Coordinated Beam Selection for Training Overhead Reduction in FDD Massive MIMO

Flavio Maschietti[‡†], Gábor Fodor[‡§], David Gesbert[†], Paul de Kerret[†]
[‡]Ericsson Research, Kista, Sweden [§]KTH, Stockholm, Sweden [†]EURECOM, Sophia-Antipolis, France
Email: {flavio.maschietti, david.gesbert, paul.dekerret}@eurecom.fr, gabor.fodor@ericsson.com

***Abstract*—While acquiring channel state information at the transmitter (CSIT) in time division duplexing systems can exploit channel reciprocity, acquiring accurate CSIT in frequency division duplexing massive multiple-input multiple-output systems is not trivial. The two main difficulties in these systems are the scalability of the downlink reference signals and the overhead associated with the required uplink feedback. Although several approaches for ensuring scalability and reducing overhead by leveraging some presumed channel properties have been studied, existing schemes do not offer a fully satisfactory solution. In this work, we propose a novel cooperative method which exploits low-rate beam-related information exchange between the mobile terminals, reduces the overhead under the assumption of the so-called grid-of-beams design, and strikes a balance between CSIT acquisition overhead, user spatial separability and coordination complexity.**

I. Introduction

Massive MIMO (mMIMO) has been identified as one important enabler for achieving higher performance in 5G communications [1]. The original mMIMO implementation is based on time division duplex (TDD) operation, which allows to design near-optimal linear precoders, as downlink channels can be estimated through orthogonal uplink sounding exploiting *channel reciprocity* [1]. In contrast, in FDD mode, channel estimation has to be carried out through downlink reference signals (RSs) and subsequent uplink feedback. In general, there exists a one-to-one correspondence between RSs and antenna elements. Therefore, training and feedback overhead is often associated with *unfeasibility* in the FDD mMIMO regime, where few resource elements are left for data transmission [2].

Nevertheless, operating in FDD remains appealing to mobile operators for crucial reasons, including *i)* most radio bands below 6 GHz are paired FDD bands, *ii)* the base stations (BSs) have higher transmit power available for the RSs than the user equipments (UEs), *iii)* overall deployment and operation costs are reduced as fewer BSs are required in FDD networks [2].

The grid-of-beams (GoB) approach has been proposed in 5G specifications to reduce such overhead [2], [3]. According to this concept, the UEs see low-dimensional virtual *(effective)* channels instead of the actual ones, where the former incorporate the precoder vectors relative to the beams. In particular, one orthogonal RS is allocated to each beam in the GoB codebook [4]. Thus, estimating such effective channels reduces the overhead, as it becomes proportional to the codebook size and independent from the number of antenna elements. Unfortunately, the substantial reduction in training (and feedback) overhead often entails a drastic performance degradation [5] as the digital precoder for data transmission is optimized for reduced channel representations which might not capture the prominent characteristics of the actual channels.

Another option consists in designing the GoB with a large number of beams, and then training a small subset of beams which contains the most relevant channel components [2], [6], [7]. The number of such components depends on several factors, including the frequency band and the radio scattering environment, which is in general beyond the designer's control. Nevertheless, when multi-antenna UEs are considered, statistical beamforming at the UE side can be exploited to let the UEs excite a suitable channel subspace, with the aim to further reduce the number of relevant components to be estimated [8], [9].

Interestingly, a so-far unexplored opportunity to further cut the training overhead arises in the multi-user case. The idea is to capitalize on the common paths which can be found among several UE channels. This paper shows that there exists in fact an interesting *trade-off* between *i)* training the beams which capture the largest channel gains for each UE, and *ii)* training the beams which might capture somewhat weaker paths but are *common* to multiple UEs, so as to reduce the training overhead. The essence of such trade-offs is captured in Fig. 1, where UE 2 can capitalize on its weaker paths to reduce the number of activated beams at the BS side. Another factor entering the trade-offs is that focusing on beams that are common to multiple UEs can lead to reduced *spatial separability* among them.

In this paper, we show that the above *trade-off* can be explored through coordination between the UEs when the GoB is used at both the BS and UE sides. To enforce coordination, we propose a low-overhead exchange protocol exploiting device-to-device (D2D) communications. In our scheme, the UEs exchange beam-related information over the *beam coherence time*, which allows for low-rate updates [10]. Finally, we propose a beam selection algorithm which exploits the exchanged information and enables to strike a balance between overhead and spatial separability, towards spectral efficiency (SE) maximization.

II. System Model

Consider a single cell mMIMO network (refer to Fig. 1), where the BS is equipped with $N_{\mathrm{BS}} \gg 1$ antennas and serves (in downlink transmission) $K \ll N_{\mathrm{BS}}$ UEs with N_{UE} antennas each.

F. Maschietti, D. Gesbert and P. de Kerret are supported by the ERC under the EU's Horizon 2020 research and innovation program (Agreement no. 670896 PERFUME). Part of this work has been carried out while F. Maschietti was visiting the radio department of Ericsson Research in Kista, Sweden. The authors would like to thank G. Klang at Ericsson Research for his support.

A. Channel Estimation with Grid-of-Beams

We assume FDD operation, i.e. the downlink and uplink channels are not reciprocal, and that GoB is exploited at both the BS and UE side. Let us define the beam codebooks $\mathcal{B}_{\text{BS}}$ and $\mathcal{B}_{\text{UE}}$ used for the GoB precoding and combining, respectively, as

$$\mathcal{B}_{\text{BS}} \triangleq \{\mathbf{v}_1,\ldots,\mathbf{v}_{B_{\text{BS}}}\}, \qquad \mathcal{B}_{\text{UE}} \triangleq \{\mathbf{w}_1,\ldots,\mathbf{w}_{B_{\text{UE}}}\}, \tag{1}$$

where $\mathbf{v}_v \in \mathbb{C}^{N_{\text{BS}}\times 1}$, $v \in \{1,\ldots,B_{\text{BS}}\}$, denotes the v-th beamforming vector in $\mathcal{B}_{\text{BS}}$, and $\mathbf{w}_w \in \mathbb{C}^{N_{\text{UE}}\times 1}$, $w \in \{1,\ldots,B_{\text{UE}}\}$, denotes the w-th beamforming vector in $\mathcal{B}_{\text{UE}}$. To ease the notation, we assume that $\mathcal{B}_{\text{UE}}$ is the same across all the UEs.

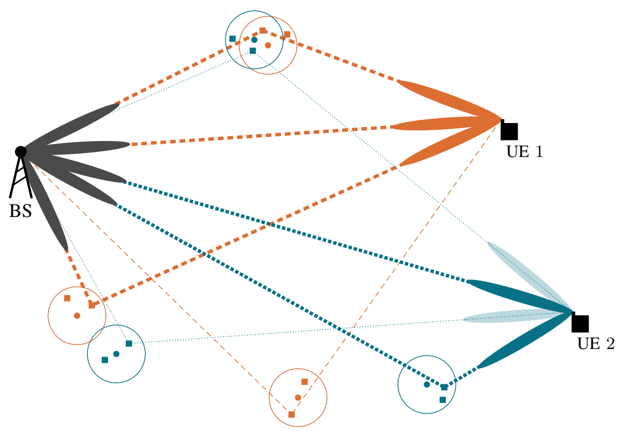

Fig. 1: Scenario example with $K=2$ UEs. Stronger paths are marked in bold. Uncoordinated SNR-based beam selection results in $M_{\text{BS}}=5$ beams to train. In this example, UE 2 could instead opt for the non-bold light blue beams to achieve $M_{\text{BS}}=3$.

A New Radio (NR)-like OFDM-based modulation scheme is assumed. We consider a resource grid consisting of T resource elements. Among those, τM_{BS} are allocated to RSs, and $T-\tau M_{\text{BS}}$ to data, where M_{BS} denotes the number of beams that are trained among the ones in $\mathcal{B}_{\text{BS}}$ and τ is the duration in OFDM symbols of their associated RSs (one each beam [4]). The received training signal $\mathbf{Y}_k \in \mathbb{C}^{M_{\text{UE}}\times\tau}$ at the k-th UE, where M_{UE} is the number of activated beams at the UE side, is expressed as

$$\mathbf{Y}_k = \sqrt{\frac{P}{T}}\mathbf{W}_k^{\text{H}}\mathbf{H}_k\mathbf{V}\mathbf{S} + \mathbf{W}_k^{\text{H}}\mathbf{N}, \qquad \forall k \in \{1,\ldots,K\} \tag{2}$$

where $\mathbf{S} \in \mathbb{C}^{M_{\text{BS}}\times\tau}$ contains the orthogonal (known) RSs, with $\mathbf{S}\mathbf{S}^{\text{H}} = \mathbf{I}_{M_{\text{BS}}}$, $\mathbf{V} \triangleq [\mathbf{v}_1 \ldots \mathbf{v}_{M_{\text{BS}}}] \in \mathbb{C}^{N_{\text{BS}}\times M_{\text{BS}}}$ is the normalized training precoder, $\mathbf{H}_k \in \mathbb{C}^{N_{\text{UE}}\times N_{\text{BS}}}$ is the channel between the BS and the k-th UE, and $\mathbf{W}_k \triangleq [\mathbf{w}_{k,1} \ldots \mathbf{w}_{k,M_{\text{UE}}}] \in \mathbb{C}^{N_{\text{UE}}\times M_{\text{UE}}}$ is the training combiner relative to the k-th UE. Note that both $\mathbf{V}$ and $\mathbf{W}_k$ contain beamforming vectors belonging to the GoB codebooks $\mathcal{B}_{\text{BS}}$ and $\mathcal{B}_{\text{UE}}$. The matrix $\mathbf{N} \in \mathbb{C}^{N_{\text{UE}}\times\tau}$, whose elements are i.i.d. $\mathcal{CN}(0,1)$, denotes the Gaussian thermal noise, while P is the total transmit power available at the BS in the considered coherent (over both time and sub-carriers) frame.

Following the training stage, the UEs are able to estimate their instantaneous GoB effective channels, defined as

$$\bar{\mathbf{H}}_k \triangleq \mathbf{W}_k^{\text{H}}\mathbf{H}_k\mathbf{V} \in \mathbb{C}^{M_{\text{UE}}\times M_{\text{BS}}}, \qquad \forall k \in \{1,\ldots,K\} \tag{3}$$

which are fed back to the BS to close the CSIT acquisition loop.

B. Data Signal Model

Let us consider a single resource element. We assume a single stream per UE and denote with $\mathbf{x} \triangleq [x_1 \ldots x_K] \in \mathbb{C}^{K\times 1}$ the Gaussian-distributed transmitted data vector, with $\mathbb{E}[\mathbf{x}\mathbf{x}^{\text{H}}] = \mathbf{I}_K$. The received data signal $\hat{x}_k$ at the k-th UE is expressed as

$$\hat{x}_k = \sqrt{\frac{P}{T}}\bar{\mathbf{w}}_k^{\text{H}}\bar{\mathbf{H}}_k\bar{\mathbf{V}}\mathbf{x} + \bar{\mathbf{w}}_k^{\text{H}}\bar{\mathbf{n}}_k, \qquad \forall k \in \{1,\ldots,K\} \tag{4}$$

where $\bar{\mathbf{V}} \triangleq [\bar{\mathbf{v}}_1,\ldots,\bar{\mathbf{v}}_K] \in \mathbb{C}^{M_{\text{BS}}\times K}$ is the normalized digital data precoder, $\bar{\mathbf{H}}_k$ is the GoB effective channel between the BS and the k-th UE, $\bar{\mathbf{w}}_k \in \mathbb{C}^{M_{\text{UE}}\times 1}$ is the digital data combiner relative to the k-th UE, and $\bar{\mathbf{n}}_k \triangleq \mathbf{W}_k^{\text{H}}\mathbf{n}_k \in \mathbb{C}^{M_{\text{UE}}\times 1}$.

Let $\mathcal{W} \triangleq \{\mathbf{W}_1,\ldots,\mathbf{W}_K\}$ and $\bar{\mathcal{W}} \triangleq \{\bar{\mathbf{w}}_1,\ldots,\bar{\mathbf{w}}_K\}$. The instantaneous SINR $\gamma_k(\mathbf{V},\bar{\mathbf{V}},\mathcal{W},\bar{\mathcal{W}})$ at the k-th UE is as follows:

$$\gamma_k(\mathbf{V},\bar{\mathbf{V}},\mathcal{W},\bar{\mathcal{W}}) \triangleq \frac{\left|\bar{\mathbf{w}}_k^{\text{H}}\bar{\mathbf{H}}_k\bar{\mathbf{v}}_k\right|^2}{\sum_{j\neq k}\left|\bar{\mathbf{w}}_k^{\text{H}}\bar{\mathbf{H}}_k\bar{\mathbf{v}}_j\right|^2 + (T/P)}, \tag{5}$$

where the dependence on $\mathbf{V}$ and $\mathcal{W}$ is due to $\bar{\mathbf{H}}_k = \mathbf{W}_k^{\text{H}}\mathbf{H}_k\mathbf{V}$.

III. Coordinated Beam Selection and Reporting

In this section, we formulate the optimal beam selection problem, highlighting the role that coordination plays in the considered scenario. To achieve this goal, we introduce the notion of relevant channel components and take a closer look at the beam reporting procedure defined in the current 5G specifications.

A. Exploiting The Relevant Channel Components

In the classical GoB implementation all the beams in the grid are trained regardless of their actual relevance, i.e. $M_{\text{BS}} = B_{\text{BS}}$. As pointed out in Section I, such an operating mode is feasible for small GoBs only, although that in turn leads to high performance loss [2]. In order to avoid exchanging performance for overhead, the intuition is to use a large GoB with few *(accurately)* selected beams to train, so as to keep M_{BS} small.

Let us define the set $\mathcal{M}_k$ containing the relevant channel components (or beam pairs) of the k-th UE as follows:

$$\mathcal{M}_k \triangleq \left\{(v,w) : \mathbb{E}_{\mathbf{H}_k}\left[\left|\mathbf{w}_w^{\text{H}}\mathbf{H}_k\mathbf{v}_v\right|^2\right] \geq \xi\right\}, \tag{6}$$

where ξ is a predefined power threshold, e.g. 10 dB.

Remark 1. The set $\mathcal{M}_k$ is solely dependent on the second order statistics of the channel $\mathbf{H}_k$, for fixed $\mathcal{B}_{\text{BS}}$ and $\mathcal{B}_{\text{UE}}$. In particular, we refer to the *beam coherence time* to denote the coherence time of such statistics. The beam coherence time T_{beam} is much longer than the channel coherence time T_{coh} [10]. □

Likewise, we define the subset $\mathcal{M}_k^{\text{BS}} \subseteq \mathcal{M}_k$ containing the relevant beam pairs relative to the k-th UE, when the latter adopts $\mathbf{W}_k$ as its receive GoB combiner, as follows:

$$\mathcal{M}_k^{\text{BS}}(\mathbf{W}_k) \triangleq \{(v,w) \in \mathcal{M}_k : \mathbf{w}_w \in \mathbf{W}_k\}, \tag{7}$$

where we introduced the notation $\mathcal{M}_k^{\text{BS}}(\cdot)$ to highlight that the set $\mathcal{M}_k^{\text{BS}}$ is dependent on the selected GoB combiner $\mathbf{W}_k$. Indeed, applying some receive beams means focusing on specific relevant beam pairs and neglecting some others. In Section III-C, we will show that a proper *(joint)* combiner selection can be made at the UEs so as to maximize the sum-rate.

B. UE Reporting and Training Overhead

The impact on the performance of the downlink GoB precoding is related to the UE reporting procedure made to assist the BS in the precoder selection. This is a standard procedure in the current 3GPP release [3]. In particular, we assume that the k-th UE has full knowledge of the second order statistics of its channel and reports to the BS the set $\mathcal{M}_k^{\mathrm{BS}}(\mathbf{W}_k)$ – also known as precoding matrix indicator (PMI) [3] – following an appropriate GoB combiner (beam) selection.

In general, UE reporting does not impose any restriction on which GoB precoder is actually used by the BS. Since the downlink precoding is transparent to the UEs – the UEs see only the effective channel – the BS can use whatever GoB precoder $\mathbf{V}$ without the need to inform the UEs [4]. Nevertheless, since combining operation is expected at the UE side, it is essential for the BS to design the GoB precoder so that it best matches the combiners indicated by the UEs. In particular, the BS is expected to train all the relevant beam pairs in $\cup_{k=1}^{K}\mathcal{M}_k^{\mathrm{BS}}(\mathbf{W}_k)$. The latter is thus critical in measuring the overall training overhead for the considered scenario.

Let us define the training overhead $\omega(\mathcal{W})$ as follows:

$$\omega(\mathcal{W}) \triangleq \frac{\tau}{T} M_{\mathrm{BS}} = \frac{\tau}{T}\Big|\bigcup_{k=1}^{K}\mathcal{M}_k^{\mathrm{BS}}(\mathbf{W}_k)\Big|. \qquad (8)$$

The largest overhead reduction is achieved when the UEs coordinate (in the beam domain) so that (8) is minimized, which is

$$\min_{\mathcal{W}} \ \omega(\mathcal{W}). \qquad (9)$$

However, a balance between achievable beamforming gain and required training overhead, as well as multi-user interference *(spatial separability)*, has to be considered in the beam decision process at the UEs. In the following, we formulate the optimal coordinated beam selection problem.

C. Optimal Coordinated Beam Selection

We define the data rate $\mathcal{R}_k$ obtained at the k-th UE as

$$\begin{aligned}&\mathcal{R}_k(\mathbf{V},\bar{\mathbf{V}},\mathcal{W},\bar{\mathcal{W}}) \triangleq \\ &\qquad \big(1-\omega(\mathcal{W})\big)\log_2\big(1+\gamma_k(\mathbf{V},\bar{\mathbf{V}},\mathcal{W},\bar{\mathcal{W}})\big). \qquad (10)\end{aligned}$$

Let $\mathcal{H} = \{\mathbf{H}_1,\dots,\mathbf{H}_K\}$. The optimal GoB beamformers $(\mathbf{V}^*, \mathcal{W}^*)$ are then found through solving the following optimization problem:

$$(\mathbf{V}^*,\mathcal{W}^*) = \underset{\mathbf{V},\mathcal{W}}{\operatorname{argmax}} \ \mathbb{E}_{\mathcal{H}}\left[\max_{\bar{\mathbf{V}},\bar{\mathcal{W}}} \sum_{k=1}^{K}\mathcal{R}_k\big(\mathbf{V},\bar{\mathbf{V}},\mathcal{W},\bar{\mathcal{W}}\big)\right]. \quad (11)$$

Solving (11) is not trivial, due to the mutual optimization of the (constrained) GoB and (unconstrained) data beamformers. A common approach consists in decoupling the design, as the GoB precoder can be optimized through long-term statistical information, whereas the unconstrained data ones can depend on the instantaneous CSIT [11]. The same approach is followed here.

D. Coordinated Beam Selection with Zero-Forcing Precoding

Towards simplification, we assume Zero-Forcing (ZF) as the data precoder at the BS side and a single receive beam per UE, i.e. $M_{\mathrm{UE}}=1$. Thus, we have $\bar{\mathbf{H}} \triangleq \big[\bar{\mathbf{h}}_1^{\mathrm{T}}\dots\bar{\mathbf{h}}_K^{\mathrm{T}}\big]^{\mathrm{T}} \in \mathbb{C}^{K\times M_{\mathrm{BS}}}$, where $\bar{\mathbf{h}}_k \triangleq \mathbf{w}_k^{\mathrm{H}}\mathbf{H}_k\mathbf{V} \in \mathbb{C}^{1\times M_{\mathrm{BS}}}$ is the effective channel vector between the BS and the k-th UE. Therefore, the received data signal $\hat{\mathbf{x}} \in \mathbb{C}^{K\times 1}$ at all the UEs can be expressed as

$$\hat{\mathbf{x}} = \sqrt{\frac{P}{T}}\beta\bar{\mathbf{H}}\bar{\mathbf{V}}\mathbf{x}+\mathbf{n}, \qquad (12)$$

where $\bar{\mathbf{V}} \triangleq \bar{\mathbf{H}}^{\dagger} = \bar{\mathbf{H}}^{\mathrm{H}}\big(\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathrm{H}}\big)^{-1} \in \mathbb{C}^{M_{\mathrm{BS}}\times K}$ is the ZF precoder, $\beta \triangleq \|\bar{\mathbf{H}}^{\dagger}\|_{\mathrm{F}}^{-1}$ is its power normalization factor, and $\mathbf{n}$ is the Gaussian thermal noise vector at all the UEs.

Based on (12), the SINR of the k-th stream (relative to the k-th UE) can be expressed as follows:

$$\gamma_k^{\mathrm{ZF}}(\mathbf{V},\mathcal{W}) \triangleq (P/T)\mathrm{Tr}\Big(\big(\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathrm{H}}\big)^{-1}\Big)^{-1}. \qquad (13)$$

The sum-rate maximization problem in (11) is then turned into a simpler long-term beam selection problem, where the optimal GoB beamformers $(\mathbf{V}^*,\mathcal{W}^*)$ can be found as follows:

$$(\mathbf{V}^*,\mathcal{W}^*) = \underset{\mathbf{V},\mathcal{W}}{\operatorname{argmax}} \ \mathbb{E}_{\mathbf{H}}\Big[\sum_{k=1}^{K}\mathcal{R}_k\big(\mathbf{V},\mathcal{W}\big)\Big], \qquad (14)$$

where

$$\mathcal{R}_k(\mathbf{V},\mathcal{W}) \triangleq \big(1-\omega(\mathcal{W})\big)\log_2\big(1+\gamma_k^{\mathrm{ZF}}(\mathbf{V},\mathcal{W})\big). \qquad (15)$$

In general, the perfect knowledge of the global instantaneous effective CSI at a central coordinator is needed to solve (14). Such information is not available without training all the possible combinations of beams in $\mathcal{B}_{\mathrm{BS}}$ and $\mathcal{B}_{\mathrm{UE}}$, which requires $M_{\mathrm{BS}} = B_{\mathrm{BS}}$ and goes against the main idea behind the training overhead reduction. Moreover, the problem in (14) is a non-trivial subset selection problem. In the following, we will propose a low-overhead approach (exploiting second order statistics) to deal with (14) in a decentralized manner, where each UE *autonomously* feeds back an optimized beam set (PMI) to assist the BS with the design of the GoB precoder $\mathbf{V}$.

IV. Decentralized Coordinated Beam Selection

In this section, we introduce a decentralized beam selection algorithm for the considered scenario. The decentralized approach allows to avoid gathering large-dimensional CSI data at a central node, e.g. the BS, thus reducing additional feedback and coordination overhead. Towards this goal, we first introduce an approximate rate metric for algorithm derivation purposes. In particular, we reformulate (14) through *Jensen's inequality*, which allows to exploit the second order channel statistics instead of the instantaneous effective CSI.

In this respect, an approximate rate $\bar{\mathcal{R}}$ for the generic UE can be defined as follows:

$$\bar{\mathcal{R}}(\mathbf{V},\mathcal{W}) \triangleq \big(1-\omega(\mathcal{W})\big)\log_2\big(1+\mathbb{E}_{\mathbf{H}}\big[\gamma^{\mathrm{ZF}}(\mathbf{V},\mathcal{W})\big]\big), \qquad (16)$$

where we have removed the dependence on k as the SINR γ^{ZF} is the same for each UE (refer to Eq. (13)).

We focus now on the log factor in (16). In particular, the expected SINR for the generic UE can be expressed as

$$\mathbb{E}_{\mathbf{H}}\left[\gamma^{\mathrm{ZF}}(\mathbf{V},\mathcal{W})\right]=(P/T)\mathbb{E}_{\mathbf{H}}\left[\left(\sum_{k=1}^{K}\lambda_k^{-1}\right)^{-1}\right], \qquad (17)$$

where λ_k is the k-th eigenvalue of the matrix $\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathrm{H}}$.

In addition, due to the inverse function being convex, the expected SINR in (17) can be upper bounded as

$$(P/T)\mathbb{E}_{\mathbf{H}}\left[\left(\sum_{k=1}^{K}\lambda_k^{-1}\right)^{-1}\right]\leq(P/T)\mathbb{E}_{\mathbf{H}}\left[K^{-2}\sum_{k=1}^{K}\lambda_k\right]. \qquad (18)$$

For algorithm derivation purposes, we will thus introduce the following approximate rate metric:

$$\hat{\mathcal{R}}(\mathbf{V},\mathcal{W})\triangleq\left(1-\omega(\mathcal{W})\right)\log_2\left(1+\hat{\gamma}(\mathbf{V},\mathcal{W})\right), \qquad (19)$$

where

$$\begin{aligned}\hat{\gamma}(\mathbf{V},\mathcal{W})&\triangleq\rho\mathbb{E}_{\mathbf{H}}\left[\sum_{k=1}^{K}\lambda_k\right]\\&=\rho\mathbb{E}_{\mathbf{H}}\left[\mathrm{Tr}\left(\bar{\mathbf{H}}\bar{\mathbf{H}}^{\mathrm{H}}\right)\right]\\&=\rho\sum_{k=1}^{K}\mathrm{Tr}\Big(\left(\mathbf{V}^{\mathrm{T}}\otimes\mathbf{w}_k^{\mathrm{H}}\right)\mathbf{\Sigma}_k\left(\mathrm{conj}(\mathbf{V})\otimes\mathbf{w}_k\right)\Big), \qquad (20)\end{aligned}$$

having defined the constant $\rho\triangleq(P/T)K^{-2}$ and the channel covariance matrix of the k-th UE $\mathbf{\Sigma}_k\triangleq\mathbb{E}_{\mathbf{H}_k}\left[\mathrm{vec}(\mathbf{H}_k)\mathrm{vec}(\mathbf{H}_k)^{\mathrm{H}}\right]$, and where $\otimes$ denotes the Kronecker product.

The optimal GoB beamformers $(\mathbf{V}^*,\mathcal{W}^*)$ can then be found through solving the following optimization problem:

$$(\mathbf{V}^*,\mathcal{W}^*)=\underset{\mathbf{V},\mathcal{W}}{\mathrm{argmax}}\ \hat{\mathcal{R}}(\mathbf{V},\mathcal{W}). \qquad (21)$$

Although no instantaneous information is needed to solve (21) through (20), the problem in (21) still requires a central coordinator knowing $\mathbf{\Sigma}_k\ \forall k$ and which dictates the beam strategies to each UE. In the following section, we turn to a heuristic approach to solve (21) in a decentralized fashion.

A. Decentralized Coordinated Beam Selection Algorithm

In order to achieve decentralized coordination, we propose to use a hierarchical information structure requiring small overhead. In particular, an *(arbitrary)* order among the UEs is established, for which the k-th UE has access to the beam decisions carried out at the (higher-ranked) UEs $k+1,\ldots,K$. This configuration is obtainable through e.g. D2D communications[1]. We further assume that such exchanged beam information is *perfectly* decoded at the intended UEs. The full signaling sequence of the proposed hierarchical beam selection is given in Fig. 2.

Let us consider w.l.o.g. the beam selection at the k-th UE, i.e. at the k-th step of the algorithm. We define the set $\mathcal{W}_{k+1}\triangleq\{\mathbf{w}_{k+1}^*,\ldots,\mathbf{w}_K^*\}$ containing the beam decisions which have been fixed prior to the k-th step. According to the hierarchical structure, the k-th UE knows the set $\mathcal{B}^{\mathrm{fix}}(\mathcal{W}_{k+1})\triangleq\cup_{j=k+1}^{K}\mathcal{M}_j^{\mathrm{BS}}(\mathbf{w}_j^*)$ (refer to Fig. 2). Therefore, the k-th UE can construct a *partial* GoB precoder $\mathbf{V}_{k+1}$ containing the beamforming vectors relative to the indexes in $\mathcal{B}^{\mathrm{fix}}(\mathcal{W}_{k+1})$. Likewise, the k-th UE can compute a *partial* $\omega(\mathcal{W}_{k+1})$.

[1]The 3GPP Release 16 is expected to support point-to-point side-links which facilitate cooperative communications among the UEs with low resource consumption [12]. The NR side-link is thus a cornerstone for the proposed scheme.

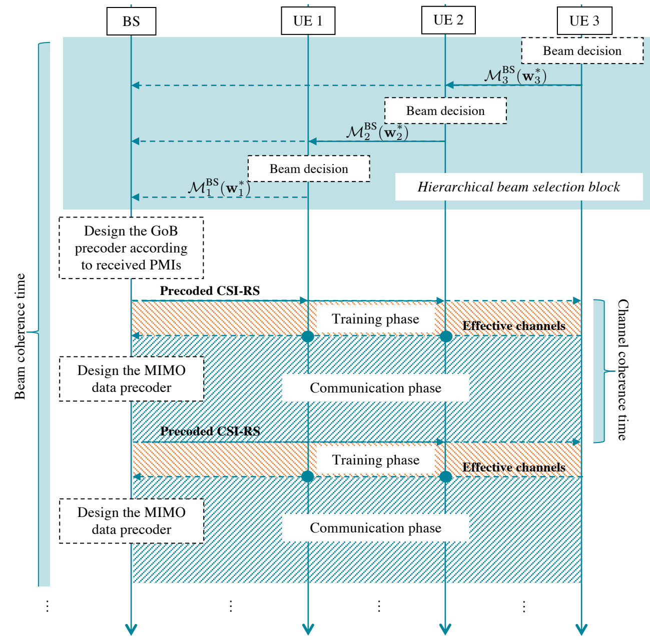

Fig. 2: Signaling sequence of the proposed coordinated beam selection for $K=3$. The beam decision made at each UE leverages the D2D-enabled beam-related information (PMI) coming from higher-ranked UEs in a hierarchical fashion.

The proposed decentralized beam selection $\mathbf{w}_k^*$ at the k-th UE can be then expressed in a recursive manner as follows:

$$\mathbf{w}_k^*=\underset{\mathbf{w}_k}{\mathrm{argmax}}\ f_k\left(\left[\mathbf{V}_k\mathbf{V}_{k+1}\right],\{\mathbf{w}_k,\mathcal{W}_{k+1}\}\right), \qquad (22)$$

where $\mathrm{col}_m(\mathbf{V}_k)=\mathbf{v}_m\ \forall m\in\mathcal{M}_k^{\mathrm{BS}}(\mathbf{w}_k)$, and

$$f_k(\mathbf{V},\mathcal{W})\triangleq\left(1-\omega(\mathcal{W})\right)\mathrm{Tr}\Big(\left(\mathbf{V}^{\mathrm{T}}\otimes\mathbf{w}_k^{\mathrm{H}}\right)\mathbf{\Sigma}_k\left(\mathbf{V}^*\otimes\mathbf{w}_k\right)\Big). \qquad (23)$$

The intuition behind the proposed algorithm is to let the k-th UE select the $\mathbf{w}_k\in\mathcal{B}_{\mathrm{UE}}$ maximizing the k-th term of the sum in (20), while taking into account the pre-log factor $(1-\omega(\mathcal{W}))$, in a *greedy* manner. This decentralized problem can be addressed using linear (exhaustive) search in the codebook $\mathcal{B}_{\mathrm{UE}}$ at each UE. In this case, the linear search does not involve a large computational burden as $|\mathcal{B}_{\mathrm{UE}}|=B_{\mathrm{UE}}$, which is expected to be in the order of N_{UE}, i.e. small in practical scenarios.

V. Simulation Results

We evaluate here the performance of the proposed algorithm for $K=7$ UEs. We assume $N_{\mathrm{BS}}=64$ and $N_{\mathrm{UE}}=4$. For the codebooks, we consider $B_{\mathrm{BS}}=32$ and $B_{\mathrm{UE}}=8$. The beamforming vectors in $\mathcal{B}_{\mathrm{BS}}$ and $\mathcal{B}_{\mathrm{UE}}$ are steering vectors sampled according to the inverse cosine function over the interval $[0,\pi]$, as in [7]. Furthermore, we assume that each UE is allowed to indicate 4 relevant beam pairs to the BS at most, i.e. the set $\mathcal{M}_k^{\mathrm{BS}}(\mathbf{w}_k^*)$ is truncated to its 4 strongest elements $\forall k$, according to 3GPP specifications [4]. We assume that the UEs use the popular least square (LS) method [8] to estimate their instantaneous effective channels, which are then fed back to the BS for ZF precoder design (refer to Fig. 2). We consider the realistic 3GPP micro-cell scenario operating in 2.1 GHz assuming the spatial channel model. Further details on the channel model can be found in [13]. All the plotted data rates are the averaged – over 10000 Monte-Carlo iterations – instantaneous sum-rates.

A. Results and Discussion

We compare the proposed algorithm with the following algorithms: *i)* uncoordinated, and *ii)* coordinated *(overhead minimization)*. In the uncoordinated benchmark, no coordination among the UEs is exploited and the beam decisions at the UEs are designed to maximize their local SNR (refer to Fig. 1). The *overhead minimization* algorithm follows the same hierarchical structure proposed in Section IV-A, but where the UEs aim to minimize the pre-log factor in (8) rather than solving (22).

In Fig. 3, we show the sum-rate as a function of the SNR, for a channel coherence time $T_{\text{coh}} = 16$ ms. Both the coordinated algorithms outperform the uncoordinated benchmark, with equal sum-rate values obtained with up to 5 dBs less. The performance gap between the coordinated algorithms decreases with the SNR, as the *overhead minimization* algorithm becomes less prone to channel estimation errors due to poor beam selection in terms of beamforming gain.

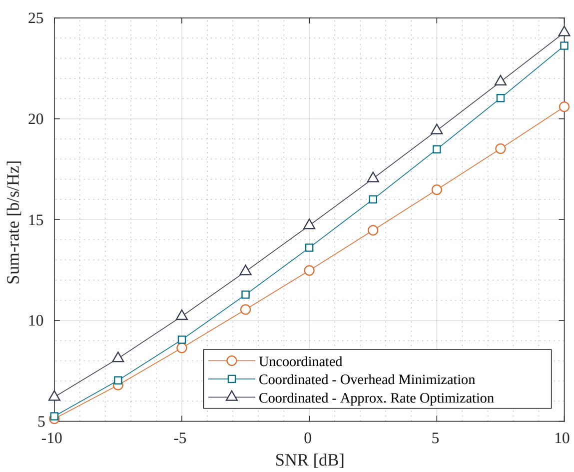

Fig. 3: Sum-rate vs SNR. The channel coherence time is 16 ms. The coordinated algorithms outperform the uncoordinated one.

In Fig. 4, we show the SE gain over the uncoordinated benchmark as a function of the channel coherence time T_{coh}. Two areas can be identified: *i)* $T_{\text{coh}} < 20$ ms, i.e. vehicular or fast pedestrian channels, where coordinated beam selection is essential and lead to performance gains up to 100%; *ii)* $T_{\text{coh}} \geq 20$ ms, where less gains (up to 10%) are achieved with the coordinated algorithms. In particular, the *overhead minimization* algorithm performs even worse than acting without coordination for $T_{\text{coh}} \geq 20$. This is because the training overhead becomes negligible for long channel coherence times, and it is more important to focus on the log factor (i.e. SINR) in (16).

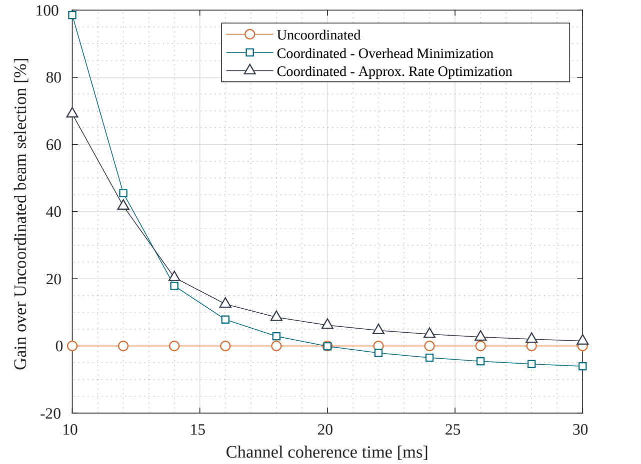

Fig. 4: Gain over uncoordinated beam selection vs channel coherence time. The SNR is 5 dB. Less gains are achieved with the proposed coordinated algorithm as T_{coh} increases.

VI. Conclusions

In this paper, we have shown that beam-domain coordination between the UEs offers a convenient mean to reduce the training overhead in FDD mMIMO networks. We have proposed a decentralized beam selection algorithm exploiting D2D communications and which allows for substantial performance gains, in particular under fast channels. In the proposed algorithm, the overhead reduction is achieved through letting the UEs align on some common effective channel subspaces. Future works include thus the extension to more effective transceivers, capable to cope with both increased spatial correlation and channel estimation errors.

References

[1] E. G. Larsson, O. Edfors, F. Tufvesson, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Commun. Mag.*, Feb. 2014.

[2] W. Zirwas, M. B. Amin, and M. Sternad, "Coded CSI reference signals for 5G — Exploiting sparsity of FDD massive MIMO radio channels," in *Proc. IEEE WSA*, Mar. 2016.

[3] 3GPP, "NR; physical layer procedures - Rel. 15," TS 38.214. Dec. 2018.

[4] E. Dahlman, S. Parkvall, and J. Sköld, *5G NR: the Next Generation Wireless Access Technology*. Academic Press, 2018.

[5] J. Flordelis, F. Rusek, F. Tufvesson, E. G. Larsson, and O. Edfors, "Massive MIMO performance — TDD versus FDD: What do measurements say?" *IEEE Trans. Wireless Commun.*, Apr. 2018.

[6] A. Adhikary, J. Nam, J. Ahn, and G. Caire, "Joint spatial division and multiplexing — The large-scale array regime," *IEEE Transactions on Information Theory*, Oct. 2013.

[7] F. Maschietti, D. Gesbert, P. de Kerret, and H. Wymeersch, "Robust location-aided beam alignment in millimeter wave massive MIMO," *Proc. IEEE GLOBECOM*, Dec. 2017.

[8] N. N. Moghadam, H. Shokri-Ghadikolaei, G. Fodor, M. Bengtsson, and C. Fischione, "Pilot precoding and combining in multiuser MIMO networks," *IEEE J. Sel. Areas Commun.*, July 2017.

[9] P. Mursia, I. Atzeni, D. Gesbert, and L. Cottatellucci, "Covariance shaping for massive MIMO systems," in *Proc. IEEE GLOBECOM*, Dec. 2018.

[10] V. Va, J. Choi, and R. W. Heath, "The impact of beamwidth on temporal channel variation in vehicular channels and its implications," *IEEE Trans. Veh. Technol.*, June 2017.

[11] A. Alkhateeb, G. Leus, and R. W. Heath, "Limited feedback hybrid precoding for multi-user millimeter wave systems," *IEEE Trans. Wireless Commun.*, Nov. 2015.

[12] 3GPP, "NR; study on NR V2X," Work item RP-181429. June 2018.

[13] ——, "SCM for MIMO simulations - Rel. 15," TR 25.996. June 2018.