Massive MIMO Performance-TDD Versus FDD: What Do Measurements Say?

Jose Flordelis, Fredrik Rusek, Fredrik Tufvesson, Erik G Larsson and Ove Edfors

The self-archived postprint version of this journal article is available at Linköping University Institutional Repository (DiVA):

<http://urn.kb.se/resolve?urn=urn:nbn:se:liu:diva-147603>

N.B.: When citing this work, cite the original publication.

Flordelis, J., Rusek, F., Tufvesson, F., Larsson, E. G, Edfors, O., (2018), Massive MIMO Performance-TDD Versus FDD: What Do Measurements Say?, *IEEE Transactions on Wireless Communications*, 17(4), 2247-2261. <https://doi.org/10.1109/TWC.2018.2790912>

Original publication available at:

<https://doi.org/10.1109/TWC.2018.2790912>

Copyright: Institute of Electrical and Electronics Engineers (IEEE)

<http://www.ieee.org/index.html>

©2018 IEEE. Personal use of this material is permitted. However, permission to reprint/republish this material for advertising or promotional purposes or for creating new collective works for resale or redistribution to servers or lists, or to reuse any copyrighted component of this work in other works must be obtained from the IEEE.

Massive MIMO Performance—TDD Versus FDD: What Do Measurements Say?

Jose Flordelis, *Student Member, IEEE*, Fredrik Rusek, *Member, IEEE*, Fredrik Tufvesson, *Fellow, IEEE*, Erik G. Larsson, *Fellow, IEEE*, and Ove Edfors, *Senior Member, IEEE*

Abstract—Downlink beamforming in Massive MIMO either relies on uplink pilot measurements—exploiting reciprocity and time-division duplexing (TDD) operation, or on the use of a predetermined grid of beams with user equipments reporting their preferred beams, mostly in frequency-division duplexing (FDD) operation. Massive MIMO in its originally conceived form uses the first strategy, with uplink pilots, whereas there is currently significant commercial interest in the second, grid-of-beams. It has been analytically shown that with isotropic scattering (independent Rayleigh fading) the first approach outperforms the second. Nevertheless, there remains controversy regarding their relative performance in practical channels. In this contribution, the performances of these two strategies are compared using measured channel data at 2.6 GHz.

Index Terms—channel measurements, FDD, Massive MIMO, performance, TDD.

I. INTRODUCTION

THE idea behind Massive MIMO is to equip base stations (BSs) in wireless networks with large arrays of phase-coherently cooperating antennas. The use of such arrays facilitates spatial multiplexing of many user equipments (UEs) in the same time-frequency resource, and yields a coherent beamforming gain that translates directly into reduced interference and improved cell-edge coverage.

The original Massive MIMO concept [1]–[4] assumes time-division duplexing (TDD) and exploits reciprocity for the acquisition of channel state information (CSI) at the BS. UEs send pilots on the uplink (UL); all UE-to-BS channels are estimated, and each antenna has its own RF electronics. The concept has, since its introduction a decade ago [1], [3], matured significantly: rigorous information-theoretic analyses are available [2], field-trials have demonstrated its performance in scenarios with moderate mobility [5]–[7], and circuit prototypes have shown the true practicality of implementations [8].

Concurrently, motivated by spectrum regulation issues, there is significant interest in developing frequency-division duplexing (FDD) versions of Massive MIMO [9]–[15]. [12], [14]

exploit the structure of the spatial correlation of the channel to reduce the amount of CSI that needs to be fed back from the UEs to the BS. Techniques for downlink (DL) training and reporting of CSI for the single-user setting, using both spatial and temporal correlations, are presented in [10], [11]. More recently, [15] has extended [10] to the multiuser setting. These methods can reduce the performance gap between FDD and TDD systems provided that channel correlations are sufficiently strong. Moreover, knowledge of the spatial covariance matrix at either the BS [12], [14], [15] or the UEs [10], [11] is assumed. It is not an easy task to facilitate this due to the large number of antennas in Massive MIMO systems. There is also interest in hybrid beamforming architectures that rely on the use of analog phase shifters and signal combiners [16]–[20], somewhat reminiscent of phased-arrays implementations of radar. With hybrid beamforming, the number of actual antennas may substantially exceed the number of RF chains. In [16], [17], the required number of RF chains to approach the performance of fully-digital schemes is considered. Furthermore, the performance of hybrid beamforming architectures using beamsteering codebooks is analyzed in [18] and [20] for single- and multiuser settings, respectively. In general, these methods work well if the channel is known to be sparse.

FDD operation and hybrid beamforming solutions both bring the same difficulty—albeit for different reasons: significant assumptions on the structure of propagation must be made for the techniques to work efficiently. Specifically:

- FDD operation requires CSI feedback from the UEs to the BS. Efficient encoding of this CSI is only possible if side information on the propagation is exploited. The resulting techniques are often called “grid-of-beams”, and have similarities to existing forms of multiuser (MU) MIMO in LTE [21].
- Hybrid-beamforming architectures inherently rely on beamforming into predetermined spatial directions, as defined by the angle-of-arrival or angle-of-departure, seen from the array. Such directions only have a well-defined operational meaning when the propagation environment offers strong direct or specular paths [22].

Whether or not practical channels possess the type of structure that is commonly assumed by FDD schemes is the overall theme of this work.

There is a long-standing debate on the relative performance of reciprocity-based (TDD) Massive MIMO and that of solutions based on grid-of-beams or hybrid-beamforming architectures with practical channels. On the one hand, the commercial

This work was supported by the Seventh Framework Programme (FP7) of the European Union under grant agreement no. 619086 (MAMMOET), ELLIT—an Excellence Center at Linköping-Lund in Information Technology, the Swedish Research Council (VR), and the Swedish Foundation for Strategic Research (SSF).

Jose Flordelis, Fredrik Rusek, Fredrik Tufvesson, and Ove Edfors are with the Department of Electrical and Information Technology, Lund University, SE-221 00 Lund, Sweden (e-mail: jose.flordelis@eit.lth.se; fredrik.rusek@eit.lth.se; fredrik.tufvesson@eit.lth.se; ove.edfors@eit.lth.se).

E. G. Larsson is with the Department of Electrical Engineering (ISY), Linköping University, SE-581 83 Linköping, Sweden (e-mail: erik.g.larsson@liu.se).

arguments for grid-of-beams solutions are clear [13], [19], but on the other hand, their real potential for high-performance spatial multiplexing has been contested [23], [24]. It is known that grid-of-beams solutions perform poorly in isotropic scattering [24], but no prior experimental results are known to the authors.

The objective of this paper is to conclusively answer this performance question in practical channels. For that we consider measurement data of real Massive MIMO channels in the 2.6 GHz band, and analyze the achievable sum-rates using beamsteering codebooks. The conclusion, summarized in detail in Sec. VI, is that except for in certain line-of-sight (LOS) environments, the original reciprocity-based TDD Massive MIMO of [1], [3] represents the only feasible implementation of Massive MIMO at the frequency bands under consideration. In other considered scenarios, the performance loss is significant for the non reciprocity-based beamforming solutions.

A. Notation

We use the following notation throughout the paper: Boldface lowercase letters represent column vectors, and boldface uppercase letters represent matrices. Using this notation, $\mathbf{I}$ is the identity matrix, $\|\mathbf{a}\|$ the Euclidean norm of vector $\mathbf{a}$, $\text{tr}(\mathbf{A})$ the trace of matrix $\mathbf{A}$, $\text{span}(\mathbf{A})$ its column space, $\mathbf{A}^T$ denotes the transpose, $\mathbf{A}^H$ the Hermitian transpose, $|\mathbf{A}|$ stands for the determinant, $\mathbf{A} \succeq \mathbf{0}$ means that $\mathbf{A}$ is positive semidefinite, $\text{diag}(\mathbf{a})$ builds a matrix having $\mathbf{a}$ along its diagonal and all other elements set to zero, $[\mathbf{A} \mid \mathbf{b}]$ denotes the matrix resulting from appending $\mathbf{b}$ to $\mathbf{A}$, and $[\mathbf{A}]_{\mathcal{I}}$ is the submatrix of $\mathbf{A}$ formed by choosing the columns of the index set $\mathcal{I}$. Furthermore, j denotes the imaginary unit, $\mathcal{CN}(\boldsymbol{\mu}, \boldsymbol{\Lambda})$ the complex Gaussian distribution with mean $\boldsymbol{\mu}$ and covariance matrix $\boldsymbol{\Lambda}$, $\mathbb{E}[\cdot]$ is the expectation operator, and $|\mathcal{I}|$ denotes the number of elements in the set $\mathcal{I}$.

II. SYSTEM MODEL

We consider the DL of a single-cell Massive MIMO system in which an M -antenna BS communicates with K single-antenna UEs in the same time-frequency resource. Orthogonal Frequency Division Multiplexing (OFDM) with L subcarriers is assumed [25]. Let $\mathbf{h}_k(\ell) \in \mathbb{C}^{M \times 1}$, for $k = 1, \dots, K$, and $\ell = 1, \dots, L$, denote the channel vector between the BS and the k^{th} UE at the ℓ^{th} subcarrier, and let

$$\mathbf{H}(\ell) = [\mathbf{h}_1(\ell) \quad \dots \quad \mathbf{h}_K(\ell)]^T \quad (1)$$

denote the corresponding $K \times M$ channel matrix. We do not place any constraints on the entries of the channel vectors $\mathbf{h}_k(\ell)$, but the average gain $\mathbb{E}\{\frac{1}{M}\|\mathbf{h}_k(\ell)\|^2\}$ to one. Then, the input-output relation of the system can be written as

$$\mathbf{y}(\ell) = \sqrt{\rho} \mathbf{H}(\ell) \mathbf{s}(\ell) + \mathbf{n}(\ell), \quad (2)$$

where $\mathbf{y}(\ell) \in \mathbb{C}^{K \times 1}$ is the vector containing the received signals of all the UEs, $\mathbf{s}(\ell) \in \mathbb{C}^{M \times 1}$ the vector of precoded transmit signals satisfying

$$\mathbb{E}\{\mathbf{s}^H(\ell) \mathbf{s}(\ell)\} = 1, \quad (3)$$

Fig. 1. Fully-digital reciprocity-based (TDD) beamforming with $K = 2$ single-antenna UEs.

ρ the signal-to-noise ratio (SNR), and $\mathbf{n}(\ell)$ is a vector of $\mathcal{CN}(0, 1)$ receiver noise at the UEs.

III. TRANSMISSION TECHNIQUES

This section outlines the beamforming techniques included in the comparison—first, fully-digital reciprocity-based (TDD) beamforming in Sec. III-A, and then, four flavors of FDD beamforming based on feedback of CSI in Sec. III-B.

A. Fully-Digital Reciprocity-Based (TDD) Beamforming

With fully-digital beamforming, no a priori assumptions are made on the propagation environment. There are no predetermined beams, but CSI is measured at the BS by observing UL pilots transmitted by the UEs. By virtue of TDD operation and reciprocity of propagation, the so-obtained UL CSI is also valid for the DL, assuming proper reciprocity calibration [26]. All signal processing takes place in the digital domain. A TDD beamforming system is schematically depicted in Fig. 1 for $K = 2$ UEs.

With full CSI at the BS, TDD performs optimally and can achieve the DL sum-capacity of the channel by dirty-paper coding (DPC) [27]. For given ρ , the sum-capacity of the ℓ^{th} subcarrier, $\mathcal{C}_{\text{TDD}}(\mathbf{H}(\ell), \rho)$, is given by the solution to the following optimization problem [28]–[31]:

$$\begin{aligned} & \underset{\boldsymbol{\Lambda}(\ell)}{\text{maximize}} \quad \log_2 \left| \mathbf{I} + \mathbf{H}^H(\ell) \boldsymbol{\Lambda}(\ell) \mathbf{H}(\ell) \right| \\ & \text{subject to} \quad \text{tr}(\boldsymbol{\Lambda}(\ell)) \leq \rho, \quad \boldsymbol{\Lambda}(\ell) \succeq \mathbf{0}, \end{aligned} \quad (4)$$

where $\boldsymbol{\Lambda}(\ell) = \text{diag}(\lambda_1(\ell), \dots, \lambda_K(\ell))$ is a diagonal power allocation matrix. The sum-capacity averaged over all the subcarriers is then

$$\bar{\mathcal{C}}_{\text{TDD}}(\rho) = \frac{1}{L} \sum_{\ell=1}^L \mathcal{C}_{\text{TDD}}(\mathbf{H}(\ell), \rho). \quad (5)$$

Problem (4) is convex and can be efficiently solved by a simple gradient search, or via a technique known as sum-power iterative waterfilling [32], [33].

A main focus of this work is to compare the performance of the various techniques of DL beamforming based on the best performance that can be extracted from them. We shall therefore assume that perfect and instantaneous CSI is available at the BS for TDD beamforming (as we will shortly see, perfect

CSI assumptions are also made for FDD beamforming). For a discussion of the effect of imperfect CSI on the sum-rates of TDD Massive MIMO systems the reader is referred to [2], which provides strict bounds on UL and DL sum-rates taking into account channel estimation errors (see Table 3.1 in [2]). In the context of our work, imperfect CSI is considered in Example 3, in Sec. V.

B. Feedback-Based FDD Beamforming with Predetermined Beams

Feedback-based beamforming relies on the reporting of quantized CSI from the UEs to the BS. Typically, CSI quantization is obtained by using a predetermined codebook consisting of M' beams, which imposes a certain structure on the precoded signals $s(\ell)$. These techniques may be applied when reliance on reciprocity is undesirable or impossible, notably in FDD operation [2], [34], [35].

We represent the M' beams through the set of M -vectors $\{\mathbf{c}_m\}_{m=1}^{M'}$. Throughout this article, we assume that these beams are given by Vandermonde vectors comprising the array response in M' directions uniformly spaced in the sine-angle domain.¹ More precisely, we define

$$\mathbf{c}_m = \frac{1}{\sqrt{M}} [1 \ e^{j\pi\psi_m} \ \dots \ e^{j\pi\psi_m(M-1)}]^T, \quad (6)$$

where $\psi_m = -1 + \frac{2m-1}{M'}$, for $m = 1, \dots, M'$. We also define the $M \times M'$ codebook matrix

$$\mathbf{C} = [\mathbf{c}_1 \ \dots \ \mathbf{c}_{M'}]. \quad (7)$$

A special case of the codebook is when $M' = M$ and the beams are orthonormal; then $\mathbf{C}^H \mathbf{C} = \mathbf{I}$. In this case, the vectors $\mathbf{c}_m$ are the columns of an $M \times M$ IDFT matrix, up to a constant shift of the origin of the phase angle ψ_m .

The UEs report their preferred beams to the BS. There are several ways that this may be done, and we consider two cases:

- 1) Each UE individually reports the indices and complex gains of a predetermined number, $N \leq M'$, of beams.
- 2) The BS, possibly based on interaction with the UEs, decides on a common set of N beams that are simultaneously used for all the UEs. Then, each UE reports the complex gains of these N beams.

In the multiuser setting, 2) potentially entails more signaling overhead than 1), since only the BS can compute the optimal set of beams, which then has to be signaled to the UEs. This is because a given UE is not aware of other UEs' channels. On the other hand, 2) can effectively combat multiuser interference, since the BS knows each UE's channel restricted to the selected subspace. It is therefore interesting to compare the achievable sum-rates of 1) and 2) above.

The structure imposed by the predetermined codebook of beams may be implemented either in the digital domain, or in the analog domain:

¹Basically, we select a beamsteering codebook that exploits the directivity of the channel. Other codebook choices are possible. For instance, if all the UEs have the same covariance matrix $\mathbf{R} = \mathbb{E}\{\mathbf{h}\mathbf{h}^H\}$ (possibly time- and frequency-varying), one can then set $\mathbf{C} = \mathbf{U}$, where $\mathbf{U}$ contains the eigenvectors of $\mathbf{R}$ corresponding to the non-zero eigenvalues. Essentially, this codebook yields the communication scheme introduced in [14].

- a) If implemented in the digital domain, the selection of the beams may be performed individually for each subcarrier.
- b) In contrast, if implemented in the analog domain, the same set of beams must be used for the entire band.

Thus, a) and b) allow us to examine the frequency selectivity of 1 and 2 above. In isotropic scattering, different beams may be selected every coherence bandwidth interval. On the other hand, for perfectly specular channels such that the angle-of-arrival of each multipath component (MPC) corresponds (exactly) to some ψ_m , frequency selectivity can be avoided whenever N is equal to or exceeds the number of MPCs. Practical channels fall somewhere in between these two extremes.

The combination of 1) and 2), respectively a) and b) above, yields four cases of interest, illustrated in Fig. 2 for $K = 2$ single-antenna UEs and $N = 2$ reported beams. These four cases are described in detail in the next four subsections. As the number of antennas increases dramatically in Massive MIMO systems, DL training in FDD operation with limited pilot and feedback overhead becomes challenging. To tackle this issue, various methods for acquisition of DL CSI have been recently proposed [10], [11], [14]. Evaluating the performance of specific DL training methods is, however, beyond the scope of this paper. In order to be as generous as possible with FDD DL beamforming techniques, throughout the paper we assume that perfect and instantaneous CSI is available to the UEs for estimating the beams and the complex gains. (As previously pointed out, in TDD operation, we shall assume perfect and instantaneous CSI at the BS.) Moreover, we assume that the feedback channels are delay- and error-free [10], [36].

Digital Grid-of-Beams (D-GOB): Each UE individually reports the indices and complex gains of a number, N , of beams. The selection and reporting of the beams is done independently for each subcarrier. This corresponds to the combination of 1) and a) above.

Let us start by computing the achievable sum-rate of D-GOB averaged over all the subcarriers, $\bar{C}_{\text{D-GOB}}(\rho)$. Each UE learns the vector of complex gains

$$\mathbf{g}_k(\ell) = \mathbf{C}^T \mathbf{h}_k(\ell), \quad (8)$$

of the M' predetermined beams. It then selects N beams, according to some criterion that will be shortly explained, and forms the set $\mathcal{Q}_k(\ell)$ of selected beam indices. Next, each UE reports $\mathcal{Q}_k(\ell)$ and the vector $\check{\mathbf{g}}_k(\ell)$ of associated complex gains to the BS. By construction, we have

$$\check{\mathbf{g}}_k(\ell) = \mathbf{B}_k^T(\ell) \mathbf{h}_k(\ell), \quad (9)$$

where the $M \times N$ matrix $\mathbf{B}_k(\ell)$ is obtained by extracting the relevant beams from $\mathbf{C}$, as dictated by $\mathcal{Q}_k(\ell)$. Accordingly, the BS may produce a version $\hat{\mathbf{h}}_k(\ell)$ of $\mathbf{h}_k(\ell)$ as given by

$$\hat{\mathbf{h}}_k(\ell) = \arg \min_{\mathbf{v} \in \text{span}(\mathbf{B}_k(\ell))} \|\mathbf{B}_k^T(\ell) \mathbf{v} - \check{\mathbf{g}}_k(\ell)\|^2. \quad (10)$$

With D-GOB, multiuser interference is in general not fully known at the BS. This is because for given UEs i and j , $i \neq j$, the sets $\mathcal{Q}_i(\ell)$ and $\mathcal{Q}_j(\ell)$ of reported beams may be different from each other. In this case, BS knowledge of the

Fig. 2. The four considered cases of feedback-based FDD beamforming. A Massive MIMO BS communicates with $K = 2$ single-antenna UEs, each reporting on $N = 2$ beams picked from a codebook of size $M' = 4$. N' is the number of RF chains. In this example, with D-GOB and H-GOB, UE 1 selects beams $\mathbf{a}$ and $\mathbf{b}$, and UE 2 selects beams $\mathbf{b}$ and $\mathbf{c}$; with D-SUB and H-SUB, UE 1 and UE 2 report on the common subspace spanned by both beams $\mathbf{a}$ and $\mathbf{b}$.

mutual interference between the two UEs is limited to the set $\mathcal{Q}_i(\ell) \cap \mathcal{Q}_j(\ell)$ of common reported beams. Since DPC requires perfect knowledge of multiuser interference, it is not feasible in the D-GOB setting. In this work, we use zero-forcing (ZF) based on the estimated channels $\hat{\mathbf{h}}_k(\ell)$, first proposed in [36], [37], as the multiuser transmission strategy. To apply ZF, we define the compound estimated channel matrix

$$\hat{\mathbf{H}}(\ell) = [\hat{\mathbf{h}}_1(\ell) \cdots \hat{\mathbf{h}}_K(\ell)]^T. \quad (11)$$

Then, from [34], the columns of the ZF precoding matrix, $\mathbf{P}(\ell)$ in Fig. 2, can be computed as

$$\mathbf{p}_k(\ell) = \mathbf{z}_k(\ell) / \|\mathbf{z}_k(\ell)\|,$$

where $\mathbf{z}_k(\ell)$ are the columns of the Moore-Penrose pseudoinverse $\hat{\mathbf{H}}^\dagger(\ell)$ of $\hat{\mathbf{H}}(\ell)$. If equal power ρ/K is allocated to each UE, the receive SINR of the k^{th} UE can be written as

$$\text{SINR}_k(\mathbf{H}(\ell), \rho) = \frac{\frac{\rho}{K} |\mathbf{h}_k^T(\ell) \mathbf{p}_k(\ell)|^2}{1 + \frac{\rho}{K} \sum_{i \neq k} |\mathbf{h}_k^T(\ell) \mathbf{p}_i(\ell)|^2}, \quad (12)$$

from which the achievable sum-rate is computed as [34]

$$\mathcal{C}_{\text{D-GOB}}(\mathbf{H}(\ell), \rho) = \sum_{k=1}^K \log_2 \left(1 + \text{SINR}_k(\mathbf{H}(\ell), \rho) \right). \quad (13)$$

The sum-rate averaged over all the subcarriers, $\bar{\mathcal{C}}_{\text{D-GOB}}(\rho)$, is then defined similar to (5). Note that even though the precoders $\mathbf{P}(\ell)$ are designed according to the ZF principle,

the multiuser cross-talk terms $|\mathbf{h}_k^T(\ell) \mathbf{p}_i(\ell)|^2$, $i \neq k$, in the denominator of (12) do not vanish in general. In fact, precoding that completely suppresses interference is impossible here since complete CSI cannot be obtained at the BS, unless $N = \min(M', M)$.

Next, we briefly discuss the problem of beam selection by the UEs, which we formulate as the solution to the optimization problem [18], [36], [38]

$$\begin{aligned} \arg \min_{\mathcal{Q}_k(\ell)} \quad & \|\mathbf{h}_k(\ell) - \hat{\mathbf{h}}_k(\ell)\|^2 \\ \text{subject to} \quad & \mathcal{Q}_k(\ell) \subset \{1, \dots, M'\}, \quad |\mathcal{Q}_k(\ell)| = N, \end{aligned} \quad (14)$$

where $\hat{\mathbf{h}}_k(\ell)$ depends on $\mathcal{Q}_k(\ell)$ through $\mathbf{B}_k(\ell)$ as given by (10). Thus, the strategy is to minimize the reconstruction error of the channel vectors. Generally, (14) is a hard combinatorial problem, and can be solved exactly only for fairly small values of N . (A special case is when $\mathbf{C}^H \mathbf{C} = \mathbf{I}$, in which case one simply needs to pick the N strongest entries in the vector $\mathbf{g}_k(\ell)$ defined by (8).) Because of this, a heuristic rather than optimal algorithm to solve (14) is favored in this work. For the particulars on the algorithm, the reader is referred to Appendix A.

Digital Subspace Beamforming (D-SUB): The BS, possibly based on interaction with the UEs, decides on a common set of N beams that are used for all the UEs. Beams are selected independently for each subcarrier. Thus, we have the combination of 2) and a) above.

We seek to find a beamforming matrix $\mathbf{B}(\ell)$, formed from the columns of $\mathbf{C}$, such that the resulting channel $\mathbf{H}(\ell)\mathbf{B}(\ell)$ maximizes the sum-rate for given ρ . Let $\mathcal{C}_{\text{D-SUB}}(\mathbf{H}(\ell), \rho)$ denote the optimal sum-rate. The structure of D-SUB beamforming is shown in Fig. 2. Clearly, the precoder $\mathbf{P}(\ell)$ needs to be designed jointly with $\mathbf{B}(\ell)$. For this, we adopt a two-step approach. First, we address the problem of designing $\mathbf{P}(\ell)$ when $\mathbf{B}(\ell)$ and ρ are given. Then, we return to the original problem of jointly designing $\mathbf{P}(\ell)$ and $\mathbf{B}(\ell)$ for given ρ , and apply the results of the first step.

For given $\mathbf{B}(\ell)$ and ρ , let $\mathcal{C}_{\text{BC}}(\mathbf{H}(\ell)\mathbf{B}(\ell), \rho)$ denote the maximum sum-rate of the MIMO broadcast channel (BC) $\mathbf{H}(\ell)\mathbf{B}(\ell)$. It is shown in Appendix B that $\mathcal{C}_{\text{BC}}(\mathbf{H}(\ell)\mathbf{B}(\ell), \rho)$ can be found as the solution to the optimization problem

$$\begin{aligned} & \underset{\mathbf{\Lambda}(\ell)}{\text{maximize}} \quad \log_2 \left| \mathbf{I} + \mathbf{U}^H(\ell) \mathbf{H}^H(\ell) \mathbf{\Lambda}(\ell) \mathbf{H}(\ell) \mathbf{U}(\ell) \right| \\ & \text{subject to} \quad \mathbf{\Lambda}(\ell) \succeq 0, \quad \text{tr}(\mathbf{\Lambda}(\ell)) \leq \rho, \end{aligned} \quad (15)$$

where $\mathbf{\Lambda}(\ell) = \text{diag}(\lambda_1(\ell), \dots, \lambda_K(\ell))$ is a diagonal power allocation matrix, and $\mathbf{U}(\ell)$ is an $M \times N$ matrix such that $\mathbf{B}(\ell) = \mathbf{U}(\ell)\mathbf{L}(\ell)$ with $\mathbf{U}^H(\ell)\mathbf{U}(\ell) = \mathbf{I}$, and $\mathbf{L}(\ell)$ an invertible matrix. If one defines the *effective* channel matrix $\tilde{\mathbf{H}}(\ell) = \mathbf{H}(\ell)\mathbf{U}(\ell)$, problem (15) is formally identical to (4), and hence can be solved efficiently. The optimal precoder $\mathbf{P}(\ell)$ for given $\mathbf{B}(\ell)$ and ρ is defined by the set of covariance matrices $\{\mathbf{Q}_{ij}\}_{i=1}^K$ found by (i) obtaining the effective covariance matrices $\{\tilde{\mathbf{Q}}_i(\ell)\}_{i=1}^K$ from the power allocations $\{\lambda_i(\ell)\}_{i=1}^K$ in (15) via the so-called ‘‘MAC-to-BC’’ transformation (described in, e.g., [29], [32]); and (ii) computing $\mathbf{Q}_i(\ell) = \mathbf{L}^{-1}(\ell) \tilde{\mathbf{Q}}_i(\ell) (\mathbf{L}^H(\ell))^{-1}$, $i = 1, \dots, K$.

Returning to our original problem, we can now express $\mathcal{C}_{\text{D-SUB}}(\mathbf{H}(\ell), \rho)$ as the solution to the optimization problem

$$\begin{aligned} & \underset{\mathbf{B}(\ell) \in [\mathbf{C}]_{\mathcal{Q}(\ell)}}{\text{maximize}} \quad \mathcal{C}_{\text{BC}}(\mathbf{H}(\ell)\mathbf{B}(\ell), \rho) \\ & \text{subject to} \quad \mathcal{Q}(\ell) \subset \{1, \dots, M'\}, \quad |\mathcal{Q}(\ell)| = N. \end{aligned} \quad (16)$$

Put in words, for each subcarrier, the sum-rate as given by (15) is maximized over all $M \times N$ beamformers $\mathbf{B}(\ell)$ generated by codebook $\mathbf{C}$. The sum-rate averaged over all the subcarriers, $\bar{\mathcal{C}}_{\text{D-SUB}}(\rho)$, is then defined similar to (5).

Although, in principle, one could attempt the maximization in (16) by exhaustive search, solving (15) at each step, the number of beamformers $\mathbf{B}(\ell)$ that needs to be checked with this approach is $\binom{M'}{N}$. Thus, for values of M' in the hundreds or larger, the above direct approach appears intractable, except for very small N . Therefore, alternative methods for solving (16) are needed. An efficient algorithm for approximate solution of (16) is presented in Appendix C.

Hybrid Subspace Beamforming (H-SUB): The BS, possibly based on interaction with the UEs, decides on a common set of N beams to service all the UEs. In contrast to D-SUB, this choice is applied across all subcarriers, thereby facilitating the implementation of the beamforming in analog hardware. This corresponds to the combination of 2) and b) above.

The hybrid beamforming architecture is shown in Fig. 2. The vector of precoded transmit signals, $\mathbf{s}(\ell)$, has the form

$$\mathbf{s}(\ell) = \mathbf{B}\mathbf{P}(\ell)\mathbf{x}(\ell), \quad \ell = 1, \dots, L,$$

where $\mathbf{x}(\ell)$ is a vector containing the information bits from the UEs satisfying $\mathbb{E} \{ \mathbf{x}(\ell)\mathbf{x}(\ell)^H \} = \mathbf{I}$. Importantly, the precoder $\mathbf{P}(\ell)$ is frequency-selective, but the beamforming matrix $\mathbf{B}$ is not. Hence, $\mathbf{B}$ can be realized entirely by analog hardware. An important consequence is that the number of required RF chains at the BS can be reduced from M (i.e., one RF chain per antenna element) to N (i.e., one RF chain per selected beam).

To obtain a cost-effective analog beamforming network, a certain structure is typically enforced on the matrix $\mathbf{B}$. In this work, we require that $\mathbf{B}$ be formed from the columns of the codebook matrix $\mathbf{C}$ defined by (7). Under this constraint, the analog beamforming network defined by $\mathbf{B}$ can be realized by using N phase shifters, and M N -input signal combiners, as depicted in Fig. 2. Other constraints on $\mathbf{B}$ leading to simplifications of the analog hardware are possible; the reader is referred to [22], [39] for a comprehensive survey of the field.

Optimal beam selection for H-SUB is analogous to D-SUB, except that beams are reused for all subcarriers. For given ρ , the sum-rate averaged over all subcarriers, $\bar{\mathcal{C}}_{\text{H-SUB}}(\rho)$, can be found as the solution to the optimization problem

$$\begin{aligned} & \underset{\mathbf{B} \in [\mathbf{C}]_{\mathcal{Q}}}{\text{maximize}} \quad \bar{\mathcal{C}}_{\text{H-SUB}}(\{\mathbf{H}(\ell)\mathbf{B}\}_{\ell=1}^L, \rho) \\ & \text{subject to} \quad \mathcal{Q} \subset \{1, \dots, M'\}, \quad |\mathcal{Q}| = N. \end{aligned} \quad (17)$$

where $\bar{\mathcal{C}}_{\text{H-SUB}}(\{\mathbf{H}(\ell)\mathbf{B}\}_{\ell=1}^L, \rho)$ is in turn defined as the solution to

$$\begin{aligned} & \underset{\{\mathbf{\Lambda}(\ell)\}_{\ell=1}^L}{\text{maximize}} \quad \frac{1}{L} \sum_{\ell=1}^L \log_2 \left| \mathbf{I} + \mathbf{U}^H \mathbf{H}^H(\ell) \mathbf{\Lambda}(\ell) \mathbf{H}(\ell) \mathbf{U} \right| \\ & \text{subject to} \quad \mathbf{\Lambda}(\ell) \succeq 0, \quad \text{tr}(\mathbf{\Lambda}(\ell)) \leq \rho, \end{aligned} \quad (18)$$

where, as usual, $\mathbf{\Lambda}(\ell) = \text{diag}(\lambda_1(\ell), \dots, \lambda_K(\ell))$ are power allocation matrices, and $\mathbf{B} = \mathbf{U}\mathbf{L}$ with $\mathbf{U}^H\mathbf{U} = \mathbf{I}$, and $\mathbf{L}$ an invertible matrix. Again, the efficient algorithm proposed in Appendix C can be used to solve (17).²

Hybrid Grid-of-Beams (H-GOB): Last, we have the combination of 1) and b), wherein similar to D-GOB, each UE individually reports the indices and complex gains of N beams, but wherein the choice of the beams is applied across all subcarriers. This strategy enables the implementation of the beamforming in analog hardware, as illustrated in Fig. 2. A special case is when $N = 1$, and additionally one dispenses with all the digital signal processing. This case is sometimes referred to as analog-only beamforming, and is used in communication standards such as IEEE 802.11ad [40]. The problem of beam selection can be posed as the following optimization problem:

$$\begin{aligned} & \arg \min_{\mathcal{Q}_k} \quad \frac{1}{L} \sum_{\ell=1}^L \|\mathbf{h}_k(\ell) - \hat{\mathbf{h}}_k(\ell)\|^2 \\ & \text{subject to} \quad \mathcal{Q}_k \subset \{1, \dots, M'\}, \quad |\mathcal{Q}_k| = N, \end{aligned} \quad (19)$$

²Note that, whenever perfect CSI is available at the BS, the average sum-rate of TDD beamforming with a hybrid architecture analogous to H-SUB is also given by $\bar{\mathcal{C}}_{\text{H-SUB}}(\rho)$. In this sense, the sum-rates of H-SUB can be taken as a benchmark for the maximum achievable sum-rates of the hybrid approach to TDD.

where $\hat{h}_k(\ell)$ is given by (10). The heuristic algorithm in Appendix A (with minor modifications) is proposed for solving (19).

IV. MEASURED CHANNELS

The measured channels were obtained in two different measurement campaigns conducted at the Faculty of Engineering (LTH) of Lund University, Lund, Sweden. At the BS side, a virtual uniform linear array (ULA) with 128 elements was used.³ The ULA spans 7 meters, and uses vertically-polarized, omnidirectional-in-azimuth antenna elements [35].⁴ At the UE side, vertically-polarized omnidirectional antennas of the same type were used. The measurements were acquired at a carrier frequency of 2.6 GHz, and using a bandwidth of 50 MHz. A brief description of the two campaigns and the scenarios follows:

- **Campaign A.** The UEs were located at the parking place outside the E-building of LTH, with the ULA mounted on top of the E-building, three floors above ground level. We consider five UE sites, denoted MS 1, ..., MS 5. Sites MS 1 to MS 4 have mainly LOS propagation conditions to the BS, although the LOS of site MS 4 is partially blocked by the roof edge, thereby giving rise to propagation by diffraction [5]; site MS 5 experiences NLOS. At each site, several UE locations are measured. In this work, we consider three propagation scenarios, summarized in Table I as scenarios 1, 2, and 3. For further details on Campaign A, the reader is referred to [5].
- **Campaign B.** The UEs were located in a courtyard of the E-building. The ULA was on a roof two floors above ground, while the 16 UEs were spread out at various positions in the courtyard. In this environment, the UEs experience LOS propagation conditions to the array, along with a number of strong scattered components arising from interactions with the enclosing walls, outdoor furniture (to a large extent consisting of metallic items), and vegetation. Because of this, the Ricean K -factor [22], [41], [42] is low compared to sites MS 1–MS 3 in campaign A; see Table II. In this work, we consider three propagation scenarios related to campaign B, summarized in Table I as scenarios 4, 5, and 6. For further details on Campaign B, the reader is referred to [43].

We should also mention that, prior to applying DL beamforming as described in Sec. III, the measured channels are normalized to have unit average gain, i.e., we define

$$\mathbf{h}_k(\ell) = \mathbf{h}_k^{\text{raw}}(\ell) \cdot \sqrt{\frac{L \cdot 128}{\sum_{\ell'=1}^L \|\mathbf{h}_k^{\text{raw}}(\ell')\|^2}}, \quad k = 1, \dots, K,$$

³Other array geometries can be considered. For example, for urban environments with tall buildings, two-dimensional arrays exploiting the elevation plane may be of interest [9]. In this investigation, however, we restrict consideration to the ULA case.

⁴Omnidirectional antennas at the BS can excite all the codewords in (7). In practical cellular deployments, however, it might be convenient to use directional antennas so that only parts of (7) are excited. For example, in a three-sector macrocell deployment, one might wish to use only those codewords in (7) that illuminate a certain sector. Nevertheless, due to typically limited angular spreads at the BS, the results presented in this paper can reasonably be expected to apply, for the same array aperture.

where $\mathbf{h}_k^{\text{raw}}(\ell)$ are the measured channels. This normalization step removes differences in path loss among UEs, while preserving variations across frequencies and antenna positions.

V. RESULTS AND DISCUSSION

Based on the measured channels obtained from Campaign A and Campaign B, we compare the performance of the five beamforming techniques described in Sec. III, namely, TDD beamforming,⁵ and four flavors of FDD beamforming: D-GOB, H-GOB, D-SUB, and H-SUB. Because TDD performs optimally, it serves as baseline. First, in Sec. V-A, we study how much one can reduce the number, N , of reported beams in FDD beamforming while still retaining a prescribed fraction of the sum-rate of TDD, i.e., the sum-capacity of the channel. Next, in Sec. V-B, we fix the FDD sum-rate to a desired value and address the following question: “Given N , what is the average SNR loss relative to optimal TDD?” Then, in Sec. V-C, we investigate the tradeoff between RF chains and BS antennas in FDD beamforming, subject to a sum-rate constraint. In Sec. V-D, we reevaluate the findings of Sec. V-A, but now including the overhead of DL training. Then, in Sec. V-E, we make a remark about analog-only beamforming, and in Sec. V-F we briefly discuss the applicability of our results to millimeter-wave (mm-wave) frequencies.

In the preparation of the results reported below, the following parameter settings were used. There are $M = 128$ antennas at the BS, which communicate with $K = 4, 8$, and 16 single-antenna UEs, depending on the particular scenario. Evaluations are done based on $L = 71$ subcarriers equispaced over a 50 MHz bandwidth. The value $L = 71$ has been selected because it ensures that the flat-frequency assumption holds (that is, that the subcarrier spacing is smaller than the coherence bandwidth of the channel) while keeping the number of subcarriers reasonably small. For each of the four considered FDD beamforming schemes, the “best” N beams (in the sense described in Sec. III-B) are selected from a codebook of size $M' = 512$, with N in the range from 1 to 128. We have chosen a large codebook size in order not to restrict the signal subspace. In Sec. V-A and Sec. V-D, we choose $\rho = 0$ dB. With this choice we obtain per-UE spectral efficiencies in the range 0.5–5.0 bits/s/Hz, which are representative of several wireless standards [21], [44]. Additionally for Sec. V-C, m -antenna subarrays, $K \leq m \leq M$, are considered. For each m , the m -antenna subarrays are selected so that they span the full length of the original M -antenna array. For example, when $m = 4$, 32 4-antenna subarrays are used; when $m = 16$, 8 16-antenna subarrays are used; and so on.

A. Relative Sum-rate as a Function of N

First, we examine scenarios 1, 2, and 3, for which $K = 4$ UEs. Fig. 3 (left half) shows the relative sum-rates $\bar{c}_A(\rho, N) = \bar{C}_A(\rho, N)/\bar{C}_{\text{TDD}}(\rho)$, where A is one of “D-GOB”, “H-GOB”, “D-SUB”, and “H-SUB”. The sum-capacities $\bar{C}_{\text{TDD}}(\rho)$ are given in Table III, in bits/s/Hz. For

⁵In fully-digital TDD systems spatial signatures need not be shaped as directional beams. Still we follow common usage and retain the term “beamforming” to refer to the spatial signal processing of fully-digital TDD.

TABLE I
SUMMARY OF MEASURED SCENARIOS.

Campaign A	<p>Scenario 1. $K = 4$ well-separated UEs in LOS, in which one UE from each of the sites MS 1 to MS 4 is selected. The minimum UE separation is 10 m.</p>	<p>Scenario 2. $K = 4$ co-located UEs in NLOS, in which four UEs are selected from site MS 5. The minimum UE separation is 0.5 m.</p>	<p>Scenario 3. $K = 4$ co-located UEs in LOS, in which four UEs are selected from site MS 2. The minimum UE separation is 0.5 m.</p>
Campaign B	<p>Scenario 4. $K = 4$ separated UEs in LOS and strong scattered components. We consider four sets of UEs (in different colors). The minimum UE separation is 3 m.</p>	<p>Scenario 5. $K = 8$ separated UEs in LOS and strong scattered components. We consider two sets of UEs (in different colors). The minimum UE separation is 3 m.</p>	<p>Scenario 6. $K = 16$ separated UEs in LOS and strong scattered components. The minimum UE separation is 3 m.</p>

TABLE II
MEDIAN LEVELS OF THE RICEAN K -FACTORS (IN dB).

Measurement Site	MS 1	MS 2	MS 3	MS 4	MS 5	Courtyard
Ricean K -factor	-0.5	5.2	3.7	-7.4	-11.6	-10.4

fixed N , we say that A outperforms B if $\bar{c}_A(\rho, N) > \bar{c}_B(\rho, N)$, where A and B may be applied to different scenarios.

TABLE III
ACHIEVABLE SUM-CAPACITY (IN bits/s/Hz) AT $\rho = 0$ dB.

Scenario	1	2	3	4	5	6
Number of UEs	4	4	4	4	8	16
Sum-capacity	19.9	20.0	16.7	20.0	32.2	49.0

With one exception⁶, the relative sum-rates $\bar{c}_A(\rho, N)$ increase with increasing values of N . At $N = 128$, D-SUB and H-SUB achieve the sum-capacity of the channel, and D-GOB and H-GOB achieve the sum-rate of ZF with perfect CSI. In general, D-GOB extracts a larger share of the sum-capacity than H-GOB, and D-SUB extracts a larger share than H-SUB. This must be so since, with D-GOB and D-SUB, beams are selected individually for each subcarrier, while with H-GOB and H-SUB, the same set of beams is used for the entire band. In other words, H-GOB can be seen as a subcase of D-GOB, and the same is true for H-SUB with respect to D-SUB. The horizontal gap between the curves of D-GOB and H-GOB in Fig. 3, and between those of D-SUB and H-SUB, represents the penalty due to the frequency selectivity of the channel, in

⁶In scenario 3, the relative sum-rate of H-GOB decreases slightly when N goes from 1 to 2. This can happen because ZF is used based on partial CSI.

terms of the number of additional beams needed, by H-GOB and H-SUB compared to D-GOB and D-SUB, respectively, to achieve a prescribed fraction of the sum-capacity. (Note that in our work, N itself is not frequency-dependent.) At 70% of the sum-capacity, this penalty is at most one beam for scenarios 1 and 3, and between 4 to 17 beams for scenario 2. In general, penalties are significantly larger for NLOS scenarios than for LOS ones, which can be explained by the larger frequency selectivity of NLOS channels [45].

Looking at scenarios 1 and 2, we note that D-GOB outperforms D-SUB. With $N = 4$, D-GOB can reach 82% of the sum-capacity, but D-SUB can only reach 72%; with $N = 10$, the relative sum-rates are 90% and 86%, respectively. In fact, this holds for all N , although the gap closes as N increases. This is somewhat surprising as one would expect that DPC should outperform ZF. The explanation is as follows. With D-GOB, beams are individually selected by each UE with the goal of maximizing the channel gain. With D-SUB, however, channel beamforming gains are traded off against lower multiuser interference. When channel propagation conditions are sufficiently favorable (e.g., distinct LOS directions as in scenario 1, or NLOS propagation as in scenario 2), maximizing the channel beamforming gain is the better strategy. In general, the relative performance of D-GOB and D-SUB depends on ρ : For all N , $\bar{c}_{D-SUB}(\rho, N)$ goes to 1 in the limit $\rho \rightarrow \infty$, with the difference between $\mathcal{C}_{TDD}(\rho)$ and $\mathcal{C}_{D-SUB}(\rho, N)$ constant [46], [47]. Meanwhile, $\bar{c}_{D-GOB}(\rho, N)$ goes to 1 as $\rho \rightarrow \infty$ only if multiuser interference can be perfectly removed, that is, only if all channel vectors can be perfectly reconstructed at the BS (e.g., by setting $N = 128$), and to 0 otherwise.

The situation is more involved with regards to H-GOB

and H-SUB: H-GOB beats H-SUB in scenario 1, and the opposite is true in scenario 2. This hints to a larger sensitivity to frequency selectivity of ZF compared to DPC. Turning to scenario 3, we observe that D-SUB and H-SUB vastly outperform D-GOB and H-GOB. As a matter of fact, D-SUB and H-SUB excel in this scenario. A part of the performance gap is due to the fact that DPC cannot be applied with D-GOB and H-GOB and we have to resort to, in our case, ZF for these (see footnote 3 in [37] for details). In addition, since the UEs are co-located and have LOS, addressing multiuser interference is crucial, and failure to do so (as with D-GOB and H-GOB) leads to large performance losses. We also make the obvious remark that if one desires to operate with $N < K$ beams, then D-GOB and H-GOB are the only available choices. An interesting conclusion thus far is that there is no single FDD beamforming technique, D-GOB or D-SUB, H-GOB or H-SUB, that is “best” in all cases, but which technique that is most appropriate depends largely on the propagation scenario. Broadly speaking, D-SUB and H-SUB are required whenever multiuser interference is significant, while D-GOB and H-GOB might be preferred as propagation becomes more favorable.

We now move on to scenarios 4, 5, and 6, with $K = 4, 8$, and 16 UEs, respectively, facing LOS propagation conditions with strong scattered components. Shown in Fig. 3 (right half) are the relative sum-rates $\bar{c}_A(\rho, N)$. The corresponding sum-capacities, $\bar{C}_{\text{TDD}}(\rho)$, are given in Table III, in bits/s/Hz. An important observation is that the presence of significant scatterers in the propagation environment has a notable impact on the performance of D-GOB, H-GOB, D-SUB, and H-SUB. To see this, compare in Fig 3 the reported values of N for scenario 1 with those, larger, of scenario 4. In addition to the LOS component, a substantial part of the received power in scenario 4 originates from scattered components. Therefore, more beams are needed to achieve a prescribed fraction of the sum-capacity of the channel.

We also note that the required number of beams, N , increases with the number of active UEs, K . That N should grow with K is consistent with the conventional Massive MIMO wisdom that the number of BS antennas (here, beams) should grow proportional to K [48]—this is also necessary for D-SUB and H-SUB, for which $N \geq K$ must hold, as we have already pointed out. In general, the scalability of FDD Massive MIMO as K grows is ultimately limited by the number of beams, N , that can be learnt and reported, regardless of how many antennas are added to the system. Thus, in practical systems, where N is typically small, the usefulness of FDD beamforming is thus limited to serving a small number of UEs.

From the above discussion, it should be clear that the performance of D-GOB, H-GOB, D-SUB and H-SUB is greatly influenced by the characteristics of the propagation scenario. In particular, LOS propagation conditions with large Ricean factors seem necessary to achieve reasonably good performance for small N . By contrast, TDD Massive MIMO offers high performance across a variety of propagation scenarios. In particular, LOS propagation is not required (cf. scenarios 1 and 4, and scenario 2 in Table III) for good performance. This distinguishing feature of TDD beamforming underlines

the value of fully-digital precoding and reciprocity-based CSI acquisition: With measured channels and no structural limitations on the precoded signals, NLOS channels are as good as LOS channels.

B. Required N for a Maximum Allowable SNR Loss

To obtain additional insights, we fix the sum-rate to a desired value, C^* , and investigate the impact of varying N , the number of reported beams. The required N will depend on C^* , and on the system SNR, ρ . Given $C^* > 0$ and $1 \leq N \leq 128$, it is immediate that one must use $\rho \geq \rho^*$, with ρ^* being the required SNR of TDD beamforming at C^* . We define the SNR loss δ_ρ by the expression

$$\delta_\rho := \rho^* / \rho. \quad (20)$$

Shown in Fig. 4 (left half) is the required number of beams, N , as a function of the maximum allowable SNR loss, δ_ρ , for $C^* = 12$ bits/s/Hz, and for scenarios 1, 2 and 3. In general, N increases sharply with decreasing SNR loss. In scenario 1, D-GOB is more efficient than D-SUB, and H-GOB is more so than H-SUB. At 3 dB SNR loss, D-GOB, H-GOB, D-SUB and H-SUB require 3, 4, 6, and 7 beams, respectively. If 6 dB SNR loss is allowed, D-GOB can operate with $N = 1$ beam, and H-GOB with $N = 2$ beams. On the other hand, in scenario 3, D-SUB and H-SUB greatly outperform D-GOB and H-GOB. In fact, neither D-GOB nor H-GOB can operate at less than 3 dB SNR loss, regardless of N . In scenario 2, none of the four investigated techniques can operate at low SNR loss with small N : At 3 dB SNR loss, all of them require $N > 20$.

Shown in Fig. 4 (right half) is N versus the maximum allowable SNR loss, δ_ρ , for $C^* = 12, 24$ and 48 bits/s/Hz and $K = 4, 8$, and 16 UEs as obtained from scenarios 4, 5 and 6, respectively. The required N increases rapidly with K . For a large range of the SNR loss, D-GOB outperforms D-SUB, and H-GOB outperforms H-SUB.

C. Tradeoff of Antennas versus RF Chains

We next address the following question: “Given a system (N', m) with N' RF chains and m antennas, $m \geq N'$, to which extent can one compensate for a reduction of N' by increasing m ?” For that, we consider the level curves Γ_β of the SNR loss for some fixed sum-rate C^* , where the parameter β represents the maximum allowable SNR loss. That is,

$$\Gamma_\beta = \{(r(m), m) : K \leq m \leq M\}, \quad (21)$$

where

$$r(m) = \min\{n : K \leq n \leq m, \delta_\rho(n, m) \geq \beta\}, \quad (22)$$

and where $\delta_\rho(n, m)$ is defined similarly to (20) as the SNR loss of a system with n RF chains and m antennas with respect to TDD beamforming with 128 antennas.

Let $\beta \in \{1, 3, 6, 9, 12\}$ dB. Fig. 5 shows the corresponding level curves for H-SUB, $C^* = 12$ bits/s/Hz and scenarios 1, 2 and 3. For each β , there exists a minimal-size fully-digital system (m^*, m^*) . The system is minimal in the sense that the number of RF chains, $N' = m^*$, cannot be further reduced

Fig. 3. Sum-rates relative to the achievable sum-capacity of the channel ($\rho = 0$ dB). As a baseline, the performance of small aperture $K \times K$ MU-MIMO is also shown.

Fig. 4. Reported beams, N , as a function of the allowable SNR loss. The sum-rate has been fixed to 12, 24 or 48 bits/s/Hz, depending on the number of UEs $K = 4, 8$, or 16, respectively.

without violating the SNR loss requirement, β . For example, in scenario 1, if $\beta = 1$ dB, then the minimum size is $m^* = 100$; but if $\beta = 3$ dB, then $m^* = 69$. Fig. 5 also shows that there exists a multiplicity of hybrid systems for which β is upheld, namely, those corresponding to points (N', m) , with $N' < m^*$ and $m \geq m^*$, along the relevant level curve Γ_β . These hybrid systems can be reached by starting from (m^*, m^*) and moving to the left along Γ_β . For example, in scenario 1 and under $\beta = 1$ dB, it is possible to travel from the point $(100, 100)$ to the point $(76, 100)$, essentially reducing the number of RF chains by 24 at no additional cost. To further reduce N' , one

must traverse the segment $(76, 100) - (76, 103) - (56, 103)$, which implies that 20 RF chains can additionally be saved by spending another 3 antennas. One can proceed in this way until the point $(21, 128)$ is reached. One important observation is that saving RF chains becomes more and more expensive along the way, i.e., as N' turns smaller.

The situation looks quite different for propagation scenario 2. In particular, the level curves are notably steeper. The level curve under $\beta = 1$ dB is given by the segment $(100, 100) - (91, 100) - (91, 103) - (88, 103) - \dots - (71, 128)$: A maximal saving of 29 RF chains can be obtained by spending 28

Fig. 5. Required number of BS antennas, m , versus RF chains, N' , for H-SUB transmission with $K = 4$ UEs, and 12 bits/s/Hz. The curve $N' = m$ for fully-digital has been highlighted as reference.

antennas. It is not obvious that the resulting hybrid system (71,128) is cheaper to realize than the original (100,100) system. In stark contrast, the level curves of scenario 3 are close to horizontal, suggesting that drastic reductions in the number of RF chains are possible. For example, the level curve under $\beta = 1$ dB starts at (116,116) and ends at (6,128). In other words, 110 RF chains can be saved by merely adding 12 antennas.

D. The Impact of DL Training Overhead

We next illustrate the performances of the different transmission schemes when the training overhead is taken into account. We assume a simple block-fading model, where the channel is constant for T_c samples. Typically, T_c is the length (time-bandwidth product) of the coherence interval of the channel, and ranges from just above one to a few hundred, depending on the carrier frequency, the richness of the channel (multipath), and the relative motion of the BS, UEs, and scatterers (Doppler). As an illustrative value, $T_c = 200$ corresponds to, e.g., a coherence time of 1 ms and a coherence bandwidth of 200 kHz. We assume that N_p DL pilot symbols are inserted within each coherence interval, leaving $T_c - N_p$ symbols available for data. For D-SUB and H-SUB, $N_p \geq N$ pilot symbols are needed to learn the channel.⁷ For D-GOB and H-GOB, we have that $N_p \geq \alpha N$, where α ranges from $\alpha = 1$, if all the UEs report the same beams, to $\alpha = K$, if the UEs report distinct beams. Here, we consider the worst case $\alpha = K$. Thus, we let

$$N_p(N) = \begin{cases} KN & \text{for D-GOB, H-GOB} \\ N & \text{for D-SUB, H-SUB,} \end{cases} \quad (23)$$

⁷This is true *after* the N beams have been selected. Optimal beam selection requires that the entire “beam space” is observed, implying $N_p = 128$. Nonetheless, $N_p = N$ holds approximately if one assumes that the *structure* of the beam space changes much more slowly than the particular coefficients of the beams. That is, if one assumes that the length of the stationarity regions of the channel is much larger than the length of the coherence interval [49].

Fig. 6. Sum-rates relative to the achievable sum-capacity of the channel as a function of T_c with $K = 4$ UEs, and $\rho = 0$ dB.

and compute the sum-rate $\tilde{C}_A(\rho, T_c)$ achievable over a large number of fading blocks (see [50]) by the formula

$$\tilde{C}_A(\rho, T_c) = \left(1 - \frac{N_p(N^*)}{T_c}\right) \bar{C}_A(\rho, N^*), \quad (24)$$

where the average sum-rates $\bar{C}_A(\rho, N)$ can be inferred from Fig. 3 and Table III, and the quantity N^* is defined by

$$N^* = \arg \max_{1 \leq N \leq 128} \left(1 - \frac{N_p(N)}{T_c}\right) \bar{C}_A(\rho, N). \quad (25)$$

From (25), N^* is the optimal number of beams to be activated: If $N < N^*$, the degrees of freedom offered by the channel are underused, whereas if $N > N^*$, too few symbols are left available for data.

The following example demonstrates that when the overhead of DL training is properly accounted for, D-GOB and D-SUB can nevertheless extract a sizable share of the sum-capacity of LOS channels. In NLOS conditions (scenario 2), however, these techniques do not work well.

Example 1: Let $\rho = 0$ dB, and let $T_c = 1, 2, \dots, 200$. Fig. 6 shows $\tilde{C}_A(\rho, T_c)$ relative to the sum-capacity of the channel for D-GOB and D-SUB, and the sum-capacity of 4×4 MU-MIMO. For comparison, the corresponding curve for TDD beamforming assuming that $N_p = K$ pilot symbols per coherence interval are allocated is also shown.⁸ D-GOB and D-SUB perform several times better than conventional MU-MIMO, with D-SUB consistently outperforming D-GOB. D-GOB performs poorly if UEs are co-located with LOS, and neither works well in NLOS. In all scenarios, TDD with DL training overhead performs better than D-SUB and D-GOB.

⁸ $N_p = K$ comes from assuming that one DL pilot per coherence interval, T_c , is allocated to each UE, so that the UEs can learn their instantaneous *effective* channel gains. This is a bit pessimistic as substantial evidence exists that no DL pilots may be needed at all, not even beamforming pilots [51], in which case we would have $N_p = 0$.

Fig. 7. Optimal number of active beams, N^* , as a function of T_c with $K = 4$ UEs, and $\rho = 0$ dB.

Fig. 8. Sum-rates relative to the achievable sum-capacity of channel as a function of N with $K = 4$ UEs, $\rho = 0$ dB, and $T_c = 200$.

The associated values of N^* are shown in Fig. 7 for D-GOB and D-SUB. Observe that as T_c increases, more beams should be activated. As the UEs may report distinct beams, DL training with D-GOB is more expensive, and N^* is thus pushed towards zero.

In the next example, we examine the optimal number of active beams, N^* , with H-GOB and H-SUB. It is shown that, in LOS conditions, H-GOB and especially H-SUB perform reasonably well when operated with a small excess of RF chains, i.e., $N = K + 2$, or so.

Example 2: Let $\rho = 0$ dB, and let $T_c = 200$. Fig. 8 shows $\left(1 - \frac{N_p(N)}{T_c}\right) \bar{C}_A(\rho, N)$ relative to optimal TDD as a function of N . For H-GOB, it is optimal to activate 4, 15, and 9 beams in scenarios 1, 2 and 3, respectively. For H-SUB, the numbers

Fig. 9. Sum-rate of ZF as a function of the filtering factor, ζ , with $K = 4$ UEs, and $\rho = 0$ dB. CSI is corrupted by noise with variance ζ/ρ . For D-GOB and D-SUB, the number of beams has been fixed to $N = 5$. The sum-rate of TDD is given by the thin solid lines.

are 18, 38, and 10. In fact, in LOS scenarios activating $K+2 = 6$ beams results in losses smaller than 10% of the relative sum-rate at N^* . In NLOS scenarios, however, losses at $K+2$ beams surge to 20–40% of an already much diminished peak relative sum-rate.

In the examples thus far, we have investigated the impact of DL training overhead assuming that perfect CSI is available. Next, we revisit Example 1, but this time we consider the case of imperfect CSI.

Example 3: Let $\rho = 0$ dB, and let $N = 5$ beams. (We also tried other values of N ; similar results were obtained.) Fig. 9 shows the achievable sum-rates, in bits/s/Hz, of D-GOB, D-SUB, and TDD beamforming when additive Gaussian noise with variance ζ/ρ is present during the channel training phase. For estimating the complex gains of the beams, simple scalar minimum mean-square error estimation is used. The parameter $0 \leq \zeta \leq 1$ corresponds to the amount of time-frequency filtering that is available, with $\zeta = 1$ meaning no filtering and $\zeta = 0$ perfect CSI. The filtering factor may include processing gains arising from letting the DL training length, τ , being larger than the number of pilots needed to learn the channel, N_p . For example, letting $\tau = 2 \times N_p$, say, corresponds to $\zeta = 1/2$. Frequency correlations may also be exploited to lower ζ . To compare among as practical setups as possible, ZF precoding has been applied to all of TDD, D-GOB, and D-SUB. The main observation is that if the system is designed such that $\zeta \ll 1$, then the relative advantage of TDD beamforming over D-GOB and D-SUB remains largely unchanged. In scenario 3, D-GOB and D-SUB perform better than TDD whenever $\zeta \geq 0.25$. This is possible because, in the preparation of Fig. 9, beams have been selected assuming perfect CSI (but as explained above, the associated complex gains are estimated from noisy channel observations), which effectively reduces the training noise variance of D-

GOB and D-SUB by a factor $\frac{128}{N}$ compared to TDD. In practice, however, beam selection from perfect CSI is not feasible, and noisy channel observations must be used.⁹ On top of this, we note that there are several impairments that affect FDD operation but not TDD, such as the quantization and aging of the fed-back complex coefficients, and which are not accounted for in Fig. 9. Thus, once considering these impairments, the performance of D-GOB and D-SUB is expected to be inferior to what is shown in Fig. 9.

E. On the Performance of Analog-Only Beamforming

The main remark we shall make here is that analog-only beamforming does not offer a sum-rate advantage over conventional, small aperture MU-MIMO systems, except for the very special case of well-separated UEs with LOS. For that, recall that analog-only beamforming is the same as H-GOB with $N = 1$, but wherein baseband processing has been suppressed. In fact, analysis of the measured channels shows that the sum-rates of analog-only beamforming, and those of regular H-GOB (thus *with* baseband processing) differ by less than 1%, in all scenarios. The claim follows by direct inspection of Fig. 3, in Sec. V-B.

F. On the Applicability of the Results to mm-Wave Bands

Although it is tempting to extrapolate our results to mm-wave frequencies, there are difficulties in doing so. Increasing the carrier frequency above 6 GHz and up to mm-wave bands leads to higher free-space pathloss (20 dB at factor 10 frequency increase) with isotropic antennas [35], higher penetration and diffraction losses [52], and more severe shadow fading [53]. The net result seems to be a reduction in the number of active scatterers [54], i.e., a sparsification of the channel. The additional pathloss can be overcome, e.g., by increasing the number of UE antennas [39]. The impact of a reduction in the number of scatterers depends, however, on the positions of the UEs and of the remaining scatterers. Up to date, most work on mm-wave MIMO systems has been focused on the single user setting, and there have been few results on the multiuser case. In particular, no measurement-based reports of multiuser settings in mm-wave bands are known to the authors. Thus, to evaluate the performance of the considered TDD and FDD DL beamforming techniques in practical mm-wave channels more measurements and analysis are needed.

VI. CONCLUSIONS

Using measured channels at 2.6 GHz, we have compared the performance of five techniques for DL beamforming in Massive MIMO, namely, fully-digital reciprocity-based (TDD) beamforming, and four flavors of FDD beamforming based on feedback of CSI (D-GOB, H-GOB, D-SUB, and H-SUB).

⁹One might argue that in FDD operation, long-term spatial correlations of the channel might be available to the UEs, enabling close-to-ideal selection of the beams. While this might be true, those statistics would also be available at the BS in TDD operation, and then subspace methods for channel estimation could also be applied [49].

The central result is that, while FDD beamforming with pre-determined beams may achieve a hefty share of the DL sum-rate of TDD beamforming, performance depends critically on the existence of advantageous propagation conditions, namely, LOS and high Ricean factors are both necessary. In other considered scenarios, the performance loss is significant for the non reciprocity-based beamforming solutions. Therefore, at the frequency bands under consideration, if robust operation across a wide variety of propagation conditions is required, reciprocity-based TDD beamforming is the only feasible alternative.

When the overhead of DL training in FDD operation is taken into account, D-SUB and H-SUB (for which beams are selected by the BS) achieve larger sum-rates than D-GOB and H-GOB (for which each UE individually selects its preferred beams). In LOS scenarios, D-GOB, H-GOB, D-SUB and H-SUB can perform several times better than conventional, small-aperture MU-MIMO. D-GOB and H-GOB necessitate well-separated UEs, while with co-located UEs D-SUB and H-SUB have the best relative performance. If hardware savings are desired, H-SUB or H-GOB can be used with the number of RF chains slightly exceeding the number of UEs. In NLOS scenarios, however, or if strong reflections are present, FDD beamforming was found to perform significantly worse than TDD.

APPENDIX

A. Efficient Algorithm for Approximate Solution of (14)

As noted in Sec. III-B, solving problem (14) exactly becomes computationally intractable for moderately large values of M' . Instead, we present an algorithmic solution based on the concept of greedy pursuit. The algorithm is summarized in Alg. 1. (Note that, for simplicity of notation, the indices ℓ and k have been omitted.) In short, the procedure starts by obtaining (steps 3 and 4) the index j^* such that $\mathbf{h}$ has the largest projection along $\mathbf{c}_{j^*}$. It then stores $\mathbf{c}_{j^*}$ and j^* in steps 5 and 6 to form $\mathbf{B}^{(1)}$ and $\mathcal{Q}^{(1)}$, respectively. In the next iteration, a new beam $\mathbf{c}_{j^*}$ is selected such as to maximize the projection on the subspace spanned by the columns of $[\mathbf{B}^{(1)} \mid \mathbf{c}_{j^*}]$ of $\mathbf{h}$. (Note that the desired projection is given as the result of the multiplication $[\mathbf{B}^{(i-1)} \mid \mathbf{c}_j]^\dagger [\mathbf{B}^{(i-1)} \mid \mathbf{c}_j]^\text{H} \mathbf{h}$ in step 4.) It then repeats steps 3 and 4. The algorithm continues until steps 3 to 6 have been executed exactly N times, at which point $\mathbf{B}^{(N)}$ would contain the N selected beams, and $\mathcal{Q}^{(N)}$ their indices. Computationally, Alg. 1 can be efficiently implemented by sequential Gram-Schmidt orthogonalization of the beamforming matrices $\mathbf{B}^{(1)}, \dots, \mathbf{B}^{(N)}$.

B. The Sum-Capacity of the MIMO-BC with Beamforming

For ease of notation, we drop the index ℓ . For given $\mathbf{B}$ and ρ , $C_{\text{BC}}(\mathbf{H}\mathbf{B}, \rho)$ is the sum-rate of the MIMO BC $\mathbf{H}\mathbf{B}$, and

Algorithm 1 UE-side Greedy Beam Selection

Require: $\mathbf{h}, \mathbf{C}, N$

- 1: $\mathcal{Q}^{(0)} = \emptyset, \mathbf{B}^{(0)} = [\]$
- 2: **for** $i = 1$ **to** N **do**
- 3: $\mathcal{S}^{(i)} = \{1, \dots, M'\} \setminus \mathcal{Q}^{(i-1)}$
- 4: $j^* = \arg \max_{j \in \mathcal{S}^{(i)}} \|[\mathbf{B}^{(i-1)} \mid \mathbf{c}_j]^\dagger [\mathbf{B}^{(i-1)} \mid \mathbf{c}_j]^H \mathbf{h}\|^2$
- 5: $\mathbf{B}^{(i)} = [\mathbf{B}^{(i-1)} \mid \mathbf{c}_{j^*}]$
- 6: $\mathcal{Q}^{(i)} = \mathcal{Q}^{(i-1)} \cup \{j^*\}$
- 7: **end for**
- 8: **return** $\mathcal{Q} = \mathcal{Q}^{(N)}, \mathbf{B} = \mathbf{B}^{(N)}$.

is given by the solution to [29], [30]:

$$\begin{aligned} & \underset{\{\mathbf{Q}_i\}_{i=1}^K}{\text{maximize}} \quad \sum_{i=1}^K \log_2 \left(\frac{1 + \mathbf{h}_i^T \mathbf{B} \left(\sum_{j=1}^i \mathbf{Q}_j \right) \mathbf{B}^H \mathbf{h}_i^*}{1 + \mathbf{h}_i^T \mathbf{B} \left(\sum_{j=1}^{i-1} \mathbf{Q}_j \right) \mathbf{B}^H \mathbf{h}_i^*} \right) \\ & \text{subject to} \quad \mathbf{Q}_i \succeq 0, \quad \sum_{i=1}^K \text{tr}(\mathbf{B} \mathbf{Q}_i \mathbf{B}^H) \leq \rho, \end{aligned} \quad (26)$$

where $\mathbf{Q}_1, \dots, \mathbf{Q}_K$ are covariance matrices. The objective function of (26) is nonconcave in $\mathbf{Q}_1, \dots, \mathbf{Q}_K$, and hence finding the maximum is a nontrivial problem. One would like to apply the BC-multiple access channel (MAC) duality theorem [29] so as to transform the nonconcave problem (26) into an equivalent, concave one, for which efficient solvers are known to exist [55]. However, the presence of $\mathbf{B}$ in the constraint $\sum_{i=1}^K \text{tr}(\mathbf{B} \mathbf{Q}_i \mathbf{B}^H) \leq \rho$ prevents us from invoking the BC-MAC duality theorem. Fortunately, we have the following useful result.

Lemma 1: For given $\mathbf{B} = \mathbf{U}\mathbf{L}$ with $\mathbf{U}^H \mathbf{U} = \mathbf{I}$, and $\mathbf{L}$ an invertible matrix, and for given ρ , we have that

$$\mathcal{C}_{\text{BC}}(\mathbf{H}\mathbf{B}, \rho) = \mathcal{C}_{\text{MAC}}(\mathbf{H}\mathbf{U}^H, \rho), \quad (27)$$

where $\mathcal{C}_{\text{MAC}}(\mathbf{H}\mathbf{U}^H, \rho)$ is the sum-capacity of the MIMO-MAC $\mathbf{H}\mathbf{U}^H$ [56].

Proof: By inserting $\mathbf{B} = \mathbf{U}\mathbf{L}$ into equation (26), we obtain the optimization problem

$$\begin{aligned} & \underset{\{\mathbf{Q}_i\}_{i=1}^K}{\text{maximize}} \quad \sum_{i=1}^K \log_2 \left(\frac{1 + \mathbf{h}_i^T \mathbf{U}\mathbf{L} \left(\sum_{j=1}^i \mathbf{Q}_j \right) \mathbf{L}^H \mathbf{U}^H \mathbf{h}_i^*}{1 + \mathbf{h}_i^T \mathbf{U}\mathbf{L} \left(\sum_{j=1}^{i-1} \mathbf{Q}_j \right) \mathbf{L}^H \mathbf{U}^H \mathbf{h}_i^*} \right) \\ & \text{subject to} \quad \mathbf{Q}_i \succeq 0, \quad \sum_{i=1}^K \text{tr}(\mathbf{L} \mathbf{Q}_i \mathbf{L}^H) \leq \rho, \end{aligned} \quad (28)$$

where we have used that $\text{tr}(\mathbf{U}\mathbf{L} \mathbf{Q}_i \mathbf{L}^H \mathbf{U}^H) = \text{tr}(\mathbf{L} \mathbf{Q}_i \mathbf{L}^H)$ by the cyclic property of the trace operator and the fact that $\mathbf{U}^H \mathbf{U} = \mathbf{I}$, by assumption.

Define the effective covariance matrices $\tilde{\mathbf{Q}}_i = \mathbf{L} \mathbf{Q}_i \mathbf{L}^H$, $i = 1, \dots, K$, and the effective channel $\tilde{\mathbf{H}} = \mathbf{H}\mathbf{U}$. Using

Algorithm 2 BS-side Multiuser Greedy Beam Selection

Require: $\mathbf{H}, \mathbf{C}, N, \rho$

- 1: $\mathcal{Q}^{(0)} = \emptyset, \mathbf{B}^{(0)} = [\], \mathbf{\Lambda} = \frac{\rho}{K} \mathbf{I}$
- 2: **for** $i = 1$ **to** N **do**
- 3: $\mathcal{S}^{(i)} = \{1, \dots, M'\} \setminus \mathcal{Q}^{(i-1)}$
- 4: $j^* = \arg \max_{j \in \mathcal{S}^{(i)}} \log_2 \left| \mathbf{I} + \mathbf{U}_j^T \mathbf{H}^H \mathbf{\Lambda} \mathbf{H} \mathbf{U}_j^* \right|$,
where $\mathbf{U}_j = [\mathbf{B}^{(i-1)} \mid \mathbf{c}_j]^\dagger [\mathbf{B}^{(i-1)} \mid \mathbf{c}_j]^H$.
- 5: $\mathbf{B}^{(i)} = [\mathbf{B}^{(i-1)} \mid \mathbf{c}_{j^*}]$
- 6: $\mathcal{Q}^{(i)} = \mathcal{Q}^{(i-1)} \cup \{j^*\}$
- 7: **end for**
- 8: **return** $\mathcal{Q} = \mathcal{Q}^{(N)}, \mathbf{B} = \mathbf{B}^{(N)}$.

these definitions, and the fact that $\mathbf{L}$ is invertible, (28) can be rewritten as

$$\begin{aligned} & \underset{\{\tilde{\mathbf{Q}}_i\}_{i=1}^K}{\text{maximize}} \quad \sum_{i=1}^K \log_2 \left(\frac{1 + \tilde{\mathbf{h}}_i^T \left(\sum_{j=1}^i \tilde{\mathbf{Q}}_j \right) \tilde{\mathbf{h}}_i^*}{1 + \tilde{\mathbf{h}}_i^T \left(\sum_{j=1}^{i-1} \tilde{\mathbf{Q}}_j \right) \tilde{\mathbf{h}}_i^*} \right) \\ & \text{subject to} \quad \tilde{\mathbf{Q}}_i \succeq 0, \quad \sum_{i=1}^K \text{tr}(\tilde{\mathbf{Q}}_i) \leq \rho. \end{aligned} \quad (29)$$

Crucially, because $\mathbf{L}$ is invertible, $\tilde{\mathbf{Q}}_i = \mathbf{L} \mathbf{Q}_i \mathbf{L}^H$ is an isomorphism. Thus, for every $\{\tilde{\mathbf{Q}}_i\}_{i=1}^K$ satisfying the constraints in (29) we can find $\{\mathbf{Q}_i\}_{i=1}^K$ fulfilling the constraints in (28), and the converse is also true. We may now apply the BC-MAC duality theorem [29] to (29), from which the desired result follows. ■

C. Efficient Algorithm for Approximate Solution of (16)

An algorithmic solution for beam selection in multiuser MIMO systems is presented in Alg. 2. For ease of notation, the index ℓ has been omitted. Alg. 2 is again based on the concept of greedy pursuit, and proceeds analogously to Alg. 1, although with a different objective function. In particular, the objective function in Alg. 2 needs to depend on the channel matrix $\mathbf{H}$, rather than on a single channel vector $\mathbf{h}_k$. Also, the selection of the beams depends now on the system SNR, ρ . Once the N beams (that is, the columns of the beam-former $\mathbf{B}$) have been selected, the optimal covariance matrices $\mathbf{Q}_1, \dots, \mathbf{Q}_K$ may be computed by first solving (15), and then applying the MAC-to-BC transformation—see, e.g., [29], [30], [32]. The selection of the beams along with the computation of the MIMO-BC covariance matrices is done independently for each subcarrier.

ACKNOWLEDGMENT

The presented investigations are based on data obtained in measurement campaigns performed by Xiang Gao, Fredrik Tufvesson, Ove Edfors, Tommy Hult, and Meifang Zhu, as well as Sohail Payami, and Fredrik Tufvesson.

REFERENCES

- [1] T. L. Marzetta, "Noncooperative cellular wireless with unlimited number of base station antennas," *IEEE Trans. Wireless Commun.*, vol. 9, no. 11, pp. 3590–3600, Nov. 2010.

- [2] T. L. Marzetta, E. G. Larsson, H. Yang, and H. Q. Ngo, *Fundamentals of Massive MIMO*. Cambridge: Cambridge University Press, 2016.
- [3] T. L. Marzetta, "How much training is required for multiuser MIMO?" in *Proc. ASILOMAR 2006 - 40th Conf. on Sig., Syst. and Comput. (ACSSC)*, Pacific Grove, CA, USA, Nov.–Dec. 2006, pp. 359–363.
- [4] F. Rusek, D. Persson, B. K. Lau, E. G. Larsson, T. L. 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, Jan. 2013.
- [5] X. Gao, O. Edfors, F. Rusek, and F. Tufvesson, "Massive MIMO performance evaluation based on measured propagation data," *IEEE Trans. Wireless Commun.*, vol. 14, no. 7, pp. 3899–3911, 2015.
- [6] J. Flordelis, X. Gao, G. Dahman, F. Rusek, O. Edfors, and F. Tufvesson, "Spatial separation of closely-spaced users in measured massive multi-user MIMO channels," in *Proc. ICC 2015 - IEEE Int. Conf. Commun.*, London, UK, Jun. 2015, pp. 1441–1446.
- [7] P. Harris, S. Malkowsky, J. Vieira, F. Tufvesson, W. B. Hasan, L. Liu, M. Beach, S. Armour, and O. Edfors, "Performance characterization of a real-time massive MIMO system with LOS mobile channels," *IEEE J. Sel. Areas Commun.*, vol. 35, no. 6, pp. 1244–1253, Jun. 2017.
- [8] H. Prabhu, J. Rodrigues, L. Liu, and O. Edfors, "A 60 pJ/b 300 Mb/s 128×8 massive MIMO precoder-detector in 28 nm FD-SOI," in *Proc. ISSCC 2017 - Int. Solid-State Circuits Conf.*, San Francisco, CA, Feb. 2017, pp. 171–176.
- [9] Y.-H. Nam, B. L. Ng, K. Sayana, Y. Li, J. C. Zhang, Y. Kim, and J. Lee, "Full-dimension MIMO (FD-MIMO) for next generation cellular technology," *IEEE Commun. Mag.*, vol. 51, no. 6, pp. 172–179, Jun. 2013.
- [10] J. Choi, D. J. Love, and P. Bidigare, "Downlink training techniques for FDD massive MIMO systems: Open-loop and closed-loop training with memory," *IEEE J. Sel. Topics Signal Process.*, vol. 8, no. 5, pp. 802–814, Oct. 2014.
- [11] J. Choi, D. J. Love, and T. Kim, "Trellis-extended codes and successive phase adjustment: A path from LTE-advanced to FDD massive MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 4, pp. 2007–2016, 2015.
- [12] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niu, "Achievable rates of FDD massive MIMO systems with spatial channel correlation," *IEEE Trans. Wireless Commun.*, vol. 14, no. 5, pp. 2868–2882, Jan. 2015.
- [13] H. Ji, Y. Kim, J. Lee, E. Onggosanusi, Y. Nam, J. Zhang, B. Lee, and B. Shim, "Overview of full-dimension MIMO in LTE-advanced pro," *IEEE Commun. Mag.*, vol. 55, no. 2, pp. 176–184, Feb. 2017.
- [14] A. Adhikary, N. Junyoung, A. Jae-Young, and G. Caire, "Joint spatial division and multiplexing—the large-scale array regime," *IEEE Trans. Inf. Theory*, vol. 59, no. 10, pp. 6441–6463, Jun. 2013.
- [15] S. Bazzi and W. Xu, "Downlink training sequence design for FDD multiuser massive MIMO systems," *IEEE Trans. Signal Process.*, vol. 65, no. 18, pp. 4732–4744, Sep. 2017.
- [16] F. Sotrovi and W. Yu, "Hybrid digital and analog beamforming design for large-scale antenna arrays," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 501–513, Apr. 2016.
- [17] T. E. Bogale, L. B. Le, A. Haghighat, and L. Vandendorpe, "On the number of RF chains and phase shifters, and scheduling design with hybrid analog-digital beamforming," *IEEE Trans. Wireless Commun.*, vol. 15, no. 5, pp. 3311–3326, May 2016.
- [18] O. E. Ayach, S. Rajagopal, S. Abu-Surra, Z. Pi, and R. W. Heath, "Spatially sparse precoding in millimeter wave MIMO systems," *IEEE Trans. Wireless Commun.*, vol. 13, no. 3, pp. 1499–1513, Mar. 2014.
- [19] S. Han, C.-L. I, Z. Xu, and C. Rowell, "Large-scale antenna systems with hybrid analog and digital beamforming for millimeter wave 5G," *IEEE Commun. Mag.*, vol. 53, no. 1, pp. 186–194, Jan. 2015.
- [20] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, vol. 14, no. 11, pp. 6481–6494, Nov. 2015.
- [21] E. Dahlman, S. Parkvall, J. Sköld, and P. Beming, *3G Evolution. HSPA and LTE for Mobile Broadband*, 2nd ed. London: Academic Press, 2008.
- [22] A. F. Molisch, V. V. Ratman, S. Han, Z. Li, S. L. H. Nguyen, L. Li, and K. Haneda, "Hybrid beamforming for massive mimo: A survey," *IEEE Communications Magazine*, vol. 55, no. 9, pp. 134–141, 2017.
- [23] T. L. Marzetta, "Massive MIMO: An introduction," *Bell Labs Tech. J.*, vol. 20, pp. 11–22, 2015.
- [24] E. Björnson, E. G. Larsson, and T. L. Marzetta, "Massive MIMO: Ten myths and one critical question," *IEEE Commun. Mag.*, vol. 54, no. 2, pp. 114–123, Feb. 2016.
- [25] J. G. Proakis and M. Salehi, *Digital Communications*. Great Britain: McGraw-Hill, 5th Edition, International Edition, 2014.
- [26] J. Vieira, F. Rusek, O. Edfors, S. Malkowsky, L. Liu, and F. Tufvesson, "Reciprocity calibration for massive MIMO: Proposal, modeling and validation," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 3042–3056, May 2017.
- [27] M. Costa, "Writing on dirty paper," *IEEE Trans. Inf. Theory*, vol. 29, no. 3, pp. 439–441, May 1983.
- [28] G. Caire and S. Shamai, "On the achievable throughput of a multiantenna Gaussian broadcast channel," *IEEE Trans. Inf. Theory*, vol. 49, no. 7, pp. 1691–1706, Jul. 2003.
- [29] S. Vishwanath, N. Jindal, and A. Goldsmith, "Duality, achievable rates, and sum-rate capacity of MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 49, no. 10, pp. 2658–2668, Oct. 2003.
- [30] S. Vishwanath and D. N. C. Tse, "Sum capacity of the vector Gaussian broadcast channel and uplink-downlink duality," *IEEE Trans. Inf. Theory*, vol. 49, no. 8, pp. 1912–1921, Aug. 2003.
- [31] W. Yu and J. M. Cioffi, "Sum capacity of a Gaussian vector broadcast channel," *IEEE Trans. Inf. Theory*, vol. 50, no. 9, pp. 1875–1892, Sep. 2002.
- [32] N. Jindal, W. Rhee, S. Vishwanath, A. Jafar, and A. Goldsmith, "Sum power iterative water-filling for multi-antenna Gaussian broadcast channels," *IEEE Trans. Inf. Theory*, vol. 51, no. 4, pp. 1570–1580, Apr. 2005.
- [33] P. He and L. Zhao, "Correction of convergence proof for iterative water-filling in gaussian MIMO broadcast channels," *IEEE Trans. Inf. Theory*, vol. 57, no. 4, pp. 2539–2543, 2011.
- [34] A. Paulraj, R. Nabar, and D. Gore, *Introduction to Space-Time Wireless Communications*, 1st ed. Cambridge University Press, 2008.
- [35] A. F. Molisch, *Wireless Communications*. New York: John Wiley & Sons, 2011.
- [36] N. Jindal, "MIMO broadcast channels with finite-rate feedback," *IEEE Trans. Inf. Theory*, vol. 52, no. 11, pp. 5045–5060, Nov. 2005.
- [37] —, "MIMO broadcast channels with finite rate feedback," in *Global Telecommunications Conference, GLOBECOM, 2005*, 29 Nov.–2 Dec. 2005, pp. 1520–1524.
- [38] J. Choi, K. Lee, D. J. Love, T. Kim, and R. W. Heath, "Advanced limited feedback designs for FD-MIMO using uniform planar arrays," in *Global Telecommunications Conference, GLOBECOM, 2015*, San Diego, California, USA, Dec. 2015, pp. 1–6.
- [39] R. W. Heath Jr., N. González-Prelcic, S. Rangan, W. Roh, and A. M. Sayeed, "An overview of signal processing techniques for millimeter wave MIMO systems," *IEEE J. Sel. Topics Signal Process.*, vol. 10, no. 3, pp. 436–453, Apr. 2016.
- [40] *IEEE 802.11ad, Amendment 3: Enhancements for Very High Throughput in the 60 GHz Band*, IEEE 802.11 Working Group Std., 2012.
- [41] L. J. Greenstein, D. G. Michelson, and V. Erceg, "Method-moment estimation of the Ricean K -factor," *IEEE Commun. Lett.*, vol. 3, no. 6, pp. 175–176, Jun. 1999.
- [42] C. Tepedelenglioglu, A. Abdi, and G. B. Giannakis, "The Ricean K factor: Estimation and performance analysis," *IEEE Trans. Wireless Commun.*, vol. 2, no. 4, pp. 799–810, Jul. 2003.
- [43] S. Payami and F. Tufvesson, "Channel measurements and analysis for very-large array systems at 2.6 GHz," in *Proc. EuCAP 2012 - 6th European Conf. in Ant. and Prop.*, Prague, Czech Republic, Mar. 2012, pp. 433–437.
- [44] *IEEE 802.11, Wireless Local Area Networks*, IEEE 802.11 Working Group Std.
- [45] N. Cardona, *Cooperative Radio Communications for Green Smart Environments*. Gistrup: River Publishers, 2016.
- [46] A. Lozano, A. M. Tulino, and S. Verdú, "High-SNR power offset in multiantenna communication," *IEEE Trans. Inf. Theory*, vol. 51, no. 12, pp. 4134–4151, Dec. 2005.
- [47] J. Lee and N. Jindal, "High SNR analysis for MIMO broadcast channels: Dirty paper coding versus linear precoding," *IEEE Trans. Inf. Theory*, vol. 53, no. 12, pp. 4787–4792, Dec. 2007.
- [48] H. Huh, G. Caire, H. C. Papadopoulos, and S. A. Ramprasad, "Achieving 'massive MIMO' spectral efficiency with a not-so-large number of antennas," *IEEE Trans. Wireless Commun.*, vol. 11, no. 9, pp. 3226–3239, Sep. 2012.
- [49] J. Flordelis, S. Hu, F. Rusek, O. Edfors, G. Dahman, X. Gao, and F. Tufvesson, "Exploiting antenna correlation in measured massive MIMO channels," in *IEEE 27th Int. Symp. on Personal Indoor and Mobile Radio Communications (PIMRC)*, Valencia, Spain, Sep. 2016, pp. 1–6.
- [50] F. Rusek, A. Lozano, and N. Jindal, "Mutual information of IID complex Gaussian signals on block Rayleigh-faded channels," *IEEE Trans. Inf. Theory*, vol. 58, no. 1, pp. 331–340, Jan. 2012.

- [51] H. Q. Ngo and E. G. Larsson, "No downlink pilots are needed in TDD massive MIMO," *IEEE Trans. Wireless Commun.*, vol. 16, no. 5, pp. 2921–2935, May 2017.
- [52] Y. Du, C. Cao, X. Zou, J. He, H. Yan, G. Wang, and D. Steer, "Measurements and modeling of penetration loss in the range from 2 GHz to 74 GHz," in *Global Telecommunications Conference, GLOBECOM, 2016*, Lisbon, Portugal, Dec. 2016, pp. 1–5.
- [53] R. J. Weiler, M. Peter, W. Keusgen, K. Sakaguchi, and F. Undi, "Environment induced shadowing of urban millimeter-wave access links," *IEEE Commun. Lett.*, vol. 5, no. 4, pp. 440–443, Aug. 2016.
- [54] C. Gustafson, K. Haneda, S. Wyne, and F. Tufvesson, "On mm-wave multipath clustering and channel modeling," *IEEE Trans. Antennas Propag.*, vol. 62, no. 3, pp. 1445–1455, Mar. 2014.
- [55] S. Boyd and L. Vandenberghe, *Convex Optimization*. New York: Cambridge University Press, 2004.
- [56] T. M. Cover and J. A. Thomas, *Elements of Information Theory*, 2nd ed. Wiley-Interscience, 2006.