurs in the next cycle. The base station does resource allocation for both uplink and downlink, and it informs the users through training signaling.

We assume that all training signals sent in the downlink direction provide the users with perfect CSI, while CSI feedback (in FDD) and uplink training (in TDD) might lead to imperfect CSI at the base station. This assumption enables coherent reception, thus making the conventional achievable sum rate expression a reasonable performance measure.<sup>4</sup>

<sup>3</sup>The physical channel is always reciprocal, but different transceiver hardware is typically used in the downlink and the uplink. Thus, careful calibration is necessary to utilize the reciprocity in practice.

<sup>4</sup>Many of the results herein can be extended to include imperfect CSI at the users in the resource allocation, followed by a second training stage that provides scheduled users with sufficiently accurate CSI of the precoded channels to enable coherent reception. See [23] for an example in FDD systems. The loss of having imperfect CSI also in the second training stage can be characterized as in [24].

#### B. Linear Precoding: General Problem Formulation

We consider linear precoding and the transmitted signal is

$$\mathbf{x} = \sum_{k=1}^K \mathbf{W}_k \mathbf{d}_k \quad (3)$$

where  $\mathbf{W}_k \in \mathbb{C}^{N \times d_k}$  is the precoding matrix,  $\mathbf{d}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_{d_k})$  is the data signal, and  $d_k$  is the number of multiplexed data streams to user  $k$ . Each user applies a semi-unitary receive combining matrix  $\mathbf{C}_k \in \mathbb{C}^{M \times d_k}$  (i.e.,  $\mathbf{C}_k^H \mathbf{C}_k = \mathbf{I}_{d_k}$ ) and treats inter-user interference as Gaussian noise. The achievable information rate is

$$g_k(\{\mathbf{W}_\ell\}, \mathbf{C}_k) = \log_2 \frac{\det(\mathbf{I}_{d_k} + \sum_{\ell=1}^K \mathbf{C}_k^H \mathbf{H}_k \mathbf{W}_\ell \mathbf{W}_\ell^H \mathbf{H}_k^H \mathbf{C}_k)}{\det(\mathbf{I}_{d_k} + \sum_{\ell \neq k} \mathbf{C}_k^H \mathbf{H}_k \mathbf{W}_\ell \mathbf{W}_\ell^H \mathbf{H}_k^H \mathbf{C}_k)} \quad (4)$$

where  $\{\mathbf{W}_\ell\}$  denotes the set of precoding matrices and  $\ell$  is an arbitrary user index [14]. The transmission is limited by an average power/SNR constraint of  $P$ , thus

$$\mathbb{E}\{\mathbf{x}^H \mathbf{x}\} = \sum_{k=1}^K \text{tr}(\mathbf{W}_k \mathbf{W}_k^H) \leq P. \quad (5)$$

Ideally, we would like to select  $\mathbf{W}_k, \mathbf{C}_k, d_k \forall k$  to maximize the sum rate; that is,

$$\begin{aligned} & \underset{\{\mathbf{W}_k, \mathbf{C}_k, d_k\}}{\text{maximize}} \quad \sum_{k=1}^K g_k(\{\mathbf{W}_\ell\}, \mathbf{C}_k) \\ & \text{subject to} \quad \sum_{k=1}^K \text{tr}(\mathbf{W}_k \mathbf{W}_k^H) \leq P, \\ & \quad \mathbf{C}_k^H \mathbf{C}_k = \mathbf{I}_{d_k}, \quad d_k \geq 0 \quad \forall k. \end{aligned} \quad (6)$$

Unfortunately, this resource allocation problem is NP-hard and therefore not practically solvable [10]. There are algorithms that find local optima of (6) (see [25] and references therein), but these are iterative and thus cannot be implemented under the cyclic system operation in Fig. 2.

We limit the selection of  $\{\mathbf{W}_k, \mathbf{C}_k, d_k\}$  to achieve a tractable problem formulation.

- 1) **Precoding:** Zero or minimal inter-user interference should be caused, which is possible when  $\sum_{k=1}^K d_k \leq N$ . This makes (6) partially feasible, because it becomes a convex problem for any fixed  $\mathbf{C}_k, d_k$ . This is a non-limiting assumption at high SNR [26], which is the regime where systems with high spectral efficiencies need to operate (e.g., using high power, small cells, or large antenna arrays [2], [3]).
- 2) **Receive combining:** The matrix  $\mathbf{C}_k$  is fixed at some value  $\tilde{\mathbf{C}}_k$  beforehand. This makes sense from a CSI acquisition perspective as only the effective channel  $\tilde{\mathbf{C}}_k^H \mathbf{H}_k$  needs to be obtained through feedback (in FDD) or training signaling (in TDD). The value  $\tilde{\mathbf{C}}_k$  might be the  $d_k$  strongest (left) singular vectors of  $\mathbf{H}_k$ , known as maximum ratio combining (MRC), but can also be selected to improve the CSI feedback accuracy [8], [9].
- 3) **Stream allocation:** Users are scheduled sequentially using some predefined scheduling policy. This avoids

making an exhaustive search over all data stream allocations, which is practically infeasible when  $N$  and  $K$  grow large. Greedy scheduling algorithms can perform remarkably close to optimum [11]–[14], while random selection ensures user fairness.

We now have a simplified resource allocation problem,

$$\begin{aligned} & \underset{\{\mathbf{W}_k\}}{\text{maximize}} \quad \sum_{k \in \mathcal{S}} \log_2 \det(\mathbf{I}_{d_k} + \tilde{\mathbf{C}}_k^H \mathbf{H}_k \mathbf{W}_k \mathbf{W}_k^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k) \\ & \text{subject to} \quad \sum_{k \in \mathcal{S}} \text{tr}(\mathbf{W}_k \mathbf{W}_k^H) \leq P, \\ & \quad \tilde{\mathbf{C}}_k^H \mathbf{H}_k \mathbf{W}_\ell = \mathbf{0}_{d_k \times d_\ell} \quad \forall k \in \mathcal{S}, \forall \ell \in \mathcal{S} \setminus \{k\}, \end{aligned} \quad (7)$$

where  $\mathcal{S}$  is the scheduling set given by the predefined scheduling rule and  $d_k > 0$  for  $k \in \mathcal{S}$  is the corresponding data stream allocation.

*Remark 1* (Updating the Receive Combiner). When (7) has been solved, the users are informed of the resource allocation through training signaling. This enables estimation of both the precoded channel  $\mathbf{H}_k \mathbf{W}_k$  and the second-order interference term  $\mathcal{I}_k = \sum_{\ell \neq k} \mathbf{H}_k \mathbf{W}_\ell \mathbf{W}_\ell^H \mathbf{H}_k^H$ , both being necessary for coherent reception. As a nice by-product [9], this enables user  $k$  to replace  $\tilde{\mathbf{C}}_k$  with the rate-maximizing MMSE receive combiner  $\mathbf{C}_k^{\text{MMSE}}$  containing the  $d_k$  dominating left singular vectors of  $(\mathbf{I}_M + \mathcal{I}_k)^{-1} \mathbf{H}_k \mathbf{W}_k$  [14]. This improves the information rate by balancing between signal gain and interference rejection. We consider  $\tilde{\mathbf{C}}_k$  in the analysis, while  $\mathbf{C}_k^{\text{MMSE}}$  is used in simulations.

### C. Linear Precoding: BD and ZFC

In this paper, we primarily analyze and compare two instances of (7): *block-diagonalization (BD)* [6] and *zero-forcing with combining (ZFC)* [8], [9]. These strategies allocate a fixed number of streams per scheduled user, but can be combined with any scheduling policy. There are alternative strategies that allocate different numbers of streams to different users [12], but simulations will show that these are *not* increasing the performance when the CSI acquisition overhead is treated properly.

*Definition 1. (Block-Diagonalization Precoding)* Let  $\mathcal{S}^{\text{BD}}$  be a scheduling set with at most  $\frac{N}{M}$  users. For each user  $k \in \mathcal{S}^{\text{BD}}$ , we set  $d_k = M$  and  $\mathbf{W}_k = \mathbf{W}_k^{\text{BD}} \mathbf{\Upsilon}_k^{1/2}$ , where  $\mathbf{W}_k^{\text{BD}}$  is a semi-unitary matrix that satisfies  $\mathbf{W}_k^{\text{BD},H} \mathbf{W}_k^{\text{BD}} = \mathbf{I}_M$  and  $\mathbf{H}_\ell \mathbf{W}_k^{\text{BD}} = \mathbf{0}$  for all  $\ell \in \mathcal{S}^{\text{BD}} \setminus \{k\}$ . The power allocation is given by the diagonal matrix  $\mathbf{\Upsilon}_k \succeq \mathbf{0}_M$ . The information rate is

$$g_k^{\text{BD}}(P) = \log_2 \det \left( \mathbf{I}_M + \mathbf{H}_k \mathbf{W}_k^{\text{BD}} \mathbf{\Upsilon}_k \mathbf{W}_k^{\text{BD},H} \mathbf{H}_k^H \right). \quad (8)$$

*Definition 2. (Zero-Forcing Precoding with Combining)* Each user combines its antennas using some channel-dependent unit-norm vector  $\tilde{\mathbf{c}}_k \in \mathbb{C}^{M \times 1}$ . Based on the effective channels  $\mathbf{h}_k^H = \tilde{\mathbf{c}}_k^H \mathbf{H}_k \in \mathbb{C}^{1 \times N}$ , a scheduling set  $\mathcal{S}^{\text{ZFC}}$  with at most  $N$  users is selected. For each user  $k \in \mathcal{S}^{\text{ZFC}}$ , we set  $d_k = 1$  and let  $\mathbf{W}_k = \sqrt{p_k} \mathbf{w}_k^{\text{ZFC}}$ , where  $\mathbf{w}_k^{\text{ZFC}}$  is a unit-norm vector

that satisfies  $\mathbf{h}_\ell^H \mathbf{w}_k^{\text{ZFC}} = 0$  for all  $\ell \in \mathcal{S}^{\text{ZFC}} \setminus \{k\}$ . The power  $p_k \geq 0$  is allocated to user  $k$  and the information rate is

$$g_k^{\text{ZFC}}(P) = \log_2 \left( 1 + p_k |\mathbf{h}_k^H \mathbf{w}_k^{\text{ZFC}}|^2 \right). \quad (9)$$

The sum-rate maximizing power allocations for BD and ZFC are achieved through water-filling (see [6]), but the asymptotic analysis in this paper often assumes equal power allocation (i.e.,  $\mathbf{\Upsilon}_k = \frac{P}{M|\mathcal{S}^{\text{BD}}|} \mathbf{I}_M \forall k \in \mathcal{S}^{\text{BD}}$  and  $p_k = \frac{P}{|\mathcal{S}^{\text{ZFC}}|} \forall k \in \mathcal{S}^{\text{ZFC}}$ ) since this becomes optimal in the high-SNR regime where  $P \rightarrow \infty$  [27]. Although the definitions of BD and ZFC assume perfect CSI, both strategies can be applied when the transmitter has imperfect CSI by making  $\mathbf{W}_k$  orthogonal to the acquired co-user channels [7]–[9]. The resulting loss will be quantified in later sections.

ZFC can schedule up to  $N$  users and sends one data stream per user, while BD can only schedule  $\frac{N}{M}$  users but multiplexes  $M$  streams to each of them. Although BD and ZFC are identical when each user only has one antenna, this does *not* mean that BD is a generalization of ZFC. In fact, there are good reasons for applying ZFC instead of BD when  $M > 1$ :

- 1) The base station only needs to acquire the effective channels  $\mathbf{h}_k$ ;
- 2) The effective channel  $\mathbf{h}_k$  has better properties than  $\mathbf{H}_k$  and can be adapted for interference rejection;
- 3) User devices require simpler hardware that only decodes one stream.

The interference mitigation is, on the other hand, less restrictive under BD since fewer users are involved and the mutual interference between streams sent to the same user is handled by receive processing [7]. By analyzing and comparing ZFC and BD under both perfect and imperfect CSI, we try to answer the fundamental question: should we select many multi-antenna users to enable receive combining or select few users and exploit multi-stream multiplexing?

*Remark 2* (Ambiguous Terminology). The terminology *block-diagonalization* and *zero-forcing* have been given different meanings in prior works. Herein, BD refers to the original work in [6], where each active user receives exactly  $M$  data streams. Apart from the ZFC strategy in Definition 2 (and in [8], [9]), another downlink zero-forcing strategy for multi-antenna users was proposed in [26]. In their definition, each antenna at the multi-antenna users is viewed as a separate virtual single-antenna user and the zero-forcing idea is applied to send a separate stream to each antenna with zero inter-antenna interference. That approach is nothing else than BD with stricter interference mitigation and can *never* perform better than BD. Herein, ZFC means sending one stream per user and utilizing receive combining, thus ZFC is *not* a special case of BD and can hypothetically outperform BD.

### III. COMPARISON OF BD AND ZFC WITH PERFECT CSI

In this section, we will compare BD and ZFC in the ideal scenario when both the base station and the users have perfect CSI. We derive analytic results indicating the impact of different system properties. Under perfect CSI, the achievable

sum rate in (7) asymptotically becomes (as  $P \rightarrow \infty$ ) [27]

$$\begin{aligned} f_{\text{sum}}^{\text{BD}}(P) &\cong N \log_2 \left( \frac{P}{N} \right) + \sum_{k \in \mathcal{S}^{\text{BD}}} \log_2 \det(\mathbf{H}_k \mathbf{W}_k^{\text{BD}} \mathbf{W}_k^{\text{BD},H} \mathbf{H}_k^H), \\ f_{\text{sum}}^{\text{ZFC}}(P) &\cong N \log_2 \left( \frac{P}{N} \right) + \sum_{k \in \mathcal{S}^{\text{ZFC}}} \log_2(|\mathbf{h}_k^H \mathbf{w}_k^{\text{ZFC}}|^2), \end{aligned} \quad (10)$$

for BD and ZFC, respectively. This result is based on having scheduling sets that satisfy  $|\mathcal{S}^{\text{BD}}| = \frac{N}{M}$  and  $|\mathcal{S}^{\text{ZFC}}| = N$  and on equal power allocation (which is asymptotically optimal).

For both strategies, the asymptotic sum rate behaves as  $\mathcal{M}_\infty \log_2(P) + \mathcal{R}_\infty$ , where  $\mathcal{M}_\infty$  is the multiplexing gain and  $\mathcal{R}_\infty$  is the rate offset. Both BD and ZFC achieve a multiplexing gain of  $\mathcal{M}_\infty = N$ , which is the same high-SNR slope as of the sum capacity. We thus need to compare the rate offsets  $\mathcal{R}_\infty$  to conclude which strategy is preferable in the high-SNR regime.

*Theorem 1.* Assume the receive correlation matrices  $\mathbf{R}_{R,k}$  have eigenvalues  $\lambda_{k,M} \geq \dots \geq \lambda_{k,1} > 0$  and the use of random user selection with  $|\mathcal{S}^{\text{BD}}| = \frac{N}{M}$ ,  $|\mathcal{S}^{\text{ZFC}}| = N$ . The expected asymptotic difference in sum rate between BD and ZFC (with MRC) is

$$\begin{aligned} \bar{\beta}_{\text{BD-ZFC}} &= \mathbb{E} \left\{ \lim_{P \rightarrow \infty} f_{\text{sum}}^{\text{BD}}(P) - f_{\text{sum}}^{\text{ZFC}}(P) \right\} \\ &= N \frac{\log_2(e)}{M} \sum_{i=1}^{M-1} \frac{M-i}{i} + \log_2 \left( \prod_{k \in \mathcal{S}^{\text{BD}}} \prod_{m=1}^M \lambda_{k,m} \right) - \sum_{\ell \in \mathcal{S}^{\text{ZFC}}} z_\ell \end{aligned} \quad (11)$$

where  $z_\ell = \mathbb{E}\{\log_2(\|\tilde{\mathbf{c}}_\ell^H \mathbf{H}_\ell\|_2^2)\} - \frac{\psi(N)}{\log_e(2)}$  and  $\psi(\cdot)$  is the digamma function. Furthermore,  $\log_2(\lambda_{\ell,M}) \leq z_\ell \leq \log_2(\mathbb{E}\{\|\tilde{\mathbf{c}}_\ell^H \mathbf{H}_\ell\|_2^2\}) - \frac{\psi(N)}{\log_e(2)}$  where  $\mathbb{E}\{\|\tilde{\mathbf{c}}_\ell^H \mathbf{H}_\ell\|_2^2\}$  is given by (32) in Lemma 1.

*Proof:* The proof is given in Appendix B. ■

The expected asymptotic difference in (11) has several terms. The first term is the (positive) expected gain of BD in a spatially uncorrelated scenario with homogenous user channels and no receive combining—this was considered in [27, Theorem 3]. The other terms depend on the spatial correlation and choice of receive combining. For users with homogenous channel conditions where all  $\mathbf{R}_{R,k}$  have the same eigenvalues  $\lambda_{k,m} = \lambda_m$ , we have

$$\bar{\beta}_{\text{BD-ZFC}} \leq N \frac{\log_2(e)}{M} \sum_{i=1}^{M-1} \frac{M-i}{i} + N \log_2 \frac{\prod_{m=1}^M \lambda_m^{1/M}}{\lambda_M} \quad (12)$$

where the last term contains the geometric mean of all eigenvalues divided by the largest eigenvalue. This ratio is smaller than one (or equal for uncorrelated channels) and thus its logarithm is negative and approaches  $-\infty$  as the eigenvalue spread increases. Therefore, Theorem 1 shows that BD *might* have an advantage on uncorrelated channels, but ZFC always becomes the better choice as the receive-side correlation grows. The explanation is that BD has less restrictive interference mitigation, but is more vulnerable to poor channels since it uses all channel dimensions for transmission.

We can expect a similar impact of any channel property that increases the eigenvalue spread in  $\mathbf{H}_k \mathbf{H}_k^H$ ; for example, spatial correlation at the transmitter-side or a strong (low-rank) line-of-sight component.

To illustrate the opposite effect of having users with different path losses, we assume for simplicity that there are  $\frac{N}{M}$  strong users with  $\mathbf{R}_{R,k} = \gamma \mathbf{I}_M$ , for some  $\gamma > 1$ , and  $N - \frac{N}{M}$  weak users with  $\mathbf{R}_{R,k} = \mathbf{I}_M$ . If BD only serves the strong users while ZFC serves also the weak users, we have

$$\bar{\beta}_{\text{BD-ZFC}} \leq N \frac{\log_2(e)}{M} \sum_{i=1}^{M-1} \frac{M-i}{i} + \left( N - \frac{N}{M} \right) \log_2(\gamma). \quad (13)$$

This upper bound approaches  $+\infty$  as the difference  $\gamma$  between the strong and weak users grows. Although not strictly proved, this indicates that BD is better at utilizing heterogeneous channel conditions as it requires fewer users to be close to the base station to achieve high sum rates. This benefit reduces if some fairness mechanism is used to compensate for unfavorable path losses.

The expected asymptotic difference in sum rate,  $\bar{\beta}_{\text{BD-ZFC}}$ , can be transformed into a difference  $-\frac{\bar{\beta}_{\text{BD-ZFC}}}{10N \log_{10}(2)}$  [dB] in transmit power to achieve the same sum rate in the high-SNR regime [27].

#### A. Impact of User Selection

The comparison in Theorem 1 was based on random user selection of the maximal number of users ( $\frac{N}{M}$  with BD and  $N$  with ZFC), although scheduling of spatially separated users is necessary to achieve the full potential of multi-user MIMO. This paper assumes  $K \geq N$  users, meaning that only a subset of users is scheduled at each channel use. If the users are unevenly distributed in the cell, it could be beneficial to intentionally schedule fewer users than possible. We will now analyze how the ability of selecting users with spatially compatible channels impacts performance.

In the high-SNR regime, the optimal (semi-unitary) precoding matrix  $\mathbf{W}_k^{\text{su}}$  for single-user transmission matches the channel as  $\tilde{\mathbf{C}}_k^H \mathbf{H}_k \mathbf{W}_k^{\text{su}} = \tilde{\mathbf{C}}_k^H \mathbf{H}_k$ , while the precoding matrix  $\mathbf{W}_k \in \mathbb{C}^{N \times d_k}$  of an SDMA strategy is balanced between matching the own channel and being orthogonal to the co-user channels. The expected asymptotic performance loss of having to cancel inter-user interference is therefore

$$\begin{aligned} \mathbb{E}\{\text{Loss}\} &= \mathbb{E}\{\log_2 \det(\tilde{\mathbf{C}}_k^H \mathbf{H}_k \mathbf{H}_k^H \tilde{\mathbf{C}}_k) \\ &\quad - \log_2 \det(\tilde{\mathbf{C}}_k^H \mathbf{H}_k \mathbf{W}_k \mathbf{W}_k^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k)\} \\ &= \mathbb{E}\left\{ \log_2 \frac{\det(\mathbf{\Lambda}_k \mathbf{\Lambda}_k^H)}{\det(\mathbf{\Lambda}_k \mathbf{B}_k \mathbf{W}_k \mathbf{W}_k^H \mathbf{B}_k^H \mathbf{\Lambda}_k^H)} \right\} \\ &= -\mathbb{E}\{\log_2 \det(\mathbf{B}_k \mathbf{W}_k \mathbf{W}_k^H \mathbf{B}_k^H)\} \end{aligned} \quad (14)$$

where  $\mathbf{\Lambda}_k \in \mathbb{C}^{d_k \times d_k}$  contains the non-zero singular values of  $\tilde{\mathbf{C}}_k^H \mathbf{H}_k$  and  $\mathbf{B}_k$  contains the corresponding right singular vectors.<sup>5</sup> Observe that the eigenvalues of  $\mathbf{B}_k \mathbf{W}_k \mathbf{W}_k^H \mathbf{B}_k^H$  are

<sup>5</sup>These matrices can be obtained from a compact singular value decomposition  $\tilde{\mathbf{C}}_k^H \mathbf{H}_k = \mathbf{U}_k \mathbf{\Lambda}_k \mathbf{B}_k$ . Note that  $\mathbf{B}_k$  contains an orthonormal basis of the row space of the effective channel  $\tilde{\mathbf{C}}_k^H \mathbf{H}_k$ .



smaller or equal to one, thus  $\mathbb{E}\{\text{Loss}\} \geq 0$ . The following theorem indicates how this loss is affected by user selection.

**Theorem 2.** For any given scheduling sets  $\mathcal{S}^{\text{BD}}, \mathcal{S}^{\text{ZFC}}$  (with  $|\mathcal{S}^{\text{BD}}| = \frac{N}{M}$  and  $|\mathcal{S}^{\text{ZFC}}| = N$ ), suppose we replace one of the users in each set with the best one among  $K$  random users. If the best user is the one minimizing the expected asymptotic loss in (14), these losses for BD and ZFC, respectively, can be lower bounded as

$$\begin{aligned} \mathbb{E}\{\text{Loss}_{\text{BD}}\} &\geq -M \log_2(1 - c_1 K^{-\frac{1}{M(N-M)}}) \\ \mathbb{E}\{\text{Loss}_{\text{ZFC}}\} &\geq -\log_2(1 - c_2 K^{-\frac{1}{N-M}}) \end{aligned} \quad (15)$$

when  $K$  is large ( $c_1, c_2$  are positive constants, see the proof).

*Proof:* The proof is given in Appendix C. ■

The lower bounds in this theorem indicate that it is easier to find users with near-orthogonal channels under ZFC than under BD. This seems reasonable since the random channels of BD users occupy  $M$  dimensions and should happen to be compatible to the co-users in all of them, while ZFC users only utilize one dimension and use receive combining to pick the most compatible among its  $M$  dimensions. Related observations can be made in the area of channel quantization, where fewer codewords are necessary to describe  $(N \times 1)$ -dimensional channels to a certain accuracy than are needed for  $(N \times M)$ -dimensional channels [28]. The concave structure of the information rates makes it difficult to obtain exact results, but the indications of Theorem 2 are verified by simulations herein.

### B. Numerical Illustrations under Perfect CSI

Next, the analytic properties in Theorem 1 and Theorem 2 are illustrated numerically. To this end, we adopt the simple exponential correlation model of [29], where  $0 \leq \rho \leq 1$ ,  $\iota = \sqrt{-1}$ ,  $U[\cdot, \cdot)$  denotes a uniform distribution, and

$$[\mathbf{R}(\rho, \theta)]_{ij} = \begin{cases} (\rho e^{i\theta})^{j-i}, & i \leq j, \\ (\rho e^{-i\theta})^{i-j}, & i > j, \end{cases} \quad \theta \sim U[0, 2\pi). \quad (16)$$

The magnitude  $\rho$  is the *correlation factor* between adjacent antennas, where  $\rho = 0$  means no spatial correlation and  $\rho = 1$  means full correlation. For simplicity,  $\rho$  is the same for all users while  $\theta$  is different. Note that  $\rho$  impacts the perceived spatial correlation non-linearly; a typical angular spread in a highly spatially correlated scenario is  $10 - 20$  degrees which roughly corresponds to  $\rho \approx 0.9$  [30].

The expected asymptotic difference between BD and ZFC is shown in Fig. 3 as a function of  $\rho$ , using  $N = 8$  transmit antennas and  $M = 2$  receive antennas. This simulation confirms that BD is advantageous in uncorrelated systems, while ZFC becomes beneficial as the correlation increases ( $\rho > 0.4$  under receive-side correlation,  $\rho > 0.7$  under transmit-side correlation, and  $\rho > 0.25$  when both sides are correlated). The two bounds from Theorem 1 are also shown in the Fig. 3. The lower bound is very accurate, while the upper bound is only tight at high correlation.

To exemplify the impact of user selection, we use the *capacity-based suboptimal user selection* (CBSUS) algorithm from [31], which greedily adds users sequentially to maximize

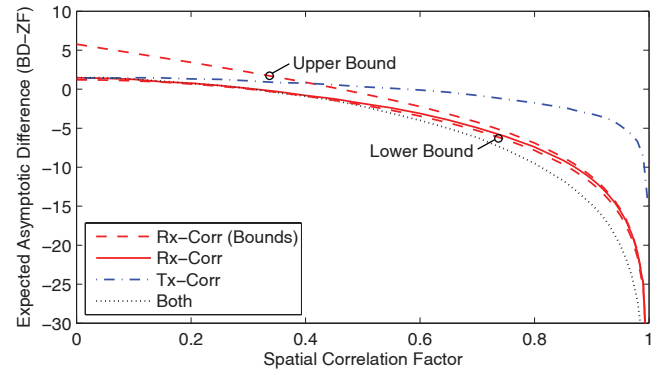


Fig. 3. The expected asymptotic difference between BD and ZFC in a system with  $N = 8$  transmit antennas,  $M = 2$  receive antennas per user, and random user selection. The impact of spatial correlation at the receiving users, transmitting base station, and both sides is shown (using the exponential correlation model from [29] with different correlation factors  $\rho$ ).

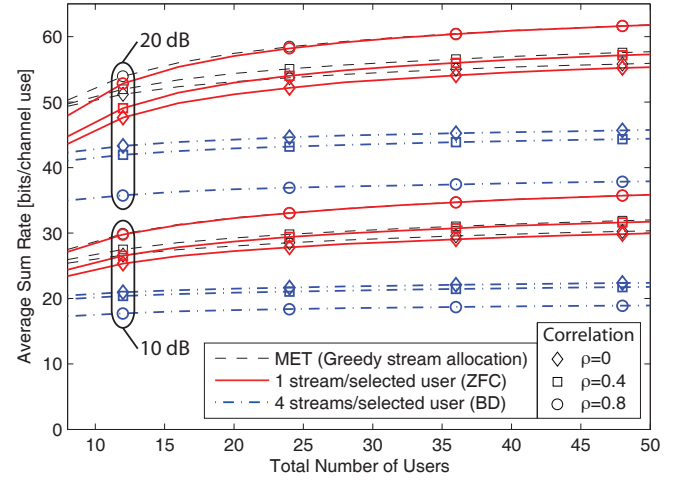


Fig. 4. The average achievable sum rate in a system with perfect CSI,  $N = 8$  transmit antennas,  $M = 4$  receive antennas, and the same average SNR among all users (10 or 20 dB). The performance with different strategies are shown as a function of the total number of users and for different correlation factors  $\rho$  among the receive antennas.

the sum rate and might give scheduling sets with fewer than  $N$  data streams. We consider a scenario with  $N = 8$  uncorrelated transmit antennas and  $M = 4$  receive antennas with correlation factor  $\rho \in \{0, 0.4, 0.8\}$ ; see [32] for another scenario. We compare ZFC (1 stream/user) and BD (4 streams/user) with multi-user eigenmode transmission (MET) from [12] where data streams are allocated greedily with zero inter-user interference and users can have different numbers of streams. We also simulated 2 streams/user, but it is not shown herein because the sum rate was always in between ZFC and BD.

Fig. 4 shows the average achievable sum rate as a function of the total number of users  $K$ . We consider the case when all users have the same average SNR (defined as  $P \frac{\mathbb{E}\{\|\mathbf{H}_k\|_F^2\}}{NM}$ ), either equal to 10 or 20 dB. Irrespective of the SNR, number of users, and receive-side correlation, ZFC outperforms BD. Thus, the scheduling-benefit of ZFC (from Theorem 2) dominates over the interference mitigation-benefit of BD (from Theorem 1)—even for spatially uncorrelated channels. As expected, the performance with ZFC improves with  $\rho$ , while correlation degrades the BD performance. MET has an advantages over ZFC since it can allocate different numbers of

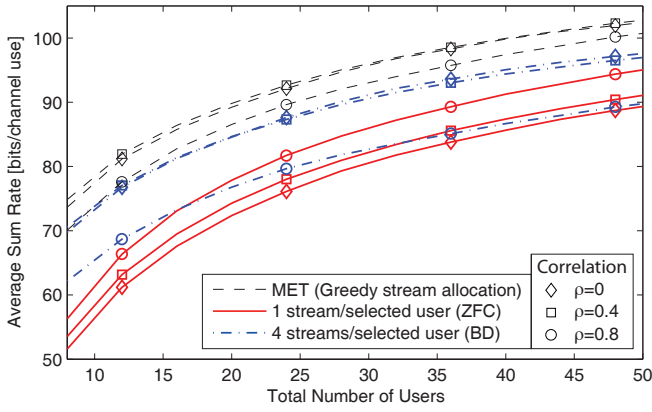


Fig. 5. The average achievable sum rate in a circular cell with perfect CSI,  $N = 8$  transmit antennas,  $M = 4$  receive antennas, and an SNR of 20 dB at the cell edge. The performance with different strategies are shown as a function of the total number of users and for different correlation factors  $\rho$  among the receive antennas.

streams to different users (based on how many singular values are strong in their channels), but this advantage is small and disappears asymptotically with the number of users; this was also observed in [12].

Next, we consider heterogeneous channel conditions by having uniformly distributed users in a circular cell with radius 250 m (minimal distance is 35 m), a path loss coefficient of 3.5, and log-normal shadow-fading with 8 dB in standard deviation. The average achievable sum rate is shown in Fig. 5 with an SNR of 20 dB at the cell edge.<sup>6</sup> The variation in path loss between users makes the results very different from the previous scenario in Fig. 4. At low receive-correlation, BD outperforms ZFC, but the difference reduces with  $K$ . ZFC is however better than BD at high correlation and many users. MET has a large advantage over the other strategies, explained by its flexible stream allocation. To comprehend the difference, the probability that a selected user is allocated a certain number of streams is shown in Fig. 6. We observe that spatial correlation reduces the number of streams per user, but the distance-dependence is even more significant; cell center users usually receive many streams while cell edge users only receive one or a few streams. This is natural since cell center users are more probable to have channel matrices with multiple relatively strong singular directions.

The conclusion is that ZFC is the method of choice in multi-user MIMO systems with perfect CSI and homogenous user conditions (since it performs very closely to the more complicated MET). On the other hand, MET and BD are better under heterogeneous user conditions. It is worth noting that the more streams allocated per user, the more channel dimensions need to be known at the base station. The next section will therefore study how practical CSI acquisition affects our results.

#### IV. COMPARISON OF BD AND ZFC WITH IMPERFECT CSI

In this section, we continue the comparison of BD and ZFC by introducing imperfect CSI, originating from either

<sup>6</sup>Such SNRs are reasonable in dense cellular systems and are necessary to compare BD and ZFC in regimes where these are supposed to work well.

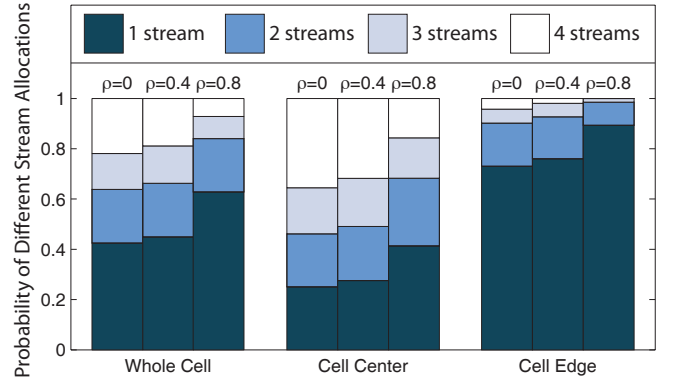


Fig. 6. The probability that a scheduled user is allocated a certain number of streams, assuming a circular cell with perfect CSI,  $N = 8$  transmit antennas,  $M = 4$  receive antennas,  $K = 20$  users, and an SNR of 20 dB at the cell edge. The whole cell has a radius of 250 meters, whereof users closer than 100 meters belong to the cell center and users further away than 200 meters belong to the cell edge.

quantized feedback in an FDD system or imperfect reverse-link estimation in a TDD system. The resources for channel acquisition are limited which has a major impact on both the number of channel dimensions that can be acquired per user and the accuracy of the acquired CSI. Theoretically, users can feed back different numbers of channel dimensions depending on some kind of long-term statistical CSI, but that would reduce the coverage (by favoring cell center users) and require a flexible system operation with additional control signaling. We therefore assume that the system acquires  $d$  dimensions/user from a randomly selected user set, where  $d \geq 1$  is fixed but depends on the intended precoding strategy. This assumption is relaxed in the numerical evaluation.

##### A. Comparison with Quantized CSI

In the FDD system operation of Fig. 2(a), each user selected for feedback conveys the  $d$ -dimensional subspace spanned by its effective channel  $\tilde{\mathbf{C}}_k^H \mathbf{H}_k$  using  $B$  bits. Similar to [7], [28], [33]–[35], we use a codebook  $\mathcal{C}_{N,d,B} = \{\mathbf{U}_1, \dots, \mathbf{U}_{2^B}\}$  with codewords  $\mathbf{U}_i \in \mathbb{C}^{N \times d}$  from the (complex) *Grassmannian manifold*  $\mathcal{G}_{N,d}$ ; that is, the set of all  $d$ -dimensional linear subspaces (passing through the origin) in an  $N$ -dimensional space. Each codeword forms an orthonormal basis, thus  $\mathbf{U}_i$  is a semi-unitary matrix satisfying  $\mathbf{U}_i^H \mathbf{U}_i = \mathbf{I}_d$ . User  $k$  selects the codeword that minimizes the chordal distance [36]:

$$\bar{\mathbf{H}}_k = \arg \min_{\mathbf{U} \in \mathcal{C}_{N,d,B}} \delta(\tilde{\mathbf{C}}_k^H \mathbf{H}_k, \mathbf{U}) \quad (17)$$

where  $\delta(\mathbf{B}, \mathbf{U}) = \sqrt{d - \text{tr}(\text{span}(\mathbf{B})^H \mathbf{U} \mathbf{U}^H \text{span}(\mathbf{B}))}$  and  $\text{span}(\cdot)$  gives a matrix containing an orthonormal basis of the row space. We assume error-free and delay-free feedback, but the conclusions of this section are expected to hold true also under feedback errors (cf. [23]).

There is a variety of ways to handle feedback errors (especially if the error structure is known), but a simple approach is to treat  $\bar{\mathbf{H}}_k$  as being the true channel [7] and calculate the precoding using a strategy developed for perfect CSI. This results in a lower bound on the performance and the



information rates with BD and ZFC becomes

$$g_k^{\text{BD-Q}}(P) = \log_2 \frac{\det(\mathbf{I}_M + \sum_{\ell \in \mathcal{S}^{\text{BD}}} \mathbf{H}_k \bar{\mathbf{W}}_\ell^{\text{BD}} \bar{\mathbf{Y}}_\ell \bar{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{H}_k^H)}{\det(\mathbf{I}_M + \sum_{\ell \in \mathcal{S}^{\text{BD}} \setminus \{k\}} \mathbf{H}_k \bar{\mathbf{W}}_\ell^{\text{BD}} \bar{\mathbf{Y}}_\ell \bar{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{H}_k^H)} \quad (18)$$

$$g_k^{\text{ZFC-Q}}(P) = \log_2 \left( 1 + \frac{\bar{p}_k |\mathbf{c}_k^H \mathbf{H}_k \bar{\mathbf{w}}_k^{\text{ZFC}}|^2}{1 + \sum_{\ell \in \mathcal{S}^{\text{ZFC}} \setminus \{k\}} \bar{p}_\ell |\mathbf{c}_k^H \mathbf{H}_k \bar{\mathbf{w}}_\ell^{\text{ZFC}}|^2} \right) \quad (19)$$

for users in the scheduling sets  $\mathcal{S}^{\text{BD}}$  and  $\mathcal{S}^{\text{ZFC}}$ , respectively.

Next, we quantify the performance loss for BD and ZFC compared with having perfect CSI. Random vector quantization (RVQ) is used for analytic convenience (as in [7], [8], [37], [38]), meaning that we average over codebooks with random codewords from the Grassmannian manifold. As any judicious codebook design is better than RVQ, the upper bounds on the performance loss that we will derive are valid for any reasonable codebook. The following theorem provides an upper bound on the performance loss under BD and extends results in [7] to include heterogeneous user conditions and spatial correlation.

*Theorem 3.* Assume that  $\frac{N}{M}$  users are scheduled randomly. The average rate loss with BD (using equal power allocation) for user  $k \in \mathcal{S}^{\text{BD}}$  due to RVQ is upper bounded as

$$\Delta_k^{\text{BD-Q}} = \mathbb{E}\{g_k^{\text{BD}}(P) - g_k^{\text{BD-Q}}(P)\} \leq \log_2 \det \left( \mathbf{I}_M + \frac{P}{M} D^{\text{BD}} \mathbf{R}_{R,k} \right) \quad (20)$$

where the average quantization distortion is

$$D^{\text{BD}} = \mathbb{E}\{\delta^2(\mathbf{H}_k, \bar{\mathbf{H}}_k)\} \approx \frac{\Gamma\left(\frac{1}{M(N-M)}\right)}{M(N-M)} \left( \frac{2^B}{(M(N-M))!} \prod_{i=1}^M \frac{(N-i)!}{(M-i)!} \right)^{-\frac{1}{M(N-M)}}. \quad (21)$$

*Proof:* The proof is given in Appendix D. ■

This theorem will be compared with the corresponding result for ZFC, but before stating that result we discuss how to select the (preliminary) receive combiner  $\tilde{\mathbf{c}}_k$ . There are primarily two factors to consider when selecting  $\tilde{\mathbf{c}}_k$ : the gain of the effective channel  $\|\tilde{\mathbf{c}}_k^H \mathbf{H}_k\|_2^2$  and the quantization distortion. The results of [39], [40] indicate that the top priority in multi-user MIMO systems is to achieve small quantization errors, because it is a prerequisite for low inter-user interference. The error can be minimized by the *quantization-based combining* (QBC) approach in [8], where the codeword and receive combiner are selected jointly as

$$(\tilde{\mathbf{c}}_k^{\text{QBC}}, \bar{\mathbf{h}}_k) = \arg \max_{\substack{\mathbf{c}: \|\mathbf{c}\|_2=1 \\ \mathbf{u} \in \mathcal{C}_{N,1,B}}} \delta(\mathbf{H}_k^H \mathbf{c}, \mathbf{u}). \quad (22)$$

The *maximum expected SINR combiner* (MESIC) in [9] achieves better practical performance by balancing effective channel gain and quantization distortion, but is asymptotically equal to QBC at high SNR. Since this is the regime of main interest herein, we will exploit the analytic simplicity of QBC. Observe that QBC and MESIC are only used for improved

feedback accuracy; the MMSE combiner in Remark 1 is used to maximize the performance during transmission (this was *not* done in the original QBC framework of [8]).

The following theorem provides an upper bound on the performance loss under ZFC and extends results in [8] to include heterogeneous user conditions and spatial correlation.

*Theorem 4.* Assume that  $\mathbf{R}_{R,k}$  has eigenvalues  $\lambda_{k,M} > \dots > \lambda_{k,1} > 0$  and that  $N$  users are selected randomly. The average rate loss for ZFC (using equal power allocation and the same  $\tilde{\mathbf{c}}_k^{\text{QBC}}$ ) due to RVQ is upper bounded as

$$\Delta_k^{\text{ZFC-Q}} = \mathbb{E}\{g_k^{\text{ZFC}}(P) - g_k^{\text{ZFC-Q}}(P)\} \leq \log_2 \left( 1 + \frac{P}{N} D^{\text{QBC}} G_k \right) \quad (23)$$

where the average quantization distortion is

$$D^{\text{QBC}} = \mathbb{E}\{\delta^2(\mathbf{h}_k, \bar{\mathbf{h}}_k)\} \approx 2^{-\frac{B}{N-M}} \left( \frac{N-1}{M-1} \right)^{-\frac{1}{N-M}} \quad (24)$$

and the average channel gain with QBC is (where  $\mu_n = \frac{1}{\lambda_{k,n}}$ )

$$G_k = \sum_{m=1}^{M-1} \sum_{n=1}^m \sum_{t=m+1}^M \frac{(N-M+1)A_{m,n,t}}{(\mu_n - \mu_t) \prod_{\substack{i=1 \\ i \neq n}}^m (\mu_n - \mu_i) \prod_{\substack{j=m+1 \\ j \neq t}}^M (\mu_j - \mu_t)},$$

$$A_{m,n,t} = \log_e \left( \frac{\mu_{m+1}^{m-1}}{\mu_m} \right) \frac{\mu_n^{m-1}}{(-\mu_t)^{2+m-M}} (m\mu_t + \mu_n(M-m-1))$$

$$+ \sum_{s=0}^{m-1} \sum_{r=0}^{M-m-1} \binom{m}{s} \binom{M-m-1}{r} \frac{\mu_n^{m-s} (-\mu_t)^{M-m-1-r}}{(-1)^s (1+s+r)}$$

$$\times \left( \frac{m-s}{\mu_n} + \frac{M-m-1-r}{\mu_t} \right) (\mu_m^{r+s} - \mu_{m+1}^{r+s}). \quad (25)$$

*Proof:* The proof is given in Appendix E. ■

The rate loss expressions in Theorem 3 and Theorem 4 for BD and ZFC, respectively, indicate the joint impact of spatial correlation (at the receiver) and CSI quantization on the performance. The main observation is that spatial correlation only has a marginal effect on the feedback accuracy; the expressions have a similar structure as for uncorrelated channels and the same scaling in the number of feedback bits is necessary to achieve the maximal multiplexing gain [7], [8].

*Corollary 1.* To achieve the maximal multiplexing gain with BD or ZFC under quantized CSI and arbitrary receive correlation, it is sufficient to scale the total number of CDI feedback bits for the scheduled users as

$$B_{\text{total}} \approx N(N-M) \log_2(P) + \mathcal{O}(1). \quad (26)$$

While this corollary only provides a sufficient condition, we can expect the scaling law in (26) to also be necessary.<sup>7</sup> In any case, the scaling law in (26) is easily satisfied by allocating (approximately)  $N-M$  channel uses for CSI feedback, since typically the uplink sum rate also behaves as  $N \log_2(P) + \mathcal{O}(1)$  in the high-SNR regime [23].

<sup>7</sup>The necessary scaling can be proved for ZFC with QBC using a technique from [41, Theorem 4], while simulations in [7] show that quantized ZFC and BD have the same scaling in the necessary number of bits.

Observe that this result is based on random user selection, while additional feedback of gain information is necessary to achieve multi-user diversity or short-term rate adaptation (cf. [42]). As BD requires  $M$  times more bits per user, ZFC can typically achieve feedback from  $M$  times more users. We therefore expect ZFC to further strengthen its advantage at finding near-orthogonal users (indicated in Theorem 2 under perfect CSI). In addition, spatial correlation at the transmitter-side (and other factors that make the channel matrices ill-conditioned) will inflict larger performance losses on BD than ZFC, just as in the case of perfect CSI.

### B. Comparison under Estimated CSI

Next, we assume that the base station acquires CSI through imperfect CSI estimation. The primary focus will be on TDD systems, where channel estimates are obtained through training signaling in the uplink (assuming perfect channel reciprocity). It is worth noting that this approach is similar to having analog CSI feedback in FDD systems, where the unquantized channel coefficients are sent on an uplink subcarrier [7], [23].<sup>8</sup>

The reciprocal uplink counterpart to the system model in (2) is

$$\tilde{\mathbf{y}}_k = \mathbf{H}_k^T \tilde{\mathbf{x}}_k + \tilde{\mathbf{n}}_k \quad (27)$$

where  $\tilde{\mathbf{y}}_k \in \mathbb{C}^{N \times 1}$  is the received uplink signal,  $\tilde{\mathbf{x}}_k \in \mathbb{C}^{M \times 1}$  is the transmitted uplink signal, and  $\tilde{\mathbf{n}}_k \sim \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$  is the noise vector.<sup>9</sup> To estimate  $\tilde{\mathbf{C}}_k^H \mathbf{H}_k \in \mathbb{C}^{d \times N}$ , user  $k$  sends  $\tilde{\mathbf{C}}_k^* \mathbf{T}_k$  over  $d$  uplink channel uses, for some known training matrix  $\mathbf{T}_k \in \mathbb{C}^{d \times d}$  and where  $(\cdot)^*$  denotes the complex conjugate. Assuming perfect statistical CSI, the MMSE estimate  $\hat{\mathbf{H}}_k$  of  $\tilde{\mathbf{C}}_k^H \mathbf{H}_k$  and the corresponding error covariance matrix  $\mathbf{E}_k$  are [22]

$$\begin{aligned} \text{vec}(\hat{\mathbf{H}}_k^T) &= \frac{1}{\sigma^2} \mathbf{E}_k \tilde{\mathbf{T}}_k^H \text{vec}(\mathbf{Y}_k), \\ \mathbf{E}_k &= \left( (\tilde{\mathbf{C}}_k^H \mathbf{R}_{R,k} \tilde{\mathbf{C}}_k \otimes \mathbf{I}_N)^{-1} + \frac{\tilde{\mathbf{T}}_k^H \tilde{\mathbf{T}}_k}{\sigma^2} \right)^{-1} \end{aligned} \quad (28)$$

where  $\tilde{\mathbf{T}}_k = (\mathbf{T}_k^T \otimes \mathbf{I}_N)$  and  $\mathbf{Y}_k$  is the received signal from training signaling. The training matrix  $\mathbf{T}_k$  has a total training power/SNR constraint  $\text{tr}(\mathbf{T}_k^H \mathbf{T}_k) = \Psi$ .

As under quantized CSI, we calculate the precoding by treating  $\hat{\mathbf{H}}_k$  as the true channel. This results in a lower bound on the performance and the information rates with BD and

<sup>8</sup>Digital/quantized feedback might be beneficial over analog/unquantized feedback when there is plenty of resources for channel estimation [23]. But if very accurate CSI is required, Corollary 1 shows that the quantization codebooks grow very large and thus the search for the best codeword might be computationally infeasible.

<sup>9</sup>The downlink noise vector was normalized towards the channel matrix in the system model of (2). To account for a different noise level at the base station,  $\sigma^2$  is the (relative) uplink noise variance.

ZFC becomes

$$g_k^{\text{BD-EST}}(P) = \log_2 \frac{\det\left(\mathbf{I}_M + \sum_{\ell \in \mathcal{S}^{\text{BD}}} \mathbf{H}_k \hat{\mathbf{W}}_\ell^{\text{BD}} \hat{\mathbf{Y}}_\ell \hat{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{H}_k^H\right)}{\det\left(\mathbf{I}_M + \sum_{\ell \in \mathcal{S}^{\text{BD}} \setminus \{k\}} \mathbf{H}_k \hat{\mathbf{W}}_\ell^{\text{BD}} \hat{\mathbf{Y}}_\ell \hat{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{H}_k^H\right)} \quad (29)$$

$$g_k^{\text{ZFC-EST}}(P) = \log_2 \left( 1 + \frac{\hat{p}_k |\mathbf{c}_k^H \mathbf{H}_k \hat{\mathbf{w}}_k^{\text{ZFC}}|^2}{1 + \sum_{\ell \in \mathcal{S}^{\text{ZFC}} \setminus \{k\}} \hat{p}_\ell |\mathbf{c}_k^H \mathbf{H}_k \hat{\mathbf{w}}_\ell^{\text{ZFC}}|^2} \right) \quad (30)$$

for users in the scheduling sets  $\mathcal{S}^{\text{BD}}$  and  $\mathcal{S}^{\text{ZFC}}$ , respectively. The following theorem provides an upper bound on the performance loss under BD due to imperfect CSI estimation.

*Theorem 5.* Assume that  $\frac{N}{M}$  users are scheduled randomly under BD. The average rate loss for user  $k \in \mathcal{S}^{\text{BD}}$  (using equal power allocation) due to CSI estimation is upper bounded as

$$\begin{aligned} \Delta^{\text{BD}} &= \mathbb{E}\{g_k^{\text{BD}}(P) - g_k^{\text{BD-EST}}(P)\} \\ &\leq \log_2 \det \left( \mathbf{I}_M + \frac{P(N-M)}{N} \left( \mathbf{R}_{R,k}^{-T} + \frac{\mathbf{T}_k^H \mathbf{T}_k}{\sigma^2} \right)^{-1} \right). \end{aligned} \quad (31)$$

*Proof:* The proof is given in Appendix F. ■

This theorem will be compared with the corresponding result for ZFC, but before stating that theorem we need to consider the impact of having MRC as the receive combiner  $\tilde{\mathbf{c}}_k$ . ZFC is similar to applying BD to the effective channels  $\mathbf{h}_k^H = \tilde{\mathbf{c}}_k^H \mathbf{H}_k$ , but an important difference is that the effective channels are not Rayleigh fading because  $\tilde{\mathbf{c}}_k$  depends on the current channel realization. The expression in (28) will therefore not give the MMSE estimate, but fortunately the linear MMSE (LMMSE) estimator from a similar expression to (28) if we know the first two moments of  $\mathbf{h}_k$  [22].

*Lemma 1.* Assume that  $\mathbf{R}_R$  has eigenvalues  $\lambda_M > \dots > \lambda_1 > 0$ , where the user indices were dropped for convenience. If  $\tilde{\mathbf{c}}$  is the dominating left singular vector of  $\mathbf{H}$ , it holds that

- the direction  $\frac{\mathbf{h}}{\|\mathbf{h}\|_2}$  of  $\mathbf{h} = \tilde{\mathbf{c}}^H \mathbf{H}$  is isotropically distributed on the unit sphere;
- the gain  $\|\mathbf{h}\|_2^2$  is independent of the direction and

$$\begin{aligned} \mathbb{E}\{\|\mathbf{h}\|_2^2\} &= \sum_{m=1}^M \sum_{\zeta \in \mathcal{A}_M} \frac{\prod_{\ell=1}^M \lambda_{\zeta_\ell}^{N-\ell+1} \prod_{\ell=N-M+1}^N (\ell-1)!}{(-1)^{\text{per}(\zeta)+m+1} \det(\Delta)} \\ &\times \sum_{\beta \in \mathcal{B}_{m,M}} \sum_{\ell=0}^{K_m(\beta)} \sum_{\tilde{k} \in \tilde{\Omega}_\ell^{(m)}} \frac{\ell!}{\tilde{k}_1! \dots \tilde{k}_m!} \frac{\left(\sum_{i=1}^m \lambda_{\zeta_{\beta_i}}^{-1}\right)^{-(\ell+1)}}{\prod_{i=1}^m \lambda_{\zeta_{\beta_i}}^{\tilde{k}_i}} \end{aligned} \quad (32)$$

where the  $ij$ th element of  $\Delta \in \mathbb{R}^{M \times M}$  is given by

$$[\Delta]_{ij} = \lambda_j^{N-i+1} (N-i)!. \quad (33)$$

In (32), the set of all permutations of  $\{1, \dots, M\}$  is denoted  $\mathcal{A}_M$ . The sign of a given permutation  $\zeta = \{\zeta_1, \dots, \zeta_M\} \in \mathcal{A}_M$  is denoted  $(-1)^{\text{per}(\zeta)}$ , where  $\text{per}(\cdot)$  is the number of

inversions<sup>10</sup> in the permuted sequence. Next,  $\mathcal{B}_{l,M}$  is the collection of all subsets of  $\mathcal{A}_M$  with cardinality  $l$  and increasing elements (i.e.,  $\beta_1 < \dots < \beta_l$  for  $\beta = \{\beta_1, \dots, \beta_l\} \in \mathcal{B}_{l,M}$ ). The upper bound in the summation over  $\ell$  is  $K_l(\beta) = \sum_{i=1}^l (N - \beta_i)$ . Finally,  $\tilde{\Omega}_\ell^{(l)}$  is the set of all  $l$ -length partitions  $\{\tilde{k}_1, \dots, \tilde{k}_l\}$  of  $\ell$  (i.e.,  $\sum_{i=1}^l \tilde{k}_i = \ell$ ) that satisfy  $0 \leq \tilde{k}_i \leq N - \beta_i$ :

$$\tilde{\Omega}_\ell^{(l)} = \left\{ \{\tilde{k}_1, \dots, \tilde{k}_l\} : \sum_{j=1}^l \tilde{k}_j = \ell, 0 \leq \tilde{k}_j \leq N - \beta_j \forall j \right\}. \quad (34)$$

*Proof:* The proof is given in Appendix G. ■

The following theorem provides an upper bound on the performance loss under ZFC due to imperfect CSI estimation.

**Theorem 6.** Assume that  $N$  users are scheduled randomly under ZFC and that MRC is applied. The average rate loss for user  $k \in \mathcal{S}^{\text{ZFC}}$  due to CSI estimation is upper bounded as

$$\begin{aligned} \Delta^{\text{ZFC-EST}} &= \mathbb{E}\{g_k^{\text{ZFC}}(P) - g_k^{\text{ZFC-EST}}(P)\} \\ &\leq \log_2 \left( 1 + \frac{P(N-1)}{N} \frac{1}{\mathbb{E}\{\|\mathbf{h}_k\|_2^2\}^{-1} + \frac{\Psi}{\sigma^2}} \right) \end{aligned} \quad (35)$$

where  $\mathbb{E}\{\|\mathbf{h}_k\|_2^2\}$  is given in (32).

*Proof:* The proof is given in Appendix H. ■

The rate loss expressions in Theorem 5 and Theorem 6 indicate the joint impact of spatial correlation and imperfect channel estimation on the performance of BD and ZFC, respectively. BD is slightly more resilient to CSI uncertainty, since the BD expression contains  $(N - M)$  where the ZFC expression has  $(N - 1)$ . But observe that the performance losses are calculated against the same precoding strategy with perfect CSI; we know from Section III that ZFC and BD have different preferable user conditions, making it hard to analytically conclude which strategy to use under imperfect CSI estimation. However, the important result is the following extension of [23] to spatially correlated scenarios with  $M \geq 1$ .

**Corollary 2.** To achieve the maximal multiplexing gain with BD or ZFC under imperfect CSI estimation and arbitrary receive correlation, it is necessary and sufficient to scale the training power  $\Psi$  as

$$\frac{P}{\Psi} \rightarrow \text{constant} < \infty \quad \text{when } P \rightarrow \infty. \quad (36)$$

*Proof:* The proof is given in Appendix I. ■

This corollary says that the training power/SNR should increase linearly with the transmit power/SNR to achieve the optimal sum rate scaling. This is, for example, satisfied by setting the total training power to  $\Psi = P$  under ZFC and  $\Psi = MP$  under BD, which corresponds to the reasonable assumption of having the same average SNR in the downlink and in the uplink.<sup>11</sup> The demands for higher CSI accuracy with increasing SNR is therefore automatically fulfilled by

<sup>10</sup>An inversion in a sequence is a pair of numbers that is in incorrect order (i.e., not in ascending order).

<sup>11</sup>Battery-powered user devices might operate at lower power budget than the base station, but Corollary 2 is satisfied as long as  $P$  and  $\Psi$  exhibit the same scaling. In practical scenarios, the path loss is the main source of SNR variations and affects the downlink and uplink equally.

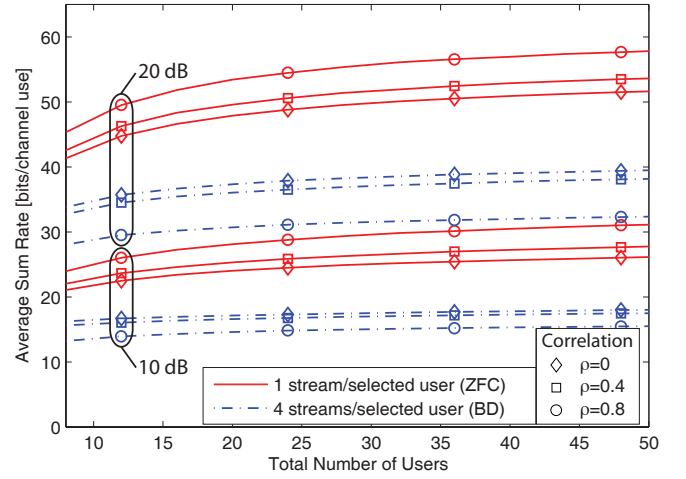


Fig. 7. The average achievable sum rate in a system with CSI estimation errors,  $N = 8$  transmit antennas,  $M = 4$  receive antennas, and the same average SNR among all users (10 or 20 dB). The performance with different strategies are shown as a function of the total number of users and for different correlation factors  $\rho$  among the receive antennas.

the reduced estimation errors. Observe that one uplink channel use is consumed per user antenna dimension that is estimated, thus creating a practical bound on how many user channels that can be estimated in block fading systems [23]. As ZFC only has one effective antenna per user, it can accommodate  $M$  times more users than BD on the same estimation overhead and thereby exploit multi-user diversity to a larger extent.

## V. NUMERICAL ILLUSTRATIONS UNDER IMPERFECT CSI

This section consists of two parts. First, the numerical illustrations in Section III-B are continued under imperfect CSI estimation. Then, we analyze the performance behavior under quantized CSI.

### A. Continuation of Section III-B under Estimated CSI

We continue the simulations in Section III-B by introducing imperfect CSI estimation. We use the MSE-minimizing training matrices from [22, Theorem 1] and training power  $\Psi = Pd$  (for estimation of  $d$  dimensions/user). The CBSUS algorithm in [31] is modified<sup>12</sup> to include the average interference (due to CSI estimation errors) in the scheduling.

The average achievable sum rate is shown in Fig. 7 as a function of the number of users that we obtain CSI estimates for using ZFC (while BD only obtains channel estimates for  $\frac{1}{M}$  of them). All users have the same average SNR of either 10 or 20 dB. The performance loss compared with having perfect CSI is 10-20% (see Fig. 4), but the conclusion is otherwise the same and even clearer than before: ZFC outperforms BD in terms of performance with few users, in handling spatial correlation, and in exploiting multi-user diversity.

In case of a circular cell (see Section III-B for details), the average achievable sum rate is shown in Fig. 8. Recall from Fig. 5 that BD was often better than ZFC in this

<sup>12</sup>Estimation errors contribute an average interference of  $P(|\mathcal{S}| - 1)/|\mathcal{S}|\mathbf{E}_{\text{est}}$ , where  $\mathbf{E}_{\text{est}} = (\mathbf{R}_{R,k}^{-T} + \mathbf{T}_k^H \mathbf{T}_k / \sigma^2)^{-1}$  for BD and  $\mathbf{E}_{\text{est}} = (1/\mathbb{E}\{\|\mathbf{h}_k\|_2^2\} + \Psi/\sigma^2)^{-1}$  for ZFC.

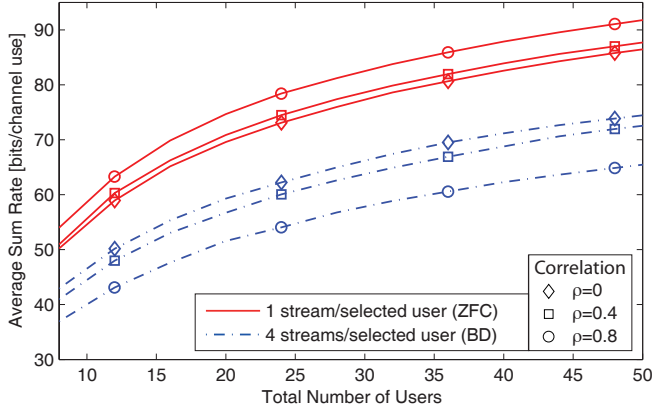


Fig. 8. The average achievable sum rate in a circular cell with CSI estimation errors,  $N = 8$  transmit antennas,  $M = 4$  receive antennas, and an SNR of 20 dB at the cell edge. The performance with different strategies are shown as a function of the total number of users and for different correlation factors  $\rho$  among the receive antennas.

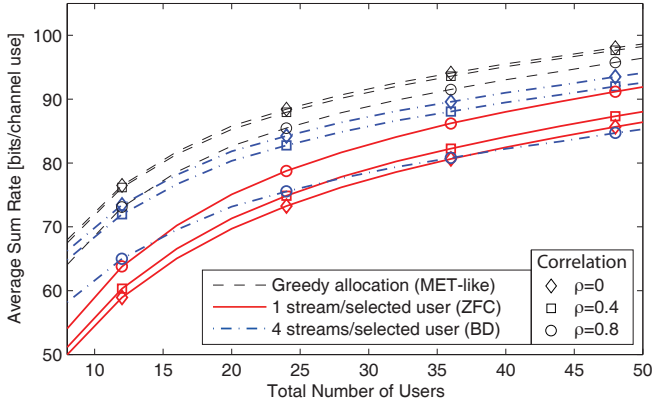


Fig. 9. The average achievable sum rate in a circular cell with CSI estimation errors,  $N = 8$  transmit antennas,  $M = 4$  receive antennas, and an SNR of 20 dB at the cell edge. The performance is shown as a function of the total number of users  $K$ , and CSI is only acquired for the users with strongest long-term statistics.

scenario under perfect CSI, but the case is completely different under imperfect CSI; ZFC outperforms the other strategies when the limited resources for CSI acquisition are taken into account. This means that the ZFC benefit of easily finding near-orthogonal users (among  $M$  times more users than with BD) dominates the BD benefit of multi-stream multiplexing (preferably to cell center users). We also tested a MET-like strategy with greedy stream allocation (we took the optimum among feeding back 1, 2 or 4 channel dimensions per active user), but it was always identical to ZFC (in both Fig. 4 and Fig. 5)—this further confirms our conclusion.

The users selected for feedback were chosen randomly (e.g., in a round-robin fashion) in Figs. 7 and 8, but could theoretically be based on some kind of long-term statistical CSI. This could for instance mean that ZFC acquires one dimension from each of the  $K$  users, while BD acquires  $M$  dimensions from the  $\frac{K}{M}$  users with the strongest long-term statistics  $\text{tr}(\mathbf{R}_{T,k})\text{tr}(\mathbf{R}_{R,k})$ . The greedy stream allocation strategy MET in [12] can be generalized to this scenario by finding the  $K$  strongest statistical eigendirections among the users and acquire CSI for an equivalent number of dimensions per user. Under these assumptions, the average achievable sum

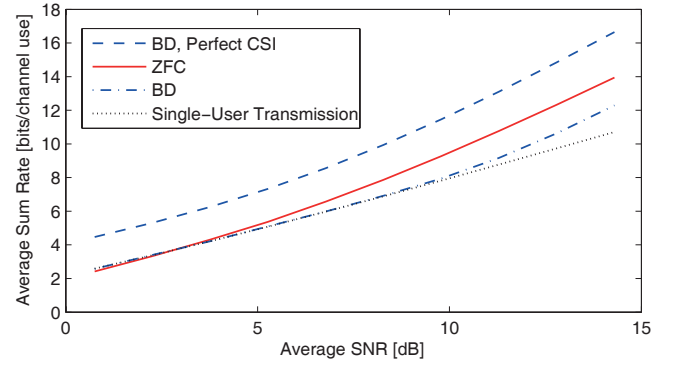


Fig. 10. The average achievable sum rate with BD and ZFC, quantized CSI feedback,  $N = 4$  transmit antennas,  $M = 2$  receive antennas, uncorrelated channels, and varying SNR. The number of feedback bits is scaled with the transmit power according to Corollary 1 and [7, Eq. (17)].

rate for the circular cell is shown in Fig. 9. The performance behavior is quite similar to the case with perfect CSI in Fig. 5; BD is better than ZFC, except at high correlation, and there is a large gain from greedy stream allocation. However, we stress that this scenario is unrealistic as CSI is only acquired for cell center users, thus reducing the coverage and destroying user fairness as cell edge users are not even considered when their channels are relatively strong.

### B. Observations under Quantized CSI

Next, we consider quantized CSI and let the number of feedback bits (per channel dimension) be scaled as  $(N - M) \log_2(P) - \text{constant}$ , where the constant is selected as in [7, Eq. (17)] to maintain a 3 dB gap between BD with perfect and quantized CSI. We consider  $N = 4$  transmit antennas,  $M = 2$  receive antennas, and RVQ. We also modify<sup>13</sup> the CBSUS algorithm in [31] to include the average interference due to quantization.

First, we compare BD (having either quantized or perfect CSI) with quantized ZFC using MESC-MMSE combining [9] and with single-user SVD-based transmission (to a randomly selected user). The quantized effective channels are obtained from 8 users under ZFC, while the entire channels are quantized for 4 users under BD. The average achievable sum rate is shown in Fig. 10 as a function of the average SNR. At low SNRs, quantized BD only selects one user and performs similar to single-user transmission. As two data streams are transmitted to the selected user, both strategies are slightly better than ZFC in this regime. But quantized ZFC quickly improves with SNR and becomes the method of choice at practical SNRs. The simulation was stopped at  $P = 14.3$  dB where BD requires feedback of 22 bits per user, meaning that the best codeword is selected in a codebook with over a million entries.<sup>14</sup> BD is therefore suboptimal both in terms of sum rate and computational complexity.

<sup>13</sup>Quantization errors contribute an average interference of  $P(|S| - 1)/|S|\mathbf{E}_{\text{quant}}$ , where  $\mathbf{E}_{\text{quant}} = N/(M(N - M))D^{\text{BD}}\mathbf{R}_{R,k}$  for BD and  $\mathbf{E}_{\text{quant}} = D^{\text{QBC}}G/(N - 1)$  for ZFC. When calculating  $D^{\text{BD}}$  and  $D^{\text{QBC}}$ , BD uses  $M$  times more feedback bits per user than ZFC.

<sup>14</sup>An approach to emulate RVQ for very large random codebooks was proposed in [7], but this does not change the fact that the quantization complexity becomes infeasible much faster under BD than under ZFC.

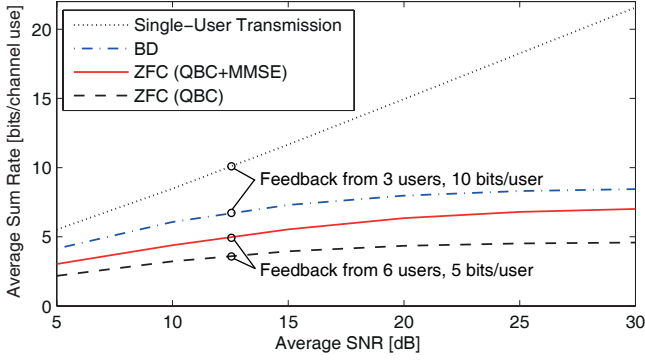


Fig. 11. Comparison of single-user transmission, BD, and different forms of ZFC under quantized CSI feedback. The scenario is the same as in [7, Fig. 6], where the superior single-user strategy was not included.

This observation stands in contrast to the numerical results in [7], where BD clearly beats ZFC under quantized CSI. To explain the difference, we repeat the simulation in [7, Fig. 6] with  $N = 6$  transmit antennas and  $M = 2$  receive antennas. In this simulation, the RVQ codebooks contain 10 bits/user under BD and 5 bits/user under ZFC. The achievable sum rate is shown in Fig. 11 for the quantized BD approach in [7] and the ZFC-QBC approach in [8]. We have also included: 1) an improved version of ZFC-QBC where the MMSE receive combiner is applied during transmission; and 2) single-user SVD-based transmission to a randomly selected user. Our simulation confirms that BD is better than ZFC in this scenario, but the difference becomes much smaller when the MMSE combiner is applied. However, none of these strategies should be used in this scenario since single-user transmission is vastly superior. The explanation is that the number of feedback bits is fixed at a number that only satisfies/exceeds the feedback scaling law in (26) and [7, Eq. (17)] at low SNRs (cf. [7, Fig. 2]), while the strict interference mitigation in BD and ZFC is only practically meaningful at high SNR. The observation in [7] is thus misleading and does not contradict the superiority of ZFC under proper feedback loads.

Conclusions from the mathematical and numerical analysis are summarized in the next section.

## VI. CONCLUSION

This paper analyzed how to divide data streams among users in a downlink system with many multi-antenna users; should few users be allocated many streams, or many users be allocated few streams? New and generalized analytic results were obtained to study this tradeoff under spatial correlation, user selection, heterogeneous user channel conditions, and practical CSI acquisition.

The main conclusion is that sending one stream per selected user and exploiting receive combining is the best choice under realistic conditions. This is good news as it reduces the hardware requirements at the users, compared with multi-stream multiplexing, and enables computationally efficient resource allocation as in [43]. The result is explained by a stronger resilience towards spatial correlation and larger benefit from user selection. To arrive at alternative conclusions, one has to consider a scenario with heterogeneous user conditions with

either perfect CSI (unrealistic) or where CSI is only acquired for the strongest users (destroys coverage and fairness). It should however be noted that if only very inaccurate CSI can be acquired, then inter-user interference will limit performance thus making single-user transmission advantageous.

## APPENDIX A COLLECTION OF LEMMAS

This appendix contains two lemmas that are essential for proving the theorems of this paper. The first result shows how spatial correlation at the receiver affects the channel directions.

**Lemma 2.** Let  $\mathbf{A} \succ \mathbf{0}_M$  be any Hermitian positive-definite matrix and let  $\tilde{\mathbf{H}} \in \mathbb{C}^{M \times N}$  be an arbitrary matrix. Then,  $\text{span}(\tilde{\mathbf{H}}) = \text{span}(\mathbf{A}\tilde{\mathbf{H}})$ , where  $\text{span}(\cdot)$  denotes the row space.

*Proof:* Let  $\mathbf{A} = \mathbf{U}_A \mathbf{\Lambda}_A \mathbf{U}_A^H$  be an eigen decomposition of  $\mathbf{A}$ . The lemma follows by observing that  $\mathbf{U}_A$  only rotates the basis vectors of the row space and  $\mathbf{\Lambda}_A$  scales the rows without affecting their span. ■

The second result generalizes the bounding of performance loss under imperfect CSI in [7], [8].

**Lemma 3.** Let  $\mathbf{W}_k, \tilde{\mathbf{W}}_k$  be isotropically distributed on the Grassmannian manifold  $\mathcal{G}_{N,d_k}$  and independent of  $\mathbf{H}_k$ , then

$$\begin{aligned} & \mathbb{E} \left\{ \log_2 \det \left( \mathbf{I}_{d_k} + \frac{P}{N} \tilde{\mathbf{C}}_k^H \mathbf{H}_k \mathbf{W}_k \mathbf{W}_k^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k \right) \right\} \\ & - \mathbb{E} \left\{ \log_2 \frac{\det \left( \mathbf{I}_{d_k} + \frac{P}{N} \sum_{\ell} \tilde{\mathbf{C}}_k^H \mathbf{H}_k \tilde{\mathbf{W}}_{\ell} \tilde{\mathbf{W}}_{\ell}^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k \right)}{\det \left( \mathbf{I}_{d_k} + \frac{P}{N} \sum_{\ell \neq k} \tilde{\mathbf{C}}_k^H \mathbf{H}_k \tilde{\mathbf{W}}_{\ell} \tilde{\mathbf{W}}_{\ell}^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k \right)} \right\} \\ & \leq \log_2 \det \left( \mathbf{I}_{d_k} + \frac{P}{N} \sum_{\ell \neq k} \mathbb{E} \{ \tilde{\mathbf{C}}_k^H \mathbf{H}_k \tilde{\mathbf{W}}_{\ell} \tilde{\mathbf{W}}_{\ell}^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k \} \right). \end{aligned} \quad (37)$$

*Proof:* This lemma follows from two inequalities. First,

$$\begin{aligned} & \mathbb{E} \left\{ \log_2 \det \left( \mathbf{I}_{d_k} + \frac{P}{N} \tilde{\mathbf{C}}_k^H \mathbf{H}_k \mathbf{W}_k \mathbf{W}_k^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k \right) \right\} \\ & - \mathbb{E} \left\{ \log_2 \det \left( \mathbf{I}_{d_k} + \frac{P}{N} \sum_{\ell} \tilde{\mathbf{C}}_k^H \mathbf{H}_k \tilde{\mathbf{W}}_{\ell} \tilde{\mathbf{W}}_{\ell}^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k \right) \right\} \leq 0 \end{aligned} \quad (38)$$

since  $\tilde{\mathbf{C}}_k^H \mathbf{H}_k \mathbf{W}_k \mathbf{W}_k^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k$  and  $\tilde{\mathbf{C}}_k^H \mathbf{H}_k \tilde{\mathbf{W}}_{\ell} \tilde{\mathbf{W}}_{\ell}^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k$  have the same distribution, and the second term contains additional positive semi-definite matrices. Second, applying Jensen's inequality on the concave function  $\log_2 \det(\cdot)$  gives

$$\begin{aligned} & \mathbb{E} \left\{ \log_2 \det \left( \mathbf{I}_{d_k} + \frac{P}{N} \sum_{\ell \neq k} \tilde{\mathbf{C}}_k^H \mathbf{H}_k \tilde{\mathbf{W}}_{\ell} \tilde{\mathbf{W}}_{\ell}^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k \right) \right\} \\ & \leq \log_2 \det \left( \mathbf{I}_{d_k} + \frac{P}{N} \sum_{\ell \neq k} \mathbb{E} \{ \tilde{\mathbf{C}}_k^H \mathbf{H}_k \tilde{\mathbf{W}}_{\ell} \tilde{\mathbf{W}}_{\ell}^H \mathbf{H}_k^H \tilde{\mathbf{C}}_k \} \right). \end{aligned} \quad (39)$$

The lemma follows from combining (38) and (39). ■



## APPENDIX B

*Proof of Theorem 1:* Using (10), the expected asymptotic difference is

$$\bar{\beta}_{\text{BD-ZFC}} = \mathbb{E} \left\{ \log_2 \left( \frac{\prod_{k \in \mathcal{S}^{\text{BD}}} \det(\mathbf{H}_k \mathbf{W}_k^{\text{BD}} \mathbf{W}_k^{\text{BD},H} \mathbf{H}_k^H)}{\prod_{\ell \in \mathcal{S}^{\text{ZFC}}} |\tilde{\mathbf{c}}_\ell^H \mathbf{H}_\ell \mathbf{w}_\ell^{\text{ZFC}}|^2} \right) \right\}. \quad (40)$$

The direction  $\frac{\tilde{\mathbf{c}}_\ell^H \mathbf{H}_\ell}{\|\tilde{\mathbf{c}}_\ell^H \mathbf{H}_\ell\|_2}$  is isotropically distributed on the unit sphere, according to Lemma 1. This enables us to rewrite (40) as

$$\begin{aligned} \bar{\beta}_{\text{BD-ZFC}} = & \mathbb{E} \left\{ \log_2 \left( \frac{\prod_{k \in \mathcal{S}^{\text{BD}}} \det(\tilde{\mathbf{H}}_k \mathbf{W}_k^{\text{BD}} \mathbf{W}_k^{\text{BD},H} \tilde{\mathbf{H}}_k^H)}{\prod_{k \in \mathcal{S}^{\text{ZFC}}} |\tilde{\mathbf{h}}_k^H \mathbf{w}_k^{\text{ZFC}}|^2} \right) \right\} \\ & + \sum_{k \in \mathcal{S}^{\text{BD}}} \log_2 \det(\mathbf{R}_{R,k}) - \sum_{\ell \in \mathcal{S}^{\text{ZFC}}} \mathbb{E}\{z_\ell\} \end{aligned} \quad (41)$$

where  $z_\ell = \mathbb{E} \left\{ \log_2 \left( \frac{\|\tilde{\mathbf{c}}_\ell^H \mathbf{H}_\ell\|_2^2}{\|\tilde{\mathbf{h}}_\ell\|_2^2} \right) \right\}$  and  $\tilde{\mathbf{h}}_\ell \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ .<sup>15</sup> The first term in (41) equals the first term in (11) by applying [27, Theorem 3]. The cited theorem was stated for uncorrelated channels, but can be applied in our scenario since  $\mathbf{W}_k^{\text{BD}}$  is not affected by the receive-side correlation matrices  $\mathbf{R}_{R,\ell} \forall \ell \in \mathcal{S}^{\text{ZFC}}$ ; see Lemma 2 in Appendix A. Two bounds on  $z_\ell$  are given in the theorem. The lower bound is achieved by the suboptimal choice of  $\tilde{\mathbf{c}}_\ell$  as the dominating eigenvector of  $\mathbf{R}_{R,k}$ ; this makes  $\tilde{\mathbf{c}}_\ell^H \mathbf{H}_\ell \sim \mathcal{CN}(\mathbf{0}, \lambda_{\ell,M} \mathbf{I}_N)$ . The upper bound is achieved from Lemma 1 by applying Jensen's inequality and computing  $\mathbb{E}\{\log_2(\|\tilde{\mathbf{h}}_\ell\|_2^2)\} = \frac{\psi(N)}{\log_e(2)}$  using standard integration.

## APPENDIX C

*Proof of Theorem 2:* We begin with BD and assume that there are  $K$  candidates to become the new user  $k$ ,  $\mathcal{K} = \{1, \dots, K\}$ , while the other users in  $\mathcal{S}^{\text{BD}}$  are fixed. Since  $|\mathcal{S}^{\text{BD}}| = \frac{N}{M}$ , all available degrees of freedom are consumed by the interference cancellation. The precoding matrix  $\mathbf{W}_k^{\text{BD}}$  is therefore completely determined by the common null space of the co-users' channels and fixed in this proof.

Minimizing (14) corresponds to finding the user  $\ell \in \mathcal{K}$  with the row space of  $\mathbf{H}_\ell$  most compatible with  $\mathbf{W}_k^{\text{BD}}$ . For a user candidate  $\ell \in \mathcal{K}$ , we can lower bound (14) as

$$\begin{aligned} & -\mathbb{E}\{\log_2 \det(\mathbf{B}_\ell \mathbf{W}_k^{\text{BD}} \mathbf{W}_k^{\text{BD},H} \mathbf{B}_\ell^H)\} \\ & = -M \mathbb{E}\{\log_2 (\det(\mathbf{B}_\ell \mathbf{W}_k^{\text{BD}} \mathbf{W}_k^{\text{BD},H} \mathbf{B}_\ell^H)^{1/M})\} \\ & \geq -M \mathbb{E} \left\{ \log_2 \left( \frac{\text{tr}(\mathbf{B}_\ell \mathbf{W}_k^{\text{BD}} \mathbf{W}_k^{\text{BD},H} \mathbf{B}_\ell^H)}{M} \right) \right\} \\ & \geq -M \log_2 \left( \frac{\mathbb{E}\{\text{tr}(\mathbf{B}_\ell \mathbf{W}_k^{\text{BD}} \mathbf{W}_k^{\text{BD},H} \mathbf{B}_\ell^H)\}}{M} \right) \\ & = -M \log_2 \left( 1 + \frac{\mathbb{E}\{\text{tr}(\mathbf{B}_\ell \mathbf{W}_k^{\text{BD}} \mathbf{W}_k^{\text{BD},H} \mathbf{B}_\ell^H) - M\}}{M} \right). \end{aligned} \quad (42)$$

The first inequality is the classic inequality between arithmetic and geometric means, while the second inequality follows

<sup>15</sup>The vector  $\tilde{\mathbf{h}}_\ell$  is correlated with  $\tilde{\mathbf{c}}_\ell^H \mathbf{H}_\ell$  (they have the same direction), but this property does not affect the proof.

from applying Jensen's inequality on the convex function  $-\log_2 \det(\cdot)$ . The final expression in (42) contains  $M - \text{tr}(\mathbf{B}_\ell \mathbf{W}_k^{\text{BD}} \mathbf{W}_k^{\text{BD},H} \mathbf{B}_\ell^H)$ , which is the squared chordal distance between  $\mathbf{B}_\ell$  and  $\mathbf{W}_k^{\text{BD}}$ .

Since the matrices  $\mathbf{B}_\ell$ , for  $\ell \in \mathcal{K}$ , are independent and isotropically distributed on the Grassmannian manifold  $\mathcal{G}_{N,M}$  irrespective of the receive-side correlation (see Lemma 2), we can bring in results from [28] on quantization of Grassmannian manifolds using  $K$  random codewords. From [28, Theorem 4], we have the following lower bound on the average squared chordal distance (for sufficiently large  $K$ ):

$$\begin{aligned} \min_{\ell \in \mathcal{K}} \mathbb{E} \left\{ M - \text{tr}(\mathbf{B}_\ell \mathbf{W}_k^{\text{BD}} \mathbf{W}_k^{\text{BD},H} \mathbf{B}_\ell^H) \right\} \\ \geq \frac{M(N-M)}{M(N-M)+1} c_{N,M,M,2}^{-\frac{1}{M(N-M)}} K^{-\frac{1}{M(N-M)}} \end{aligned} \quad (43)$$

where  $c_{N,M,M,2}$  is a positive constant defined in [28, Eq. (8)]. Plugging (43) into (42) yields the lower bound for BD in the theorem.

A similar approach can be taken under ZFC (by setting  $M = 1$  in the derivation), but the  $M$  receive antennas provide degrees of freedom to select the effective channel as the vector in the row space of  $\mathbf{H}_\ell$  that minimizes the chordal distance to  $\mathbf{w}_k^{\text{ZFC}}$ . This is done by the QBC approach in [8], which was derived for uncorrelated channels but can be applied under receive correlation due to Lemma 2. We apply [8, Lemma 1], which says that the minimal chordal distance is the minimum of  $K$  independent  $\beta(N-M, M)$ -distributed random variables. This quantity can be lower bounded by taking the minimum of  $K$  independent  $\beta(N-M, 1)$  variables and further lower bounded by the quantization bound in [28, Theorem 4]:

$$\min_{\ell \in \mathcal{K}} \mathbb{E} \left\{ 1 - \left| \frac{\mathbf{h}_\ell^H \mathbf{w}_k^{\text{ZFC}}}{\|\mathbf{h}_\ell\|_2} \right|^2 \right\} \geq \frac{(N-M)K^{-\frac{1}{(N-M)}}}{(N-M)+1} c_{N-M+1,1,1,2}^{-\frac{1}{(N-M)}} \quad (44)$$

where  $c_{N-M+1,1,1,2}$  is a positive constant defined in [28, Eq. (8)]. Plugging (44) into (42) for  $M = 1$  yields the lower bound for ZFC in the theorem.

## APPENDIX D

*Proof of Theorem 3:* Using Lemma 2, the row space of the correlated channel  $\mathbf{H}_k = \mathbf{R}_{R,k}^{1/2} \tilde{\mathbf{H}}_k$  is the same as for the uncorrelated channel  $\tilde{\mathbf{H}}_k$ . Consequently,  $\bar{\mathbf{W}}_k^{\text{BD}}$  will be isotropically distributed on the Grassmannian manifold  $\mathcal{G}_{N,M}$ , just as proved for uncorrelated channels in [7, Theorem 1]. The performance loss can therefore be bounded using Lemma 3 and it only remains to characterize  $\mathbb{E}\{\mathbf{H}_k \bar{\mathbf{W}}_\ell^{\text{BD}} \bar{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{H}_k^H\}$  for  $\ell \neq k$ . Observe that

$$\begin{aligned} & \mathbb{E}\{\mathbf{H}_k \bar{\mathbf{W}}_\ell^{\text{BD}} \bar{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{H}_k^H\} \\ & = \mathbf{R}_{R,k}^{1/2} \mathbb{E}\{\mathbf{L}_k \mathbf{Q}_k \bar{\mathbf{W}}_\ell^{\text{BD}} \bar{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{Q}_k^H \mathbf{L}_k^H\} \mathbf{R}_{R,k}^{1/2,H} \end{aligned} \quad (45)$$

using that  $\mathbf{H}_k = \mathbf{R}_{R,k}^{1/2} \tilde{\mathbf{H}}_k = \mathbf{R}_{R,k}^{1/2} \mathbf{L}_k \mathbf{Q}_k$ , where  $\mathbf{L}_k \in \mathbb{C}^{M \times M}$  is the lower triangular matrix and  $\mathbf{Q}_k \in \mathbb{C}^{M \times N}$  is the semi-unitary matrix in an LQ decomposition of  $\tilde{\mathbf{H}}_k$ . Observe that

$\mathbf{L}_k$  and  $\mathbf{Q}_k$  are independent, thus we can calculate their expectations sequentially as

$$\begin{aligned} \mathbb{E}\{\mathbf{L}_k \mathbf{Q}_k \bar{\mathbf{W}}_\ell^{\text{BD}} \bar{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{Q}_k^H \mathbf{L}_k^H\} \\ = \frac{D^{\text{BD}}}{N-M} \mathbb{E}\{\mathbf{L}_k \mathbf{I}_M \mathbf{L}_k^H\} = \frac{ND^{\text{BD}}}{N-M} \mathbf{I}_M. \end{aligned} \quad (46)$$

The first equality follows from [7, Eq. (43)–(45)], while the second follows from  $\mathbb{E}\{\mathbf{L}_k \mathbf{L}_k^H\} = N\mathbf{I}_M$  (since  $\mathbb{E}\{\tilde{\mathbf{H}}_k \tilde{\mathbf{H}}_k^H\} = N\mathbf{I}_M$ ). Plugging (46) into Lemma 3 yields

$$\Delta^{\text{BD}} \leq \log_2 \det \left( \mathbf{I}_M + \frac{P}{N} \left( \frac{N}{M} - 1 \right) \frac{ND^{\text{BD}}}{N-M} \mathbf{R}_{R,k} \right) \quad (47)$$

from which (20) follows directly. The approximate expression for  $D^{\text{BD}}$  is given in [7, Eq. (26)].

#### APPENDIX E

*Proof of Theorem 4:* This proof follows along the lines of [8, Theorem 1], with the difference that 1) we have spatial correlation at the receiver; and 2) we use QBC also under perfect CSI. Using Lemma 2, we observe that the row space of the correlated channel  $\mathbf{H}_k = \mathbf{R}_{R,k}^{1/2} \tilde{\mathbf{H}}_k$  is the same as for the uncorrelated channel  $\tilde{\mathbf{H}}_k$ . Since the gain of the effective channel is ignored in (22), the error-minimizing codeword is the same as for uncorrelated channels and we can apply [8, Lemma 2] to conclude that the direction of the effective channel  $\mathbf{h}_k = \mathbf{H}_k^H \tilde{\mathbf{c}}_k^{\text{QBC}}$  is isotropically distributed. The beamforming vector  $\bar{\mathbf{w}}_k^{\text{ZFC}}$  is independent of  $\mathbf{h}_k$  and also isotropic, thus the performance loss can be bounded using Lemma 3. It only remains to characterize  $\mathbb{E}\{|\mathbf{h}_k^H \bar{\mathbf{w}}_\ell^{\text{ZFC}}|^2\} = \mathbb{E}\{\|\mathbf{h}_k\|_2^2\} \mathbb{E}\{|\frac{\mathbf{h}_k^H}{\|\mathbf{h}_k\|_2} \bar{\mathbf{w}}_\ell^{\text{ZFC}}|^2\}$  for  $\ell \neq k$ . The second factor equals  $\frac{D^{\text{QBC}}}{N-1}$  using [41, Lemma 2] and [8, Eq. (17)], while computing the average norm  $\mathbb{E}\{\|\mathbf{h}_k\|_2^2\}$  of the effective channel is nontrivial. To enable reuse of results from [8], let  $\tilde{\mathbf{c}}_k^{\text{U-QBC}}$  be the QBC for the uncorrelated channel  $\tilde{\mathbf{H}}_k$  and observe that

$$\tilde{\mathbf{c}}_k^{\text{QBC}} = \frac{\mathbf{R}_{R,k}^{-1/2} \tilde{\mathbf{c}}_k^{\text{U-QBC}}}{\|\mathbf{R}_{R,k}^{-1/2} \tilde{\mathbf{c}}_k^{\text{U-QBC}}\|_2}. \quad (48)$$

We can therefore express the effective channel as

$$\mathbf{h}_k = \mathbf{H}_k^H \tilde{\mathbf{c}}_k^{\text{QBC}} = \mathbf{H}_k^H \frac{\mathbf{R}_{R,k}^{-1/2} \tilde{\mathbf{c}}_k^{\text{U-QBC}}}{\|\mathbf{R}_{R,k}^{-1/2} \tilde{\mathbf{c}}_k^{\text{U-QBC}}\|_2} = \frac{\tilde{\mathbf{H}}_k^H \tilde{\mathbf{c}}_k^{\text{U-QBC}}}{\|\mathbf{R}_{R,k}^{-1/2} \tilde{\mathbf{c}}_k^{\text{U-QBC}}\|_2} \quad (49)$$

and its squared norm will be

$$\|\mathbf{h}_k\|_2^2 = \|\tilde{\mathbf{H}}_k^H \tilde{\mathbf{c}}_k^{\text{U-QBC}}\|_2^2 \frac{1}{\tilde{\mathbf{c}}_k^{\text{U-QBC},H} \mathbf{R}_{R,k}^{-1} \tilde{\mathbf{c}}_k^{\text{U-QBC}}}. \quad (50)$$

The first factor is the same as under uncorrelated fading and satisfies  $\mathbb{E}\{\|\tilde{\mathbf{H}}_k^H \tilde{\mathbf{c}}_k^{\text{U-QBC}}\|_2^2\} = N - M + 1$  (see [8, Lemma 4]), while the second factor depends on  $\mathbf{R}_{R,k}$ . Since both the quantization codebook and  $\tilde{\mathbf{H}}_k$  are isotropically distributed,  $\tilde{\mathbf{c}}_k^{\text{U-QBC}}$  is also isotropic and the two terms in (50) are independent. To characterize the second term, observe that  $\tilde{\mathbf{c}}_k^{\text{U-QBC}}$  can be viewed as a normalized uncorrelated circular-symmetric complex Gaussian vector. By using that the eigenvectors

of  $\mathbf{R}_{R,k}$  are not affecting the distribution and that squared magnitudes of  $\mathcal{CN}(0, 1)$ -variables are exponentially distributed [44], we conclude that the second term of (50) has the same distribution as

$$\frac{\sum_{i=1}^M \xi_i}{\sum_{i=1}^M \frac{\xi_i}{\lambda_{k,i}}} \quad (51)$$

for some independent exponentially distributed  $\xi_i \sim \text{Exp}(1)$ . For any  $a$  such that  $\lambda_{k,m} \leq a \leq \lambda_{k,m+1}$ , we can write the CDF as

$$\begin{aligned} \Pr \left\{ \frac{\sum_{i=1}^M \xi_i}{\sum_{i=1}^M \frac{\xi_i}{\lambda_{k,i}}} \leq a \right\} \\ = \Pr \left\{ \sum_{i=1}^m \underbrace{\left( \frac{a}{\lambda_{k,i}} - 1 \right)}_{\geq 0} \xi_i - \sum_{i=m+1}^M \underbrace{\left( 1 - \frac{a}{\lambda_{k,i}} \right)}_{\geq 0} \xi_i \geq 0 \right\}. \end{aligned} \quad (52)$$

This is the difference of two sums of exponentially distributed variables (with distinct positive variances). The PDF of each sum is characterized by [44, Theorem 4] and by calculating their convolution and integrating over all positive values, we achieve the CDF

$$\begin{aligned} \Pr \left\{ \frac{\sum_{i=1}^M \xi_i}{\sum_{i=1}^M \frac{\xi_i}{\lambda_{k,i}}} \leq a \right\} \\ = \sum_{n=1}^m \sum_{t=m+1}^M \frac{(\mu_n - a^{-1})^m (a^{-1} - \mu_t)^{M-m-1}}{(\mu_n - \mu_t) \prod_{\substack{i=1 \\ i \neq n}}^m (\mu_n - \mu_i) \prod_{\substack{j=m+1 \\ j \neq t}}^M (\mu_j - \mu_t)} \end{aligned} \quad (53)$$

using the simplifying notation  $\mu_n = \frac{1}{\lambda_{k,n}}$ . The corresponding mean value is achieved from the CDF by simply taking the derivative and sum up the mean values over each  $a$ -interval. By multiplying the mean value expression with  $N - M + 1$  (i.e., the contribution of the first part in (50)), we achieve the expression for  $G_k$ .

#### APPENDIX F

*Proof of Theorem 5:* The proof follows along the lines of Theorem 3, but we consider CSI estimation errors instead of quantization errors. First, observe that both  $\mathbf{W}_\ell^{\text{BD}}$  and  $\tilde{\mathbf{W}}_\ell^{\text{BD}}$  are isotropically distributed on the Grassmannian manifold  $\mathcal{G}_{N,M}$  (since receive-side correlation is not affecting the row space of  $\mathbf{H}_k$  and  $\tilde{\mathbf{H}}_k$ ; see Lemma 2). The performance loss can therefore be bounded using Lemma 3 and it only remains to characterize  $\mathbb{E}\{\mathbf{H}_k \tilde{\mathbf{W}}_\ell^{\text{BD}} \tilde{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{H}_k^H\}$  for  $\ell \neq k$ . From (28) we have

$$\mathbf{H}_k = \hat{\mathbf{H}}_k + \mathbf{R}_{E,k}^{1/2} \tilde{\mathbf{E}}_k \quad (54)$$

where the second term is the estimation error,  $\mathbf{R}_{E,k} = \left( \mathbf{R}_{R,k}^{-T} + \frac{\mathbf{T}_k^H \mathbf{T}_k}{\sigma^2} \right)^{-1}$ , and  $\tilde{\mathbf{E}}_k$  has  $\mathcal{CN}(0, 1)$ -entries. By using that  $\hat{\mathbf{H}}_k \tilde{\mathbf{W}}_\ell^{\text{BD}} = \mathbf{0}$  for  $\ell \neq k$ , we achieve

$$\mathbb{E}\{\mathbf{H}_k \tilde{\mathbf{W}}_\ell^{\text{BD}} \tilde{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{H}_k^H\} = \mathbf{R}_{E,k}^{1/2} \mathbb{E}\{\tilde{\mathbf{E}}_k \tilde{\mathbf{W}}_\ell^{\text{BD}} \tilde{\mathbf{W}}_\ell^{\text{BD},H} \tilde{\mathbf{E}}_k^H\} \mathbf{R}_{E,k}^{1/2} \quad (55)$$

where  $\mathbb{E}\{\tilde{\mathbf{E}}_k \widehat{\mathbf{W}}_\ell^{\text{BD}} \widehat{\mathbf{W}}_\ell^{\text{BD},H} \tilde{\mathbf{E}}_k^H\} = M \mathbf{I}_M$  since  $\tilde{\mathbf{E}}_k$  is complex Gaussian and independent of  $\widehat{\mathbf{W}}_\ell^{\text{BD}}$ . Therefore,  $\mathbb{E}\{\mathbf{H}_k \widehat{\mathbf{W}}_\ell^{\text{BD}} \widehat{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{H}_k^H\} = M \left( \mathbf{R}_{R,k}^{-T} + \frac{\mathbf{T}_k^H \mathbf{T}_k}{\sigma^2} \right)^{-1}$ .

#### APPENDIX G

*Proof of Lemma 1:* Observe that  $\mathbf{H} = \mathbf{R}_R^{1/2} \tilde{\mathbf{H}}$  has the same distribution as  $\mathbf{H}\mathbf{U}$  for any unitary matrix  $\mathbf{U}$ . Thus, we can rotate  $\mathbf{h}$  arbitrarily without changing the statistics, meaning that  $\frac{\mathbf{h}}{\|\mathbf{h}\|_2}$  must be isotropically distributed. Next, note that  $\|\mathbf{h}\|_2^2 = \|\tilde{\mathbf{c}}^H \mathbf{H} \mathbf{U}\|_2^2 = \|\tilde{\mathbf{c}}^H \mathbf{H}\|_2^2$ , thus unitary rotations will not affect the effective channel gain meaning that the direction and the channel gain are statistically independent.  $\|\mathbf{h}\|_2^2$  is the dominating eigenvalue of the correlated complex Wishart matrix  $\mathbf{H}\mathbf{H}^H \in \mathcal{W}_M(N, \mathbf{R}_R)$ . The expectation in (32) achieved directly from [45, Theorem 3] or by using the moment generating function in [46] (which gives an equivalent expression that looks slightly different).

#### APPENDIX H

*Proof of Theorem 6:* This theorem is proved in the same way as Theorem 5. The only notable difference is that we use the effective channel  $\mathbf{h}_k$ , which has a single effective receive antenna, instead of the original channel  $\mathbf{H}_k$ . The effective channel is zero-mean and has an average channel gain  $\mathbb{E}\{\|\mathbf{h}_k\|_2^2\}$  given by (32). Thus, the effective channel and its channel estimate is related as  $\mathbf{h}_k^H = \hat{\mathbf{h}}_k^H + \left( \frac{1}{\mathbb{E}\{\|\mathbf{h}_k\|_2^2\}} + \frac{\Psi}{\sigma^2} \right)^{-1/2} \tilde{\mathbf{e}}_k^H$ , where  $\tilde{\mathbf{e}}_k \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$ .

#### APPENDIX I

*Proof of Corollary 2:* The sufficiency is easily achieved from Theorem 5 and Theorem 6. To obtain the necessity with BD, consider the interference term  $\mathbf{H}_k \widehat{\mathbf{W}}_\ell^{\text{BD}} \hat{\mathbf{Y}}_\ell \widehat{\mathbf{W}}_\ell^{\text{BD},H} \mathbf{H}_k^H = \mathbf{R}_{E,k}^{1/2} \tilde{\mathbf{E}}_k \widehat{\mathbf{W}}_\ell^{\text{BD}} \widehat{\mathbf{W}}_\ell^{\text{BD},H} \tilde{\mathbf{E}}_k^H \mathbf{R}_{E,k}^{1/2}$ , where the true and estimated are related as in (54). This term must be bounded as  $P \rightarrow \infty$ . Since the product  $\tilde{\mathbf{E}}_k \widehat{\mathbf{W}}_\ell^{\text{BD}}$  is non-zero almost surely and  $\hat{\mathbf{Y}}_\ell \rightarrow \frac{P}{M[\text{SBD}]} \mathbf{I}_M$ , it is necessary that  $\frac{1}{P} \mathbf{R}_{E,k}$  has bounded elements. This makes (36) a necessary condition. The proof for ZFC is analogous.

#### REFERENCES

- [1] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Sälzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, 2007.
- [2] F. Rusek, D. Persson, B. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *IEEE Signal Process. Mag.*, vol. 30, no. 1, pp. 40–60, 2013.
- [3] J. Hoydis, M. Kobayashi, and M. Debbah, "Green small-cell networks," *IEEE Veh. Technol. Mag.*, vol. 6, no. 1, pp. 37–43, 2011.
- [4] H. Weingarten, Y. Steinberg, and S. Shamai, "The capacity region of the Gaussian multiple-input multiple-output broadcast channel," *IEEE Trans. Inf. Theory*, vol. 52, no. 9, pp. 3936–3964, 2006.
- [5] R. Roy and B. Ottersten, "Spatial division multiple access wireless communication systems," US Patent, 1991, 5515378.
- [6] Q. Spencer, A. Swindlehurst, and M. Haardt, "Zero-forcing methods for downlink spatial multiplexing in multiuser MIMO channels," *IEEE Trans. Signal Process.*, vol. 52, no. 2, pp. 461–471, 2004.
- [7] N. Ravindran and N. Jindal, "Limited feedback-based block diagonalization for the MIMO broadcast channel," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1473–1482, 2008.
- [8] N. Jindal, "Antenna combining for the MIMO downlink channel," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3834–3844, 2008.
- [9] M. Trivellato, F. Boccardi, and H. Huang, "On transceiver design and channel quantization for downlink multiuser MIMO systems with limited feedback," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1494–1504, 2008.
- [10] Z.-Q. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Sel. Topics Signal Process.*, vol. 2, no. 1, pp. 57–73, 2008.
- [11] A. Tölli and M. Juntti, "Scheduling for multiuser MIMO downlink with linear processing," in *Proc. IEEE PIMRC*, 2005, pp. 156–160.
- [12] F. Boccardi and H. Huang, "A near-optimum technique using linear precoding for the MIMO broadcast channel," in *Proc. IEEE ICASSP*, vol. 3, 2007, pp. 17–20.
- [13] R. Chen, Z. Shen, J. Andrews, and R. Heath, "Multimode transmission for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 56, no. 7, pp. 3294–3302, 2008.
- [14] C. Guthy, W. Utschick, R. Hunger, and M. Joham, "Efficient weighted sum rate maximization with linear precoding," *IEEE Trans. Signal Process.*, vol. 58, no. 4, pp. 2284–2297, 2010.
- [15] A. Bayesteh and A. Khandani, "On the user selection for MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1086–1107, 2008.
- [16] W. Rhee, W. Yu, and J. Cioffi, "The optimality of beamforming in uplink multiuser wireless systems," *IEEE Trans. Wireless Commun.*, vol. 3, no. 1, pp. 86–96, 2004.
- [17] M. Dohler, R. Heath, A. Lozano, C. Papadias, and R. Valenzuela, "Is the PHY layer dead?" *IEEE Commun. Mag.*, vol. 49, no. 4, pp. 159–165, 2011.
- [18] L. Zheng and D. Tse, "Diversity and multiplexing: A fundamental tradeoff in multiple-antenna channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 5, pp. 1073–1096, 2003.
- [19] A. Lozano and N. Jindal, "Transmit diversity vs. spatial multiplexing in modern MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 9, no. 1, pp. 186–197, 2010.
- [20] D. Chizhik, J. Ling, P. Wolniansky, R. Valenzuela, N. Costa, and K. Huber, "Multiple-input-multiple-output measurements and modeling in Manhattan," *IEEE J. Sel. Areas Commun.*, vol. 21, no. 3, pp. 321–331, 2003.
- [21] K. Yu, M. Bengtsson, B. Ottersten, D. McNamara, P. Karlsson, and M. Beach, "Modeling of wide-band MIMO radio channels based on NLoS indoor measurements," *IEEE Trans. Veh. Technol.*, vol. 53, no. 3, pp. 655–665, 2004.
- [22] E. Björnson and B. Ottersten, "A framework for training-based estimation in arbitrarily correlated Rician MIMO channels with Rician disturbance," *IEEE Trans. Signal Process.*, vol. 58, no. 3, pp. 1807–1820, 2010.
- [23] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser MIMO achievable rates with downlink training and channel state feedback," *IEEE Trans. Inf. Theory*, vol. 56, no. 6, pp. 2845–2866, 2010.
- [24] T. Yoo and A. Goldsmith, "Capacity and power allocation for fading MIMO channels with channel estimation error," *IEEE Trans. Inf. Theory*, vol. 52, no. 5, pp. 2203–2214, 2006.
- [25] Q. Shi, M. Razaviyayn, Z.-Q. Luo, and C. He, "An iteratively weighted MMSE approach to distributed sum-utility maximization for a MIMO interfering broadcast channel," *IEEE Trans. Signal Process.*, vol. 59, no. 9, pp. 4331–4340, 2011.
- [26] T. Yoo and A. Goldsmith, "On the optimality of multi-antenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 3, pp. 528–541, 2006.
- [27] J. Lee and N. Jindal, "High SNR analysis for MIMO broadcast channels: Dirty paper coding vs. linear precoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4787–4792, 2007.
- [28] W. Dai, Y. Liu, and B. Rider, "Quantization bounds on Grassmann manifolds and applications to MIMO communications," *IEEE Trans. Inf. Theory*, vol. 54, no. 3, pp. 1108–1123, 2008.
- [29] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, no. 9, pp. 369–371, 2001.
- [30] E. Björnson, D. Hammarwall, and B. Ottersten, "Exploiting quantized channel norm feedback through conditional statistics in arbitrarily correlated MIMO systems," *IEEE Trans. Signal Process.*, vol. 57, no. 10, pp. 4027–4041, 2009.
- [31] Z. Shen, R. Chen, J. Andrews, R. Heath, and B. Evans, "Low complexity user selection algorithms for multiuser MIMO systems with block diagonalization," *IEEE Trans. Signal Process.*, vol. 54, no. 9, pp. 3658–3663, 2006.

- [32] E. Björnson, M. Bengtsson, and B. Ottersten, "Receive combining vs. multistream multiplexing in multiuser MIMO systems," in *Proc. IEEE Swe-CTW*, 2011, pp. 109–114.
- [33] D. Love, R. Heath, and T. Strohmer, "Grassmannian beamforming for multiple-input multiple-output wireless systems," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2735–2747, 2003.
- [34] V. Raghavan, R. Heath, and A. Sayeed, "Systematic codebook designs for quantized beamforming in correlated MIMO channels," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1298–1310, 2007.
- [35] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341–1365, 2008.
- [36] S. Zhou and B. Li, "BER criterion and codebook construction for finite-rate precoded spatial multiplexing with linear receivers," *IEEE Trans. Signal Process.*, vol. 54, no. 5, pp. 1653–1665, 2006.
- [37] W. Santipach and M. Honig, "Asymptotic capacity of beamforming with limited feedback," in *Proc. IEEE ISIT*, 2004.
- [38] C. Au-Yeung and D. Love, "On the performance of random vector quantization limited feedback beamforming in a MISO system," *IEEE Trans. Wireless Commun.*, vol. 6, no. 2, pp. 458–462, 2007.
- [39] M. Kountouris, R. de Francisco, D. Gesbert, D. Slock, and T. Sälzer, "Multiuser diversity - multiplexing tradeoff in MIMO broadcast channels with limited feedback," in *Proc. ACSSC*, 2006, pp. 364–368.
- [40] N. Ravindran and N. Jindal, "Multi-user diversity vs. accurate channel feedback for MIMO broadcast channels," in *Proc. IEEE ICC*, 2008, pp. 3684–3688.
- [41] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, 2006.
- [42] T. Yoo, N. Jindal, and A. Goldsmith, "Multi-antenna downlink channels with limited feedback and user selection," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1478–1491, 2007.
- [43] E. Björnson and E. Jorswieck, "Optimal resource allocation in coordinated multi-cell systems," *Foundations and Trends in Communications and Information Theory*, vol. 9, no. 2-3, pp. 113–381, 2013.
- [44] D. Hammarwall, M. Bengtsson, and B. Ottersten, "Acquiring partial CSI for spatially selective transmission by instantaneous channel norm feedback," *IEEE Trans. Signal Process.*, vol. 56, no. 3, pp. 1188–1204, 2008.
- [45] E. Björnson and B. Ottersten, "Post-user-selection quantization and estimation of correlated Frobenius and spectral channel norms," in *Proc. IEEE PIMRC*, 2008.
- [46] M. McKay, I. Collings, and P. Smith, "Capacity and SER analysis of MIMO beamforming with MRC," in *Proc. IEEE ICC*, 2006.



**Emil Björnson** (S'07-M'12) was born in Malmö, Sweden, in 1983. He received the M.S. degree in Engineering Mathematics from Lund University, Lund, Sweden, in 2007. He received the Ph.D. degree in Telecommunications from the Signal Processing Lab at KTH Royal Institute of Technology, Stockholm, Sweden, in 2011. He is the first author of the book "Optimal Resource Allocation in Coordinated Multi-Cell Systems" published in 2013.

Dr. Björnson was one of the first recipients of the International Postdoc Grant from the Swedish Research Council. This grant is currently funding a joint postdoctoral research fellowship at the Alcatel-Lucent Chair on Flexible Radio, SUPELEC, Paris, France, and the Signal Processing Lab at KTH Royal Institute of Technology, Stockholm, Sweden. His research interests include multi-antenna communications, resource allocation, random matrix theory, estimation theory, stochastic signal processing, and optimization.

For his work on optimization of multi-cell MIMO communications, he received a Best Paper Award at the 2009 International Conference on Wireless Communications & Signal Processing (WCSP'09) and a Best Student Paper Award at the 2011 IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP'11).



**Marios Kountouris** (S'04-M'08) received the Diploma in Electrical and Computer Engineering from the National Technical University of Athens, Greece, in 2002 and the M.Sc. and Ph.D. degrees in Electrical Engineering from the Ecole Nationale Supérieure des Télécommunications (Telecom Paris-Tech), France, in 2004 and 2008, respectively. His doctoral research was carried out at Eurecom Institute, France, funded by France Telecom - Orange Labs, France. From February 2008 to May 2009, he has been with the Department of Electrical and Computer Engineering at the University of Texas at Austin, USA, as a postdoctoral research associate, working on wireless ad hoc networks under DARPA's ITMANET program. Since June 2009 he has been with the Department of Telecommunications at SUPELEC (Ecole Supérieure D'Electricité), France where he is currently an Assistant Professor. His research interests include multiuser multi-antenna communications, heterogeneous wireless networks, interference modeling, and network information theory. He is currently an Editor for the EURASIP Journal on Wireless Communications and Networking and Vice Chair of IEEE SIG on Green Cellular Networks. He received the Best Paper Award in Communication Theory Symposium at the IEEE Globecom conference in 2009 and the 2012 IEEE SPS Signal Processing Magazine Award. He is a Member of the IEEE and a Professional Engineer of the Technical Chamber of Greece.



**Mats Bengtsson** (M'00-SM'06) received the M.S. degree in computer science from Linköping University, Linköping, Sweden, in 1991 and the Tech. Lic. and Ph.D. degrees in electrical engineering from the KTH Royal Institute of Technology, Stockholm, Sweden, in 1997 and 2000, respectively.

From 1991 to 1995, he was with Ericsson Telecom AB Karlstad. He currently holds a position as Associate Professor at the Signal Processing Laboratory, School of Electrical Engineering, KTH. His research interests include statistical signal processing and its applications to communications, multi-antenna processing, cooperative communication, radio resource management, and propagation channel modelling. Dr. Bengtsson served as Associate Editor for the IEEE Transactions on Signal Processing 2007-2009 and was a member of the IEEE SPCOM Technical Committee 2007-2012.



**Björn Ottersten** (S'87-M'89-SM'99-F'04) was born in Stockholm, Sweden, 1961. He received the M.S. degree in electrical engineering and applied physics from Linköping University, Linköping, Sweden, in 1986. In 1989 he received the Ph.D. degree in electrical engineering from Stanford University, Stanford, CA. Dr. Ottersten has held research positions at the Department of Electrical Engineering, Linköping University, the Information Systems Laboratory, Stanford University, the Katholieke Universiteit Leuven, Leuven, and the University of Luxembourg.

During 96/97 Dr. Ottersten was Director of Research at ArrayComm Inc, a start-up in San Jose, California based on Ottersten's patented technology. He has co-authored journal papers that received the IEEE Signal Processing Society Best Paper Award in 1993, 2001, and 2006 and 3 IEEE conference papers receiving Best Paper Awards. In 1991 he was appointed Professor of Signal Processing at the KTH Royal Institute of Technology, Stockholm. From 1992 to 2004 he was head of the department for Signals, Sensors, and Systems at KTH and from 2004 to 2008 he was dean of the School of Electrical Engineering at KTH.

Currently, Dr. Ottersten is Director for the Interdisciplinary Centre for Security, Reliability and Trust at the University of Luxembourg. As Digital Champion of Luxembourg, he acts as an adviser to European Commissioner Neelie Kroes. Dr. Ottersten has served as Associate Editor for the IEEE Transactions on Signal Processing and on the editorial board of IEEE Signal Processing Magazine. He is currently editor in chief of EURASIP Signal Processing Journal and a member of the editorial board of EURASIP Journal of Applied Signal Processing. Dr. Ottersten currently serves on the IEEE Signal Processing Society Board of Governors and is a Fellow of the IEEE and EURASIP. In 2011 he received the IEEE Signal Processing Society Technical Achievement Award. He is a first recipient of the European Research Council advanced research grant. His research interests include security and trust, reliable wireless communications, and statistical signal processing.