KTH Electrical Engineering

Receive Combining vs. Multistream Multiplexing in Multiuser MIMO Systems

Proceedings of IEEE Swedish Communication Technologies Workshop (Swe-CTW)

19-21 October, Stockholm, Sweden, 2011

EMIL BJÖRNSON, MATS BENGTSSON, AND BJÖRN OTTERSTEN

Stockholm 2011

KTH Royal Institute of Technology
ACCESS Linnaeus Center
Signal Processing Lab

DOI: 10.1109/Swe-CTW.2011.6082475

Receive Combining vs. Multistream Multiplexing in Multiuser MIMO Systems

Emil Björnson
ACCESS Linnaeus Center
KTH Royal Institute of Technology
SE-100 44 Stockholm, Sweden
Email: emil.bjornson@ee.kth.se

Mats Bengtsson
ACCESS Linnaeus Center
KTH Royal Institute of Technology
SE-100 44 Stockholm, Sweden
Email: mats.bengtsson@ee.kth.se

Björn Ottersten
ACCESS Linnaeus Center
KTH Royal Institute of Technology
SE-100 44 Stockholm, Sweden
Email: bjorn.ottersten@ee.kth.se

***Abstract*—In single-user transmission, the receive antennas should preferably be used to enable multiplexing. The situation is different under multiuser transmission, where only the number of transmit antennas limits the multiplexing gain. The system therefore has the choice between sending one stream per scheduled user (i.e., combining receive antennas for diversity) or selecting a smaller number of users and multiplex multiple streams to each of them. This tradeoff is investigated herein, based on zero-forcing (with receive antenna combining) and block-diagonalization precoding which represents the two extremes. Based on asymptotic analysis and numerical examples, the unexpected conclusion is that each user only should receive one stream and use its antennas to achieve a receive combining gain. This is explained by zero-forcing having a stronger resilience towards spatial correlation and larger benefit from multiuser diversity. This fundamental result has positive implications for the design of multiuser systems as it reduces the hardware constraints at the user devices.**

I. Introduction

Multiantenna techniques introduce spatial dimensions to wireless systems that can be exploited to enhance reliability and/or spatial multiplexing of multiple data streams with controlled interference [1]. The downlink single-cell sum capacity behaves as

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

where N is the number of base station antennas, K is the number of M-antenna users, and P is the total transmit power. The number of users is typically large (i.e., $K \geq N$) and thus the optimal *multiplexing gain* is $\min(N, MK) = N$.

The sum capacity can be achieved using dirty-paper coding [2], but this non-linear scheme has impractical complexity and is sensitive to having imperfect channel state information (CSI). Two linear precoding schemes that are both simple and can achieve the optimal multiplexing gain are *block-diagonalization (BD)* [3] and *zero-forcing with receive combining (ZFC)* [4], [5]. These schemes concentrate on interference mitigation, which is the limiting factor for increasing the performance in future cellular networks (with high cell density and large signal-to-noise ratios (SNRs)). The main difference between BD and ZFC lies in how the receive antennas are utilized at each user. ZFC only sends one data stream per scheduled user and each user can combine the received signals on its antennas to achieve receive diversity (i.e., an effective channel with better properties). BD selects fewer users than ZFC but multiplexes M streams to each of them, which relaxes the interference mitigation and enables joint/iterative detection of each user's streams. In other words, ZFC and BD represent the two extremes in the use of receive antennas at the user device: *receive combining* and *multistream multiplexing*. The multimode switching technique in [6] is an example of a scheme that operates in between these extremes.

This paper considers the tradeoff between receive combining and multistream multiplexing in a wireless system using asymptotic analysis and numerical examples. Despite the similar terminology, this problem is fundamentally different from the *transmit diversity-spatial multiplexing tradeoff* [7] that considers whether the transmitter should try to achieve the full multiplexing gain or send fewer streams to enhance reliability. For any given total number of streams, the question herein is whether these should be allocated to many or few users. Recent works show that spatial multiplexing outperforms transmit diversity in single-user transmission [8], while the CSI accuracy decides the total number of streams under multi-user transmission [9].

A prominent related work is [10], where BD and ZFC are compared under quantized CSI feedback. The analysis reveals a distinct advantage of BD, but is limited to uncorrelated channels and does not include scheduling. Similar numerical observations are made in [6] under perfect CSI and practical scheduling algorithms, but the authors also observe that certain zero-forcing schemes (not ZFC) are beneficial if the channels have high spatial correlation. In the limit of many users (i.e., $K \to \infty$), [11] proves analytically that zero-forcing achieves the sum capacity. This result is used in [5] to justify choosing ZFC over BD, but this is not very convincing because: 1) asymptotic optimality of BD can also be proved; 2) the performance at practical number of users is unknown (K needs to be very large to approach capacity); and 3) the analysis creates a multiuser diversity gain, which is a modeling artifact of Rayleigh fading channels [12]. This paper provides a comprehensive comparison of BD and ZFC that includes user selection, spatial correlation, and imperfect CSI.

B. Ottersten is also with Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg. 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 No. 228044.

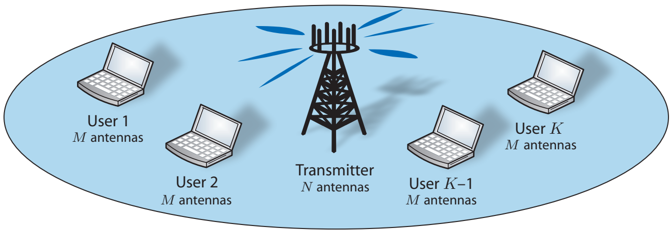

Fig. 1. Illustration of the downlink multiuser MIMO system herein.

The main contributions of this paper are:

- We use asymptotic analysis to show that ZFC is more resilient to spatial correlation and benefits more from multiuser diversity than BD. We derive upper bounds on the performance loss under imperfect CSI estimation and show that both ZFC and BD achieve the full multiplexing gain if the power used for channel estimation increases linearly with the transmit power.
- We provide numerical examples showing that ZFC outperforms BD. This indicates that receive combining is preferable over multistream multiplexing. A notable exception is under poor CSI accuracy, but single-user transmission is the best choice in these scenarios.

The mathematical proofs are omitted due to space limitations, but the results are verified by simulations.

II. System Model

We consider a downlink multiuser MIMO system where a single base station with N antennas communicates with $K \geq N$ users, where the kth user has $M < N$ antennas. For analytical convenience, we assume that $\frac{N}{M}$ is an integer. The narrowband, flat-fading channel to user k is represented in the complex-baseband by $\mathbf{H}_k \in \mathbb{C}^{M \times N}$ and is independent between the users. The received signal at this user is

$$\mathbf{y}_k = \mathbf{H}_k \mathbf{x} + \mathbf{n}_k \tag{2}$$

where $\mathbf{x} \in \mathbb{C}^{N \times 1}$ is the transmitted signal and $\mathbf{n}_k \in \mathcal{CN}(\mathbf{0}, \mathbf{I}_M)$ is the (normalized) noise vector. For analytical convenience and motivated by measurements, we employ the Kronecker model with $\mathbf{H}_k = \mathbf{R}_{R,k}^{1/2} \widetilde{\mathbf{H}}_k \mathbf{R}_{T,k}^{1/2}$, where $\mathbf{R}_{T,k}$ and $\mathbf{R}_{R,k}$ are the correlation matrices at the transmitter and receiver side, respectively, and $\widetilde{\mathbf{H}}_k$ has independent $\mathcal{CN}(0,1)$-entries. The correlation matrices are generally different for each users, describing different spatial properties.

We consider linear precoding and the transmitted signal is

$$\mathbf{x} = \sum_{k=1}^{K} \mathbf{V}_k \mathbf{d}_k \tag{3}$$

where $\mathbf{V}_k \in \mathbb{C}^{N \times d_k}$ is the precoding vector, $\mathbf{d}_k \in \mathcal{CN}(\mathbf{0}, \mathbf{I}_{d_k})$ is the data signal, and d_k is the number of multiplexed data streams to user k. The achievable data rate of this user is

$$g_k(P) = \log_2 \frac{\det\left(\mathbf{I}_M + \sum_{\bar{k}=1}^{K} \mathbf{H}_k \mathbf{V}_{\bar{k}} \mathbf{V}_{\bar{k}}^H \mathbf{H}_k^H\right)}{\det\left(\mathbf{I}_M + \sum_{\bar{k} \neq k} \mathbf{H}_k \mathbf{V}_{\bar{k}} \mathbf{V}_{\bar{k}}^H \mathbf{H}_k^H\right)} \tag{4}$$

and the transmission is limited by a total power constraint

$$\mathbb{E}\{\mathbf{x}^H \mathbf{x}\} = \sum_{k=1}^{K} \mathrm{tr}\{\mathbf{V}_k \mathbf{V}_k^H\} \leq P. \tag{5}$$

A. Block-Diagonalization and Zero-Forcing Precoding

Two linear precoding schemes that are both simple and have good asymptotic properties are: *block-diagonalization (BD)* [3] and *zero-forcing with receive combining (ZFC)* [4], [5]. Under perfect CSI, these schemes are defined as follows.

Definition 1. *(Block-Diagonalization)* Let $\mathcal{S}^{\mathrm{BD}}$ be a scheduling set with at most $\frac{N}{M}$ users. For each user $k \in \mathcal{S}^{\mathrm{BD}}$, we set $d_k = M$ and $\mathbf{V}_k = \sqrt{P/(M|\mathcal{S}^{\mathrm{BD}}|)}\mathbf{V}_k^{\mathrm{BD}}$, where $\mathbf{V}_k^{\mathrm{BD}}$ is a semi-unitary matrix that satisfies $\mathbf{V}_k^{\mathrm{BD},H} \mathbf{V}_k^{\mathrm{BD}} = \mathbf{I}_M$ and $\mathbf{H}_{\bar{k}} \mathbf{V}_k^{\mathrm{BD}} = \mathbf{0}$ for all $\bar{k} \in \mathcal{S}^{\mathrm{BD}} \setminus \{k\}$. The data rate is

$$g_k^{\mathrm{BD}}(P) = \log_2 \det\left(\mathbf{I}_M + \frac{P}{M|\mathcal{S}^{\mathrm{BD}}|} \mathbf{H}_k \mathbf{V}_k^{\mathrm{BD}} \mathbf{V}_k^{\mathrm{BD},H} \mathbf{H}_k^H\right). \tag{6}$$

Definition 2. *(Zero-Forcing with Receive Combining)* Each user combines its antennas using a channel-dependent unit-norm vector $\mathbf{r}_k \in \mathbb{C}^{M \times 1}$ (e.g., dominating left singular vector). Based on the effective channels $\mathbf{h}_k^H = \mathbf{r}_k^H \mathbf{H}_k \in \mathbb{C}^{1 \times N}$, a scheduling set $\mathcal{S}^{\mathrm{ZFC}}$ with at most N users is selected. For each user $k \in \mathcal{S}^{\mathrm{ZFC}}$, we set $d_k = 1$ and $\mathbf{V}_k = \sqrt{P/|\mathcal{S}^{\mathrm{ZFC}}|}\mathbf{v}_k^{\mathrm{ZFC}}$ where $\mathbf{v}_k^{\mathrm{ZFC}}$ is a unit-norm vector satisfying $\mathbf{h}_{\bar{k}}^H \mathbf{v}_k^{\mathrm{ZFC}} = 0$ for all $\bar{k} \in \mathcal{S}^{\mathrm{ZFC}} \setminus \{k\}$. The data rate is

$$g_k^{\mathrm{ZFC}}(P) = \log_2\left(1 + \frac{P}{|\mathcal{S}^{\mathrm{ZFC}}|} |\mathbf{r}_k^H \mathbf{H}_k \mathbf{v}_k^{\mathrm{ZFC}}|^2\right). \tag{7}$$

Observe that these definitions assumed equal power allocation, which is asymptotically optimal. The numerical examples will however include optimal waterfilling power allocation.

Both BD and ZFC are designed to create zero co-user interference, but the difference is that ZFC only sends one data stream per scheduled user while BD selects fewer users but multiplexes M streams to each of them. BD and ZFC are identical when each user only has one antenna (i.e., $M = 1$), but this does *not* mean that BD is generalization of ZFC. There are many good reasons for applying ZFC when $M > 1$:

1) The effective channels $\mathbf{h}_k$ have good properties (e.g., only the strongest singular direction of $\mathbf{H}_k$ is used);
2) Simplified hardware that only decodes one stream/user;
3) Only the effective channels $\mathbf{h}_k$ need to be known.

The interference mitigation is, on the other hand, less restrictive under BD since fewer users and dimensions are involved. This provides more degrees of freedom for efficient precoding to the scheduled users [10]. We will investigate which approach that is advantageous in terms of achievable sum rate performance. This will provide us with insight on whether the M receive antennas should be used for multiplexing (as in BD) or be combined to achieve receive diversity (as in ZFC).

The considered ZFC should not be confused with the zero-forcing generalization in [13], [14] that sends a separate stream to each of the M receive antennas at each user and creates zero inter-antenna interference. That approach always performs worse than BD, which will not be the case for ZFC.

III. Performance Analysis: Perfect CSI

In this section, we assume that both the base station and the users have perfect CSI. We will analyze the sum rate

$$f_{\text{sum}}(P) = \sum_{k=1}^{K} g_k(P), \tag{8}$$

which asymptotically (as $P \to \infty$) becomes [14]

$$f_{\text{sum}}^{\text{BD}}(P) \cong N \log_2\left(\frac{P}{N}\right) + \sum_{k \in \mathcal{S}^{\text{BD}}} \log_2 \det(\mathbf{H}_k \mathbf{V}_k^{\text{BD}} \mathbf{V}_k^{\text{BD},H} \mathbf{H}_k^H),$$

$$f_{\text{sum}}^{\text{ZFC}}(P) \cong N \log_2\left(\frac{P}{N}\right) + \sum_{k \in \mathcal{S}^{\text{ZFC}}} \log_2(|\mathbf{r}_k^H \mathbf{H}_k \mathbf{v}_k^{\text{ZFC}}|^2), \tag{9}$$

for BD and ZFC (with optimal power allocation), respectively.

For both schemes, 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 $\mathcal{M}_\infty = N$, which is the same high-SNR slope as for the sum capacity. We thus need to compare the rate offsets $\mathcal{R}_\infty$ to conclude which scheme that is preferable.

Theorem 1. Assume that the transmit antennas are uncorrelated (i.e., $\mathbf{R}_{T,k} = \mathbf{I}_N \ \forall k$), receive correlation matrices $\mathbf{R}_{R,k}$ have eigenvalues $\lambda_{k,M} \geq \ldots \geq \lambda_{k,1} > 0$, and random user selection with $|\mathcal{S}^{\text{BD}}| = \frac{N}{M}$, $|\mathcal{S}^{\text{ZFC}}| = N$. The expected asymptotic difference between BD and ZFC is bounded as

$$\bar{\beta} = \lim_{P \to \infty} \mathbb{E}\{f^{\text{BD}}(P) - f^{\text{ZFC}}(P)\} \leq N \frac{\log_2(e)}{M} \sum_{i=1}^{M-1} \frac{M-i}{i} + \log_2\Big(\frac{\prod_{k \in \mathcal{S}^{\text{BD}}} \prod_{m=1}^{M} \lambda_{k,m}}{\prod_{\bar{k} \in \mathcal{S}^{\text{ZFC}}} \lambda_{\bar{k},M}}\Big). \tag{10}$$

The expected asymptotic difference in Theorem 1 consists of two terms. The first term is positive and corresponds to the expected gain of BD in an uncorrelated system (see [14, Theorem 3]). The second term contains a ratio of eigenvalues. If all $\mathbf{R}_{R,k}$ have the same eigenvalues $\lambda_{k,m} = \lambda_m$, it becomes $N \log_2((\prod_{m=1}^{M} \lambda_m)^{1/M})/\lambda_M)$ which contains the geometric mean of all eigenvalues divided by the largest eigenvalue. This ratio is smaller than one (or equal in uncorrelated channels). As we take the logarithm of this quantity, the second term is negative and approaches $-\infty$ with increasing eigenvalue spread. In other words, Theorem 1 shows that BD has a (bounded) advantage on uncorrelated channels, while ZFC becomes the best choice as the receive side correlation grows. The explanation is that BD has less restrictive interference mitigation, but is more sensitive to poor channels since it uses all dimensions for transmission. We can expect a similar impact of spatial correlation at the transmitter side, although not included in the theorem since it complicates analysis.

Observe that this result was derived assuming random selection of the maximal number of users. Since $K \geq N$, only a subset of users are scheduled at each channel use. If users are unevenly distributed in the cell, it could be beneficial to intentionally schedule fewer users. Scheduling is also used to achieve user fairness and exploit multiuser diversity. The following theorem considers its impact on BD and ZF.

Theorem 2. Assume that the transmit antennas are uncorrelated (i.e., $\mathbf{R}_{T,k} = \mathbf{I}_N \ \forall k$). For any given $\mathcal{S}^{\text{BD}}, \mathcal{S}^{\text{ZFC}}$ (with $|\mathcal{S}^{\text{BD}}| = \frac{N}{M}$ and $|\mathcal{S}^{\text{ZFC}}| = N$), suppose that 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[1] compared with single-user transmission, the expected asymptotic losses for BD and ZFC, respectively, are lower bounded by

$$\begin{aligned} \mathbb{E}\{\text{Loss}_{\text{BD}}\} &\geq -M \log_2(1 - c_1 K^{-\frac{1}{M(N-M)}}) \\ \mathbb{E}\{\text{Loss}_{\text{ZFC}}\} &\geq -\log_2(1 - c_2 K^{-\frac{1}{N-M}}) \end{aligned} \tag{11}$$

when K is large (c_1, c_2 are positive constants).

This theorem indicates that it is easier to find users with near-orthogonal channels under ZFC than under BD, which is reasonable since the channels of BD users should have good fitting in M dimensions while ZFC users only have one dimension. If the scheduling also prioritizes strong users, recall from Theorem 1 that the performance of ZFC depends on the dominating eigenvalue, while BD depends on the geometric mean of the eigenvalues. The dominating eigenvalue among K users certainly scales faster than the geometric mean, thus ZFC should also be superior in exploiting strong channels.

A. Numerical Examples

Next, the analytical properties in Theorem 1 and Theorem 2 will be illustrated numerically. To this end, we adopt the simple exponential correlation model of [15], where

$$[\mathbf{R}(\rho)]_{ij} = \begin{cases} \rho^{j-i}, & i \leq j, \\ \rho^{*(i-j)}, & i > j, \end{cases} \quad |\rho| \leq 1, \ \angle\rho \in U[0, 2\pi). \tag{12}$$

The magnitude $|\rho|$ is the *correlation factor* between adjacent antennas, where $|\rho| = 0$ means no spatial correlation and $|\rho| = 1$ means full correlation. We assume for simplicity that all users have the same path loss. The dominating left singular vector of $\mathbf{H}_k$ is used as receive combiner $\mathbf{r}_k$ in the transmission design (to maximize $\|\mathbf{h}\|^2$) and is replaced by the MMSE combiner during data reception (see [5]):

$$\mathbf{r}_k^{\text{MMSE}} = \frac{\Big(\mathbf{I}_M + \frac{P}{|\mathcal{S}^{\text{ZFC}}|} \sum_{\bar{k} \in \mathcal{S}^{\text{ZFC}} \setminus \{k\}} \mathbf{H}_k \mathbf{v}_{\bar{k}}^{\text{ZFC}} \mathbf{v}_{\bar{k}}^{\text{ZFC},H} \mathbf{H}_k^H\Big)^{-1} \mathbf{H}_k \mathbf{v}_k^{\text{ZFC}}}{\Big\|\Big(\mathbf{I}_M + \frac{P}{|\mathcal{S}^{\text{ZFC}}|} \sum_{\bar{k} \in \mathcal{S}^{\text{ZFC}} \setminus \{k\}} \mathbf{H}_k \mathbf{v}_{\bar{k}}^{\text{ZFC}} \mathbf{v}_{\bar{k}}^{\text{ZFC},H} \mathbf{H}_k^H\Big)^{-1} \mathbf{H}_k \mathbf{v}_k^{\text{ZFC}}\Big\|}. \tag{13}$$

The expected asymptotic difference between BD and ZFC is shown in Fig. 2 as a function of the correlation factor, 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

[1] If $\mathbf{B}_k = \text{span}(\mathbf{H}_k)$ is an orthonormal basis of the row space of $\mathbf{H}_k$, then the expected asymptotic loss is $-\mathbb{E}\{\log_2 \det(\mathbf{B}_k \mathbf{V}_k \mathbf{V}_k^H \mathbf{B}_k^H)\}$.

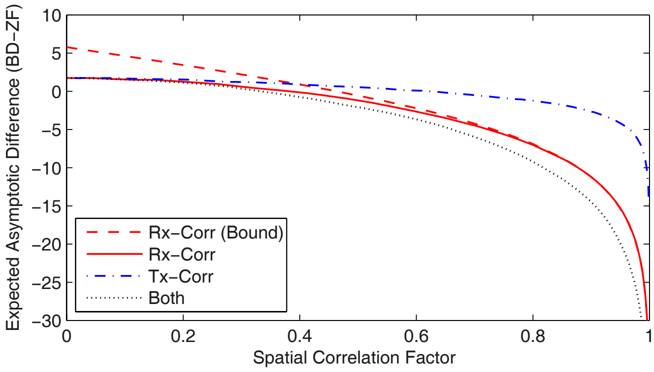

Fig. 2. The expected asymptotic difference between BD and ZFC in a system with $N = 8$ transmit antennas, $M = 2$ receive antennas per user, and random user selection. The impact of spatial correlation at the receiver, transmitter, and both sides is shown (using the exponential correlation model of [15]).

are correlated). The bound in Theorem 1 is only tight at high correlation. It is worth noting that $|\rho|$ impacts the perceived spatial correlation non-linearly; typical spatially correlated environments correspond to $|\rho| \approx 0.9$ [16].

The impact of user selection (also known as multiuser diversity) on the average sum rate is shown in Fig. 3, using the low-complexity BD capacity-based scheduling algorithm from [17]. We consider $N = 8$ transmit antennas, $M = 2$ or $M = 4$ receive antennas, and a total transmit power of $P = 10$ dB or $P = 20$ dB. As the system is spatially uncorrelated, it is remarkable that the inclusion of user selection makes ZFC superior to BD. Fewer than N streams are used when the number of users is small, while the number of active streams increases with K. ZFC clearly benefits more from multiuser diversity than BD, which confirms Theorem 2.

The conclusion is that ZFC is the method of choice in multiuser MIMO systems with perfect CSI. Although it is possible to find scenarios when BD is better than ZFC, these disregard spatial correlation and user selection.

IV. Performance Analysis: Imperfect CSI

Next, we relax the assumption of perfect CSI at the base station. While each user knows its own channel matrix perfectly, the base station only has imperfect channel estimates. The primary focus is on time division duplex (TDD) systems, where channel estimates are obtained through training signaling on the uplink (assuming perfect channel reciprocity). But it is worth noting that this approach is similar to analog CSI feedback in frequency division duplex (FDD) systems, where the unquantized channel coefficients are sent on an uplink subcarrier (to avoid the codebook complexity) [10], [18].[2]

The reciprocal uplink counterpart to the channel in (2) is

$$\tilde{\mathbf{y}}_k = \mathbf{H}_k^T \tilde{\mathbf{x}}_k + \mathbf{e}_k \tag{14}$$

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 $\mathbf{e}_k \in \mathcal{CN}(\mathbf{0}, \sigma^2 \mathbf{I}_N)$ is the noise vector. To estimate $\mathbf{H}_k$, user k sends a known training matrix $\mathbf{T}_k \in \mathbb{C}^{M\times M}$ over M uplink channel uses. Assuming perfect statistical CSI, the MMSE estimate $\widehat{\mathbf{H}}_k$ and the corresponding error covariance matrix $\mathbf{C}_k$ are [19]

$$\begin{aligned} \text{vec}(\widehat{\mathbf{H}}_k^T) &= \frac{1}{\sigma^2}\mathbf{C}_k \widetilde{\mathbf{T}}_k^H \text{vec}(\mathbf{Y}_k), \\ \mathbf{C}_k &= \left((\mathbf{R}_{R,k} \otimes \mathbf{R}_{T,k}^T)^{-1} + \frac{1}{\sigma^2} \widetilde{\mathbf{T}}_k^H \widetilde{\mathbf{T}}_k \right)^{-1} \end{aligned} \tag{15}$$

where $\widetilde{\mathbf{T}}_k = (\mathbf{T}_k^T \otimes \mathbf{I}_M)$ and $\mathbf{Y}_k$ is the received signal from training signaling. The training matrix $\mathbf{T}_k$ has a training power constraint $\frac{1}{M}\text{tr}\{\mathbf{T}_k^H \mathbf{T}_k\} = q$. We assume the use of an MSE-minimizing training matrix, given by [19, Theorem 1].

A. Block-Diagonalization Precoding

Under imperfect CSI, BD precoding cannot cancel all co-user interference. There is a variety of ways to handle the channel uncertainty, but a simple approach is to treat $\widehat{\mathbf{H}}_1, \ldots, \widehat{\mathbf{H}}_K$ as being the true channels [10]. This results in a lower bound on the performance and the data rate of user $k \in \mathcal{S}^{\text{BD}}$ is

$$g_k^{\text{BD-EST}}(P) = \log_2 \frac{\det\left(\mathbf{I}_M + \frac{P}{M|\mathcal{S}^{\text{BD}}|} \sum_{\bar{k}\in\mathcal{S}^{\text{BD}}} \mathbf{H}_k \widehat{\mathbf{V}}_{\bar{k}}^{\text{BD}} \widehat{\mathbf{V}}_{\bar{k}}^{\text{BD},H} \mathbf{H}_k^H \right)}{\det\left(\mathbf{I}_M + \frac{P}{M|\mathcal{S}^{\text{BD}}|} \sum_{\bar{k}\in\mathcal{S}^{\text{BD}}\setminus\{k\}} \mathbf{H}_k \widehat{\mathbf{V}}_{\bar{k}}^{\text{BD}} \widehat{\mathbf{V}}_{\bar{k}}^{\text{BD},H} \mathbf{H}_k^H \right)}. \tag{16}$$

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

Theorem 3. Assume that the transmit antennas are uncorrelated (i.e., $\mathbf{R}_{T,k} = \mathbf{I}_N \ \forall k$), that $\mathbf{R}_{R,k}$ has arbitrary correlation, and that $\frac{N}{M}$ users are scheduled randomly. The average rate loss for user $k \in \mathcal{S}^{\text{BD}}$ due to CSI estimation is bounded as

$$\begin{aligned} \Delta^{\text{BD}} &= \mathbb{E}\{g_k^{\text{BD}}(P) - g_k^{\text{BD-EST}}(P)\} \\ &\leq \log_2 \det \left(\mathbf{I}_M + \frac{P(N-M)}{N} (\mathbf{R}_{R,k}^{-T} + \frac{\mathbf{T}_k^H \mathbf{T}_k}{\sigma^2})^{-1} \right). \end{aligned} \tag{17}$$

B. Zero-Forcing Precoding with Receive Combining

When users are unaware of co-user channels, a reasonable combining strategy is to select $\mathbf{r}_k$ as the dominating left singular vector of $\mathbf{H}_k$ (to maximize $\|\mathbf{h}_k\|^2$). ZFC precoding is very similar to applying BD to the effective channels $\mathbf{h}_k^H = \mathbf{r}_k^H \mathbf{H}_k$, but an important difference is that the effective channels are not Rayleigh fading (since $\mathbf{r}_k$ is based on the current channel realization). This makes it difficult to derive the MMSE estimator, but fortunately the linear MMSE (LMMSE) estimator is also given by (15) [19]. To apply this LMMSE estimator, we need the statistics of $\mathbf{h}_k$.

Lemma 1. Assume that the transmit antennas are uncorrelated (i.e., $\mathbf{R}_T = \mathbf{I}_N$) and that $\mathbf{R}_R$ has eigenvalues $\lambda_M > \ldots > \lambda_1 > 0$, where the user indices was dropped for convenience. If $\mathbf{r}$ is the dominating left singular vector of $\mathbf{H}$, it holds that

- the direction $\mathbf{h}/\|\mathbf{h}\|$ is isotropically distributed on the unit sphere of $\mathbb{C}^N$;
- the gain $\|\mathbf{h}\|^2$ is independent of the direction and

[2]Digital/quantized feedback can offer substantial performance improvements over analog/unquantized feedback [18]. But if very accurate CSI is required, the quantization codebooks grow very large and the search for the best codeword becomes computationally infeasible.

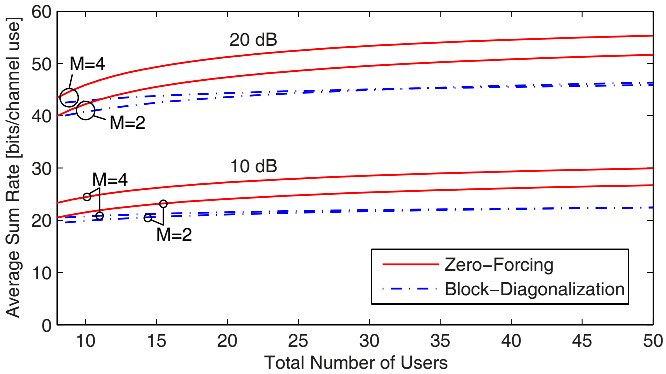

Fig. 3. The average sum rate with BD and ZFC in a system with $N = 8$ transmit antennas and different number of receive antennas. A low-complexity user selection algorithm from [17] is used.

$$\mathbb{E}\{\|\mathbf{h}\|^2\} = \sum_{l=1}^{M} \sum_{\boldsymbol{\alpha} \in \mathcal{A}_M} \frac{\prod_{\bar{k}=1}^{M} \lambda_{\alpha_{\bar{k}}}^{N-\bar{k}+1} \prod_{\bar{k}=N-M+1}^{N} (\bar{k}-1)!}{(-1)^{\text{per}(\boldsymbol{\alpha})+l+1} \det(\boldsymbol{\Delta})} \times \sum_{\boldsymbol{\beta} \in \mathcal{B}_{l,M}} \sum_{\bar{k}=0}^{\sum_{i=1}^{l}(N-\beta_l)} \sum_{\tilde{k} \in \widetilde{\Omega}_{\bar{k}}^{(l)}} \frac{\bar{k}!}{\tilde{k}_1! \cdots \tilde{k}_l!} \frac{\left(\sum_{i=1}^{l} \lambda_{\alpha_{\beta_i}}^{-1}\right)^{-(\bar{k}+1)}}{\prod_{i=1}^{l} \lambda_{\alpha_{\beta_i}}^{\tilde{k}_i}}, \tag{18}$$

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

$$[\boldsymbol{\Delta}]_{ij} = \lambda_j^{N-i+1} (N-i)!. \tag{19}$$

In (18), the set of all permutations of $\{1, \ldots, M\}$ is denoted $\mathcal{A}_M$. The sign of a given permutation $\boldsymbol{\alpha} = \{\alpha_1, \ldots, \alpha_M\} \in \mathcal{A}_M$ is denoted $(-1)^{\text{per}(\boldsymbol{\alpha})}$, where $\text{per}(\cdot)$ is the number of inversions 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., $\beta_1 < \ldots < \beta_l$ for $\boldsymbol{\beta} = \{\beta_1, \ldots, \beta_l\} \in \mathcal{B}_{l,M}$). Finally, $\widetilde{\Omega}_{\bar{k}}^{(l)}$ is the set of all l-length partitions $\{\tilde{k}_1, \ldots, \tilde{k}_l\}$ of $\bar{k}$ (i.e., $\sum_{i=1}^{l} \tilde{k}_i = \bar{k}$) that satisfy $0 \leq \tilde{k}_i \leq N - \beta_i$:

$$\widetilde{\Omega}_{\bar{k}}^{(l)} = \left\{ \{\tilde{k}_1, \ldots, \tilde{k}_l\} : \sum_{j=1}^{\bar{k}} \tilde{k}_j = \bar{k}, 0 \leq \tilde{k}_j \leq N - \beta_j \ \forall j \right\}. \tag{20}$$

As for BD precoding, we treat $\widehat{\mathbf{h}}_1, \ldots, \widehat{\mathbf{h}}_K$ as being the true channels and the resulting data rate of user $k \in \mathcal{S}^{\text{ZFC}}$ is

$$g_k^{\text{ZFC-EST}}(P) = \log_2 \left(1 + \frac{\frac{P}{|\mathcal{S}^{\text{ZFC}}|} |\mathbf{r}_k^H \mathbf{H}_k \widehat{\mathbf{v}}_k^{\text{ZFC}}|^2}{1 + \frac{P}{|\mathcal{S}^{\text{ZFC}}|} \sum_{\bar{k} \in \mathcal{S}^{\text{ZFC}} \setminus \{k\}} |\mathbf{r}_k^H \mathbf{H}_k \widehat{\mathbf{v}}_{\bar{k}}^{\text{ZFC}}|^2} \right). \tag{21}$$

Once the base station has fixed the precoding, user k can improve performance by applying the MMSE combiner in (13).[3] The following theorem provides an upper bound on the performance loss due to imperfect CSI estimation.

Theorem 4. Assume that the transmit antennas are uncorrelated (i.e., $\mathbf{R}_{T,k} = \mathbf{I}_N \ \forall k$) and that N users are scheduled randomly. The average rate loss for user $k \in \mathcal{S}^{\text{ZFC}}$ due to CSI estimation is bounded as (with $\mathbf{r}_k$ as in Lemma 1)

$$\begin{aligned} \Delta^{\text{ZFC}} &= \mathbb{E}\{g_k^{\text{ZFC}}(P) - g_k^{\text{ZFC-EST}}(P)\} \\ &\leq \log_2 \left(1 + \frac{P(N-1)}{N} \frac{1}{\mathbb{E}\{\|\mathbf{h}_k\|^2\}^{-1} + q/\sigma^2} \right). \end{aligned} \tag{22}$$

[3]The receive combiner can be updated with limited overhead, since the precoded channels already have to be estimated to enable coherent reception.

C. Comparison of BD and ZFC Under Imperfect CSI

The rate loss expressions in Theorem 3 and Theorem 4 indicate the impact of spatial correlation and estimation on the performance. It seems like BD is slightly more resilient to channel uncertainty, since the BD expression contains $(N - M)$ where the ZFC expression has $(N - 1)$. However, the important result is the following, which extends [18].

Corollary 1. To achieve the full multiplexing gain with BD or ZFC under arbitrary receive correlation, it is sufficient to scale the training power q as

$$P/q \to \texttt{constant} < \infty \quad \text{when } P \to \infty. \tag{23}$$

This corollary says that the training power should increase linearly with the transmit power to achieve the optimal sum rate scaling. This is typically satisfied in practice, where we have the same average SNR in the downlink and uplink. 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 [18]. As ZFC only has one effective antenna per user, it can accommodate M times more users than BD on the same estimation overhead. There is a natural connection to quantized CSI feedback, where the number of feedback bits should scale towards infinity (at a speed proportional to the increase in uplink capacity) to achieve the full multiplexing gain [10] and where only some users should send CSI to limit the feedback overhead [20].

Next, we compare BD and ZFC by repeating the simulation in Fig. 3 for the case of imperfect CSI estimation. The BD and ZFC schemes are implemented as described earlier in this section. We use training power $q = P$ and the capacity-based scheduling algorithm in [17] is modified to include the average interference due to CSI estimation errors. The average sum rate is shown in Fig. 4 as a function of the number of users that we obtain CSI estimates for under ZFC (while BD only obtains channel estimates for $1/M$ of them). We consider $N = 8$ transmit antennas, $M = 2$ or $M = 4$ receive antennas, and a total transmit power of $P = 10$ dB or $P = 20$ dB. The performance loss compared with having perfect CSI is below 10%, but the conclusions are otherwise the same: ZFC outperforms BD both in terms of performance with a few users and in exploiting multiuser diversity.

This conclusion stands in contrast to the numerical results in [10], where BD clearly beats ZFC under quantized CSI. To explain the difference, we repeat the simulation in [10, Fig. 6] with $N = 6$ transmit antennas and $M = 2$ receive antennas. The quantization codebooks are generated using random vector quantization (RVQ) and contain 10 bits/user under BD and 5 bits/user under ZFC. The sum rate is shown in Fig. 5 for the

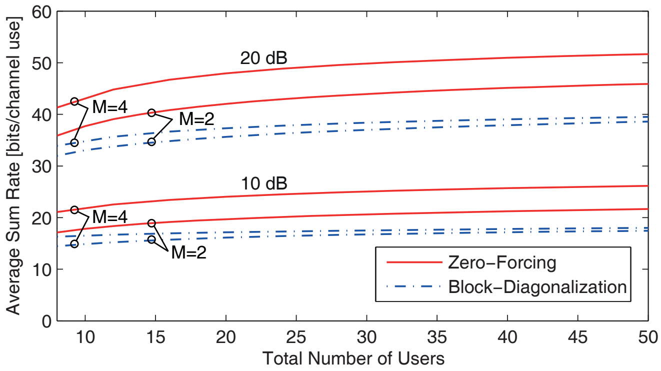

Fig. 4. The average sum rate with BD and ZFC under imperfect CSI estimation and using a low-complexity user selection algorithm.

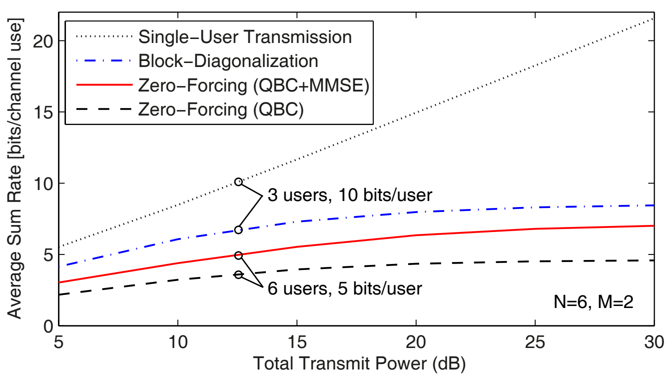

Fig. 5. Comparison of single-user transmission, BD, and different forms of ZFC under quantized CSI feedback. The scenario is the same as in [10, Fig. 6], where the superior single-user scheme was not considered.

quantized BD approach in [10] and the quantized ZF-QBC approach in [4]. We have also included an improved version of ZF-QBC where the MMSE combiner in (13) is applied during transmission and 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 schemes should be used in this scenario since the single-user transmission is vastly superior.

To summarize, ZFC achieves better performance than BD if the CSI accuracy is sufficient to support multiuser transmission. If the base station has inaccurate CSI (compared with the SNR at which the transmission takes place), single-user transmission should be used instead of BD and ZFC.

V. Conclusion

There are two modes in multiuser MIMO systems: single-user and multi-user transmission. When the base station has relatively accurate CSI (e.g., power for CSI estimation scales linearly with transmit power), multi-user transmission is beneficial as it can achieve the full multiplexing gain. In this scenario, we have analyzed whether the receive antennas at each user should be combined for achieving receive diversity (as with ZFC) or for multistream multiplexing (as with BD). The unexpected conclusion is that receive combining is the favorable choice, since ZFC is better at exploiting spatial correlation and multiuser diversity. This result has positive implications for the design of multiuser systems as it reduces the hardware constraints, especially in terms of processing at the user devices and the required CSI accuracy per user. When only inaccurate CSI can be achieved, the performance is limited by co-user interference making single-user transmission advantageous. Under these conditions, BD outperforms ZFC since it creates less co-user interference but this is unimportant since none of these two schemes should be used.

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] 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.

[3] 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.

[4] N. Jindal, "Antenna combining for the MIMO downlink channel," *IEEE Trans. Wireless Commun.*, vol. 7, no. 10, pp. 3834–3844, 2008.

[5] 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, pp. 1494–1504, 2008.

[6] 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.

[7] 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.

[8] 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.

[9] J. Zhang, R. Heath, M. Kountouris, and J. Andrews, "Mode switching for the multi-antenna broadcast channel based on delay and channel quantization," *EURASIP J. on Adv. in Signal Process.*, 2009.

[10] 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.

[11] 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.

[12] M. Dohler, R. Heath, A. Lozano, C. Papadias, and R. Valenzuela, "Is the PHY layer dead?" *IEEE Commun. Mag.*, vol. 49, pp. 159–165, 2011.

[13] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE J. Sel. Areas Commun.*, vol. 24, pp. 528–541, 2006.

[14] 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.

[15] S. Loyka, "Channel capacity of MIMO architecture using the exponential correlation matrix," *IEEE Commun. Lett.*, vol. 5, pp. 369–371, 2001.

[16] 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.

[17] 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.

[18] 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.

[19] 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.

[20] N. Ravindran and N. Jindal, "Multi-user diversity vs. accurate channel feedback for MIMO broadcast channels," in *Proc. IEEE ICC'08*, 2008, pp. 3684–3688.