Coded CSI Reference Signals for 5G - Exploiting Sparsity of FDD Massive MIMO Radio Channels

Wolfgang Zirwas, Muhammad Bilal Amin
Nokia Networks
Munich, Germany
Email: {bilal.amin.ext, wolfgang.zirwas}@nokia.com

Mikael Sternad
Signals and Systems
Uppsala University, Sweden
Email: Mikael.Sternad@signal.uu.se

***Abstract*—Future 5G systems are expected to provide higher performance, partly unleashed by massive MIMO as well as tight cooperation like joint transmission CoMP. For paired and unpaired spectrum below 6 GHz RF-frequency bands, frequency division duplex as well as time division duplex (FDD/TDD) has to be supported. The use of large cooperation areas over several cells together with massive MIMO downlink transmission is challenging in particular for FDD systems, due to two requirements. First, the channel state information (CSI) for downlinks from a large number of antennas has to be obtained without unreasonable overhead due to the transmission of orthogonal downlink reference (pilot) signals from these antennas. Second, relevant channel estimates have to be made available at the network side without an unrealistic uplink control signaling overhead. Pilot contamination has been extensively discussed in the literature as upper bounding performance, due to either exploding overhead for orthogonal reference signals or due to limited CSI accuracy, which is detrimental especially for sensitive interference cancellation schemes. We here propose a strategy for channel estimation of a large number of FDD downlink channels that works without an unreasonable pilot overhead. Analysis of channel statistics for urban macro scenarios applying massive MIMO - potentially combined with strong UE beamforming - reveal a sparse nature of the typical channel matrices. We propose a coded allocation of CSI reference signals, inherently exploiting this sparse nature. It allows accurate CSI estimation of UE individual subsets of relevant channel components despite a very low reference signal overhead of typically less than 5 percent.**

Keywords — massive MIMO; channel estimation; CoMP; pilot contamination;

I. INTRODUCTION

Massive MIMO or, according to 3GPP terminology, full dimension MIMO (FD MIMO) is one of the main features for the evolution of LTE as well as for future 5G systems. Expected benefits range from very high spectral efficiency, high coverage and capacity to high energy efficiency. From the beginning, pilot contamination - i.e. the mutual interference between multiple reference signals (pilot signals) - has been identified as one of the main challenges that limits massive MIMO performance [1]. To allow the estimation of the massive number of channels for antenna elements, time division duplex (TDD) transmission that generates channel reciprocity is often assumed. UEs will then transmit a single to few uplink (UL) sounding reference signals (SRS) to the evolved Node B (eNB), where all channel state information (CSI) for all antenna elements of the massive MIMO array can be estimated simultaneously in one step. For a well calibrated antenna array, the eNB can then use these uplink channel measurements from all UEs as estimates of the downlink channels, for a proper multi user MIMO (MU MIMO) precoding of the subsequent downlink transmission.

With TDD, high spectral or energy efficiency can be achieved by strong multi stream or MU MIMO transmission, but multiple UEs then have to transmit their reference signals (RS) simultaneously. Mutual crosstalk can then be avoided by the use of orthogonal CSI RSs, but in case of a massive number of streams, the resulting pilot signaling overhead will become large. In cellular networks the situation is even worse due to the unavoidable inter cell interference. Reducing this interference would require orthogonal resources for SRS over multiple cells, i.e. some form of frequency reuse. Many proposals exist on how one might for example exploit the spatial covariance structure to reduce the overhead or to increase the achievable CSI accuracy. For example, in [2] the optimum trade-off between pilot and user data power has been analyzed. To obtain an overview of currently discussed options, the interested reader is referred to [3]. In the present paper the focus will be on a different setup, namely on massive MIMO systems in frequency division duplex (FDD) mode, where CSI has to be estimated by UEs in downlink (DL) and will be reported afterwards on UL control channels (see also [4]). For large massive MIMO arrays, the overhead for orthogonal reference signals would then become extremely large in case of per antenna element individual reference signals. From that point of view, FDD seems to be even more challenging than TDD, raising the question why to go for FDD instead of TDD? The first motivation is that lower RF frequency bands below 6 GHz are the most valuable ones due to their large coverage. Performance gains in these bands will be appreciated by mobile network operators. But, below 6 GHz many of the RF bands are paired FDD bands, motivated by higher coverage compared to TDD. For TDD, intermittent - and therefore shorter - average transmission time limits the average Tx power for a given maximal transmit power, especially in the uplink.

The use of FDD based constantly transmitted DL CSI RS also has other benefits, as eNBs have higher Tx power than UEs and do not suffer from limited battery power. Furthermore, for FDD all UEs can listen to CSI RSs and estimate radio channels for relatively longer time periods as compared to rarely transmitted SRS. This improves CSI estimation accuracy and is furthermore a prerequisite for a proper channel prediction. Channel prediction is seen as one of the main enablers for future 5G systems that rely on joint transmission cooperative multi point (JT CoMP) for interference mitigation [5].

It is therefore of interest for several reasons to explore concepts that, in combination, will enable us to use massive MIMO antennas and coherent joint processing from multiple sites for FDD downlink transmission. To accomplish this, we have to estimate the relevant channel components (CC). Here, CCs are the complex-valued channels between a certain Tx-beam and the UE Rx-beam and moreover relevant CCs are those being within a predefined power window with respect to the strongest CC of that UE. Here single rank per UE is assumed, i.e. each UE forms one single beam for its two, four or even eight UE antennas.

Special care has to be taken with respect to the overall overhead for CSI RSs. In our proposed scheme, the first step is to form a limited number of effective radio channels. A grid of beam (GoB) concept transforms for example a uniform linear array consisting of e.g. 64 x 16 antenna elements in azimuth/elevation direction to e.g. eight virtually precoded beams per cell [6]. Even for this GoB concept - typically generated from a set of DFT precoders - the overall number of effective channel components becomes large in a CoMP setting. Enlarged cooperation areas might comprise up to nine cells with eight beams each, or equivalently 72 overall beams. To avoid inter cell interference - or in other words pilot contamination - might then require hundreds of orthogonal CSI reference signals.

As a second step, we therefore here propose to use so called coded CSI reference signals. The aim is to enable the UE to estimate the up to K strongest CCs out of a much larger set of I CCs, using only K CSI RS. We may use K = 2N resource elements (RE) for a set of CSI RSs. This set of REs is partitioned into two groups, each of size N. Then, if for example K = 40 so N = 20, this will allow UEs to estimate a limited set of for example 2N = 40 relevant out of N^2 = 400 potential channel components. Another variant will be presented below that can estimate K CCs out of an even larger pool of potential CCs.

When using coded CSI reference signals, each effective channel will transmit individually precoded - or coded - CSI reference signals on a set of pre-allocated CSI resource elements. UEs reconstruct their set of relevant channel components by applying a properly calculated Moore Penrose Pseudo inverse of the precoding matrix[1].

An important point here is that each UE can reconstruct a different individual set of relevant channel components, as long as the number of relevant CCs for this UE is sparse, i.e. not larger than K. For that purpose a full rank and good condition of any potential sub matrix has to be ensured, and for this reason we propose the use of a Vandermonde like coding matrix.

This concept exploits the inherent sparseness of massive MIMO channel matrices as it has been found for typical urban macro outdoor scenarios. The sparsity reduces the number of channels that need to be estimated. It also limits the required uplink feedback reporting overhead. Below, the sparse radio channel conditions are discussed in Section II. Section III explains the coded CSI concept which is then evaluated in detail in Section IV, while Section V provides the conclusions.

[1]There exists a known concept, that is also denoted coded reference signals, where K orthogonal codes are used on K REs, similarly as in CDMA, but this can not be used for labeling I $>$ K CCs, and out of these estimate up to K relevant CCs.The present coded CSI concept instead uses a set of non-orthogonal coded CSI RS vectors to accomplish this task.

II. Sparse 5G Radio Channels

Massive MIMO downlink transmission - potentially in combination with JT CoMP covering many sites or cells - uses a large number of antenna elements. For a straight forward implementation, hundreds to even thousands of CCs would have to be estimated and in case of FDD also reported to the eNB. For a suitable system design - as being sketched in this section - the overall number of CCs can be reduced significantly. Furthermore, the channel matrix combining the relevant CCs of all simultaneously served UEs will be sparse.

System Assumptions

One of the main targets for a clean slate 5G system is to mitigate interference between cells and sites from the scratch as well as to smoothly integrate massive MIMO. Below 6 GHz RF frequency bands are scarce and precious, mandating high spectral efficiency and accordingly effective MU MIMO modes. There will be paired and unpaired bands so that FDD as well as TDD have to be supported. Furthermore, interference mitigation between cells and sites has to be integrated. This could be achieved by a so called interference mitigation framework IMF-A including JT CoMP as one main ingredient as explained in more detail in [7].

The IMF-A framework starts from enlarged cooperation areas comprising three sites or nine cells being decoupled from each other by a suitable interference floor shaping technique [8]. Due to this decoupling, the performance for a single cooperation area provides a quite accurate estimate of the performance of a full cellular network. Each cell - i.e. each sector of the three cooperating macro sites - is assumed to be equipped with a massive MIMO antenna array for example of size 16 x 16 or even larger. The massive number of antenna elements form by digital or hybrid beamforming a fixed grid of beams, i.e. downscale for example 32 x 16 = 512 physical antennas to 8 x 2 = 16 effective or virtual antenna ports (AP). The term AP has been introduced by 3GPP LTE, where each AP is mapped more or less to one individual reference signal, e.g. CSI RS.

The GoB concept reduces the number of virtual APs to e.g. 16, but orthogonality of CSI RSs is needed at least for all nine cells forming a single cooperation area being in this case 16 x 9 = 144 APs. For the typical cross polarized antenna elements the number of APs will increase further to over 288. For a 3GPP LTE system, the allocation of 288 orthogonal CSI RS resource elements per resource block bandwidth would - under the assumption of a 5 ms CSI RS periodicity - result in an overhead for CSI RSs of $>$ 35 percent. The aim of using coded CSI RSs is to reduce this overhead to more reasonable numbers like five percent or less, while not sacrificing - or even possibly improving - the achievable accuracy of the channel estimation.

Relevant Channel Components

An important prerequisite for the coded CSI concept - as being explained further below - is a sparse channel matrix, i.e. UEs see only a limited set of relevant CCs. Therefore, here in a first step the expected channel conditions for future 5G massive MIMO and CoMP scenarios are being evaluated on high level. Raytracing as well as system level simulations based on the Quadriga channel model [9] have been used to evaluate the typical channel structure of real world massive MIMO scenarios.

Ray tracing simulations have been conducted for Schwabing, close to the city center of Munich, for a single cooperation area (CA). In Figure 1, the black columns indicate the locations of the three sites with inter-site distances (ISD) close to 500 m. At each site there are three 120 degree sectors. Each sector has with a 32 x 16 massive MIMO antenna array, forming a regular GoB consisting of ten single polarized beams in this case, with a half power beam width (HPBW) of only 2 degrees. Vertically the beam patterns are those of a classical Kathrein antenna with a vertical HPBW of 6 degree and a tilting angle of 7 degree ensuring full coverage. Figure 1 verifies strong coverage within the center of the cooperation area, where a Rx power of at least -80 dBm is being achieved almost everywhere. This figure provides the Rx power for the outdoor UEs, while the locations of the buildings are indicated by the dark blue areas equivalently to -130 dBm. The strong coverage is achieved as a single user CoMP mode has been used in this case for illustration of the main CoMP area.

More interesting are the results in Figure 2, which provide for the same scenario the number of relevant channel CCs within a power window of 20 dB. The 3D plot is for better visibility color coded according to the legend. Most areas within the center of the cooperation area have low (light blue below 20) or moderate (green to yellow below 50) relevant CCs. At the border of the cooperation area, there are many red areas with high number of relevant CCs. This can be explained by the multiple reflections, together with strong shadowing for longer distances in NLOS scenarios. The result is a general interference floor with the lack of any stronger CCs. For the intended interference mitigation scheme, these border areas are of less relevance as each cooperation area serves only its CA center users.

Based on Quadriga system level simulations, the number of relevant CCs within a power window of 20 dB were on average in the range of 20 to 30 per UE. So from overall 512 x 2 x 9 = 9216 antenna elements, there remains 16 x 9 = 144 transmitted effective CCs (= Tx beams), from which the UEs receive on average only about 20 relevant CCs, or 20/144 $\approx$ 14 percent of the Tx-beams. This justifies the term sparse for the overall channel matrix $\mathbf{H}$, which combines the relevant CCs for all simultaneously served UEs. Other means might reduce the number of relevant CCs further, like for example UE sided beamforming or some opportunistic selection of subsets of Tx-beams [10]. From system level simulations, it is known that a threshold of TH = 20 to 25 dB achieves already good CoMP performance with respect to SINR or sum rate per cooperation area, with tolerable degradations compared to the full channel knowledge [7].

One should note that for future 5G systems, some type of multi carrier modulation is being assumed similar as known from 3GPP LTE. The Rx-power of a CC as discussed above has been calculated by averaging over all the sub carriers, similar as done for the reference signal received power (RSRP) measurements known from 3GPP LTE. These are essentially pathloss measurements over the system bandwidth, which for 5G is expected to be in the range of about 20 to 100 MHz or even a 400 MHz noncontiguous spectrum.

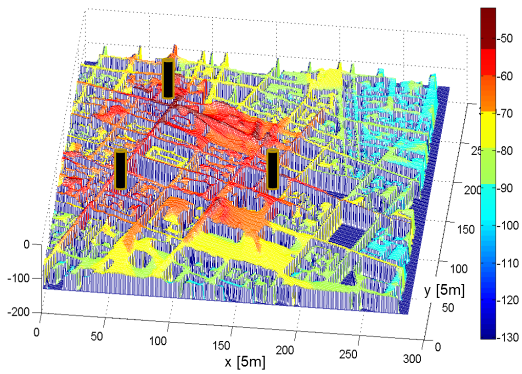

Fig. 1: Coverage of single CA in Munich Schwabing

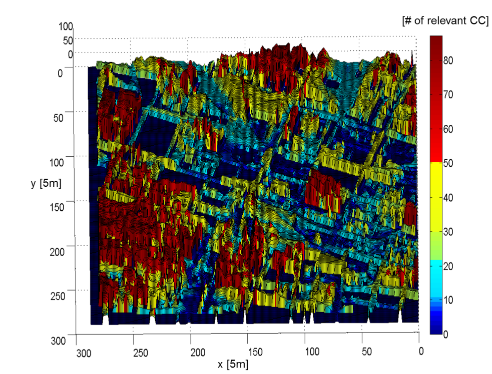

Fig. 2: Number of relevant CCs over location

III. Coded CSI Reference Signal Concept

The main target of the coded CSI RS concept is to exploit the sparse nature - as derived above - of typical future 5G scenarios for an efficient channel estimation solution with low CSI RS overhead.

In case all UEs would see the same set of sparse relevant CCs, one could easily limit the transmission of CSI RSs to those relevant antenna ports (AP), i.e. Tx-beams of the GoB beamformers. In case of MU-MIMO and UEs having individual different sub sets of relevant CCs as illustrated in Figure 3, more advanced solutions will be needed to enable UE specific estimation of relevant CCs with a limited number of CSI RSs.

With a conventional set of K orthogonal CSI RS, each AP would transmit one of them (i.e. transmit a symbol at one of the K corresponding resource units, while transmitting no power at the others). With coded CSI RS, each AP transmits on multiple CSI RS, in a pattern that is known at the receiver. The pattern is organized for sets of APs so that each of the I

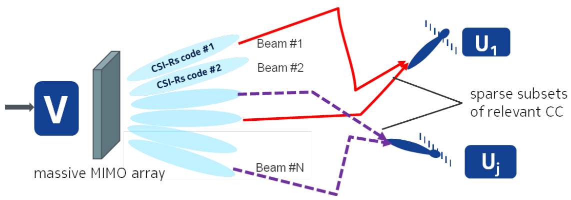

Fig. 3: UE individual subsets of relevant CCs with single UE beamformer instead of single UE antenna

$\gg$ K access points in a set will transmit a unique vector of K pilots. This is the key property that will limit the required pilot overhead.

For illustration of the basic idea of coded CSI RSs, let us consider a very simple case of a system that uses K = 6 CSI RS (placed e.g. on different subcarriers) to estimate flat-fading channels from I = 9 APs. UE-individual subsets of relevant CCs consist of at most 6 out of these 9 CCs. In Figure 4 left, the nine APs with numbers 1 to 9 are allocated in a 3 x 3 matrix. Six CSI RSs - termed CSI 1 to CSI 6 - are transmitted according to the arrows at the borders of the matrix, i.e. CSI 1 is transmitted from AP 1, 4 and 7 or CSI 4 from APs 1, 2 and 3. The reason to call the concept *coded CSI RSs* is the similarity to e.g. Reed Solomon codes for detection of erroneous symbols. For example AP 5 can be identified by the detection of a the combination of CSI 2 and 5 and similarly AP 6 by CSI 3 and 5, as visible from the circles in the left-hand figure in Figure 4. In the middle figure of Figure 4, magenta colored element indicates that the according AP (x-axis) transmits the corresponding CSI RS (y-axis). For example AP 1 transmits CSI RS 1 and 4. We have so far assumed that all active CSI RSs are transmitting the same symbol, i.e. a '1'.

In a noise-free case, UEs receive the $\mathcal{C}^{K\times1}$ vector $\mathbf{y} = \mathbf{C}^{33\_base} * \mathbf{h}$, where $\mathbf{h}$ is the $\mathcal{C}^{I\times1}$ vector comprising all CCs and $\mathbf{C}^{33\_base}$ is the K x I matrix of zeros and pilot symbols, that is exemplified in the middle figure of Figure 4. Element k of $\mathbf{y}$ represents the received signal $\mathbf{y}_k$ at reference signal position k. It is the sum over all CCs for which CSI RS k is active (transmits a '1'). Now, assume that in the given example, the only relevant CCs are those related to AP 5 and 6, while all other CCs are zero or at least weaker than a certain threshold TH. Then, the UE receives on $\mathbf{y}_5$ the sum of CC 5 and 6, on $\mathbf{y}_2$ CC 5 and on $\mathbf{y}_3$ CC 6. This can be used to construct a linear system of three equations for two variables, which can be easily solved. From another perspective one can note that each CC is in this case estimated twice from two different CSI RSs. This provides the inherent potential to improve the estimation accuracy.

The general solution for estimation of the relevant CCs of the channel vector $\mathbf{h}_{\mathbf{i}_{rel}}$ is by performing the Moore-Penrose matrix inversion of $\mathbf{C}_{rel} = \mathbf{C}^{33\_base}_{1..K,\mathbf{i}_{rel}}$, i.e. $\hat{\mathbf{h}}_{\mathbf{i}_{rel}} = pinv(\mathbf{C}_{rel}) * \mathbf{y}$, with $\mathbf{i}_{rel}$ being the indices to the relevant CCs or APs. In our example above $\mathbf{i}_{rel}$ = $\{5,6\}$. To generate the matrix $\mathbf{C}_{rel}$, the UEs have to identify the set i_{rel} of relevant antenna ports based on e.g. RSRP measurements as known from 3GPP LTE. The CSI reporting will be limited to these relevant APs or CCs $\mathbf{i}_{rel}$ so that the eNBs are informed by the UEs about their actual set of relevant CCs. As pathloss is a large scale parameter, the set of relevant CCs will have to be updated only quite rarely, e.g. every 500 ms with accordingly low extra overhead.

In principle the goal in the example above would be to allow channel estimation for any subset of relevant CCs with cardinality $|\mathbf{h}_{\mathbf{i}_{rel}}|$ = 6 since we can create six linear equations with six CSI RS. Unfortunately $\mathbf{C}^{33\_base}$ as it is setup in Figure 4 does not fulfill this target as can be easily concluded from Figure 4 middle. If the subset of relevant CCs is AP1 to AP6, the CSI RS 6 will be zero (does not transmit) for all these six APs. The rank of the according coding matrix $\mathbf{C}_{rel}$ is therefore at maximum 5 if these APs are of interest. The channels to all six APs AP1-AP6 can therefore not be estimated simultaneously based on this received information only. The alternative coding matrix $\mathbf{C}_{33}$ as depicted in Figure 4 directly beside $\mathbf{C}^{33\_base}$, has full rank - i.e. rank 6 - for any subset of 6 out of 9 relevant CCs. This can be easily verified by testing of all possible subsets in this simple case of only 9 APs.

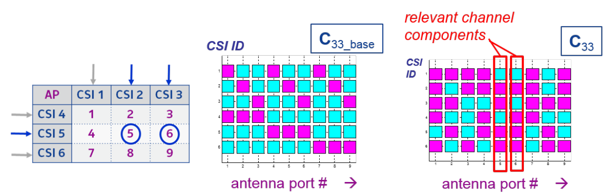

Fig. 4: Basic concept of coded CSI reference signals; magenta/blue stands for '1'/'0' code elements

Vandermonde like coding matrix

For larger coding matrices, as needed for the target scenario with some tens of relevant CCs and potentially some hundreds or more APs, testing of all possible subset combinations leads to prohibitive large complexity and finding a suitable coding matrix is even more challenging. However, full rank can be obtained by construction. To ensure inherently full rank for any subset

$$\mathbf{i}_{rel} \text{ with cardinality } \mathcal{K} \ll I$$

out of overall I APs, it is proposed to use a Vandermonde like coding matrix $\mathbf{C}_V \in \mathcal{C}^{K\times I}$ were element (k,i) is defined in equation (1) below, with $k \in 1...K$ referring to the index of the resource element (RE) being used for CSI RS k and $i \in 1...I$ as index for the APs:

$$\mathbf{C}_V(k, i, \phi_v) \quad = \quad \exp(j(k\phi_v)^i). \tag{1}$$

The scalar real-valued phase parameter ϕ_v defines the whole matrix $\mathbf{C}_V$ and for suitably chosen values one can ensure linear independence between any set of columns. The K CSI RS can now be shared by a very large set of APs, even larger than in the square design that was exemplified in Figure 4. A symbol-specific amplitude might have been defined in a similar manner as the phase ϕ_v, but due to the exponent in equation (1), the power variation might then easily get large and for channel estimation a constant power per reference signal is preferable. As illustrated in Figure 5, the matrix $\mathbf{C}_V$ allocates to each AP i a specific sequence of phase values

- i.e. an AP specific code - of length K. Each element of the code sequences is transmitted from each AP on predefined orthogonal REs. The positions of the utilized resource elements could be for example as defined since Release 10 for 3GPP LTE systems, which provides 40 REs for CSI RSs per physical resource block (PRB) bandwidth. LTE uses more or less a one to one mapping of CSI RSs/REs to APs, while here in contrast AP number i is identified by its sequence $\mathbf{C}_V(1:K, i, \phi_v)$ running over the full code length K, i.e. all REs. A particular UE, UE_j, receives the Rx vector $\mathbf{y}^j$ containing per RE the sum power of all APs multiplied by the channel vector $\mathbf{h}^j \in \mathcal{C}^K$ together with some additive white Gaussian noise (AWGN) vector

$$\mathbf{n} \sim \mathcal{CN}(0, \frac{\sigma^2}{2}\mathbf{I})$$

as described in equation (2),

$$\mathbf{y}^j = \mathbf{C}_V \mathbf{h}^j + \mathbf{n}. \quad (2)$$

Note, so far a flat radio channel per PRB is being assumed, i.e. all REs of a PRB from a certain AP i to an UE j see exactly the same radio channel h_i^j.

Equation (2) can be rewritten as done in equation (3) below, where the vector $\mathbf{i}_{rel} \in \mathcal{I}$ contains all indices to the relevant CCs being received at UE j within the power window defined by the strongest CC with Rx power P^j_{max} and P^j_{TH}, where P^j_{TH} is TH dB below the maximum CC power P^j_{max}. The Rx-signal $\mathbf{y}^j$ is a combination of the desired first term carrying all relevant CCs, a second interference term due to the irrelevant CCs $\overline{\mathbf{i}_{rel}}$ falling below P^j_{TH} - i.e. the complement set of $\mathbf{i}_{rel}$ - and the AWGN noise. For a sparse channel matrix, the interference term will be small due to the low Rx-power of the $\overline{\mathbf{i}_{rel}}$ CCs, which has to be ensured by a proper overall system design.

$$\begin{aligned} \mathbf{y}^j &= \mathbf{C}_V(1..K, \mathbf{i}_{rel})\mathbf{h}^j(\mathbf{i}_{rel}) + \underbrace{\mathbf{C}_V(1..K, \overline{\mathbf{i}_{rel}})\mathbf{h}^j(\overline{\mathbf{i}_{rel}})}_{\text{inter AP interference}} + \mathbf{n} \\ \mathbf{i}_{rel} &= \underset{i \in I}{arg}(\|\mathbf{h}^j(i)\|_2^2 \geq P^j_{TH}); \\ P^j_{TH} &= P^j_{max} - TH. \end{aligned} \quad (3)$$

Assuming that the noise and interference terms are sufficiently small, it is possible to reconstruct and estimate the relevant CCs $\mathbf{h}^j(\mathbf{i}_{rel})$ by a Moore Penrose Pseudo inverse of the matrix $\mathbf{C}_V(1..K, \mathbf{i}_{rel})$ according to equation (4). The estimation error $\mathbf{E}$ is then simply the difference between the estimated and the real radio channel for the relevant CCs, as defined in equation (5) below. It depends on the interference term in equation (3), the noise $\mathbf{n}$ as well as the rank and condition of the coding matrix $\mathbf{C}_V$.

$$\hat{\mathbf{h}}^j(\mathbf{i}_{rel}) = pinv[\mathbf{C}_V(1..K, \mathbf{i}_{rel})]\, \mathbf{y}^j; \quad (4)$$

$$\mathbf{E} = \hat{\mathbf{h}}^j(\mathbf{i}_{rel}) - \mathbf{h}^j(\mathbf{i}_{rel}). \quad (5)$$

IV. High Level Performance Evaluation

Different aspects of the coded CSI concept as described above have been evaluated for a single cooperation area of the IMF-A framework as explained in section II. The cooperation area consists of three equidistant sites, each with three cells and eight fixed beams per cell, i.e. overall 72 beams or APs. As can be seen from the exemplary Figure 6 for a threshold TH of 20dB with respect to the maximum power of the strongest CC, there will only be a limited number of relevant CCs. Most CCs have small to very small Rx-power.

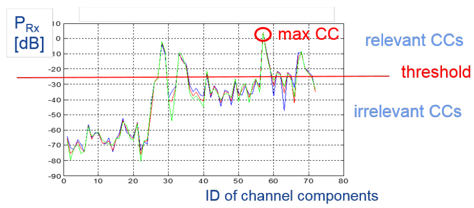

Fig. 6: Typical relative CC Rx power levels for an UE in a single cooperation area with 72 CCs and no AWGN

The estimation of the CSI for the relevant CCs $\hat{\mathbf{h}}^j(\mathbf{i}_{rel})$ involves according to equation (4) a Moore Penrose matrix inversion so that the rank and condition of $\mathbf{C}_V$ becomes important for the achievable estimation accuracy. In Figure 7, the achievable normalized mean square error (NMSE) of the error $\mathbf{E}_i$ - being the expectation of component i of $\|\mathbf{E}\|_2$ normalized to the power of $\|\mathbf{h}\|_2$ - is illustrated in the noise-free case for an example of 11 relevant out of overall 72 CCs for different ϕ_v values as parameter. As a first observation one can conclude that the NMSE varies significantly for different ϕ_v values and for different CCs.

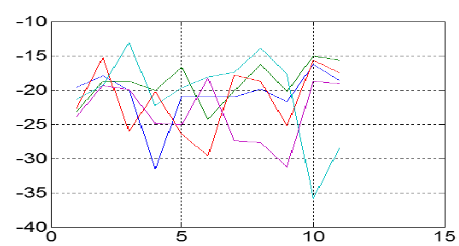

Fig. 7: NMSE [dB] for different ϕ_V values, when estimating 11 relevant channel components out of a total of 72 CCs.

A second observation can be found in Figure 8 where on the right-hand side, the condition number of $\mathbf{C}_V(1..K, \mathbf{i}_{rel})$ has been calculated for different subsets and for increasing number of relevant CCs, i.e. for different cardinality κ of the set $\{\mathbf{i}_{rel}\}$. The condition suffers over proportionally as the number of relevant CCs κ approaches the maximum value K, which in this example is K=18. As a conclusion, for a high estimation quality one should ensure high $code\,diversity$, being defined as DIV = K-κ and the minimum DIV should be for example 1 or 2. Note, the term $code\,diversity$ is motivated by the similarity to e.g. antenna or spatial diversity as known from MIMO precoding, where increasing diversity orders lead to steeper bit error rate (BER) slopes due to improved condition of the according channel matrices.

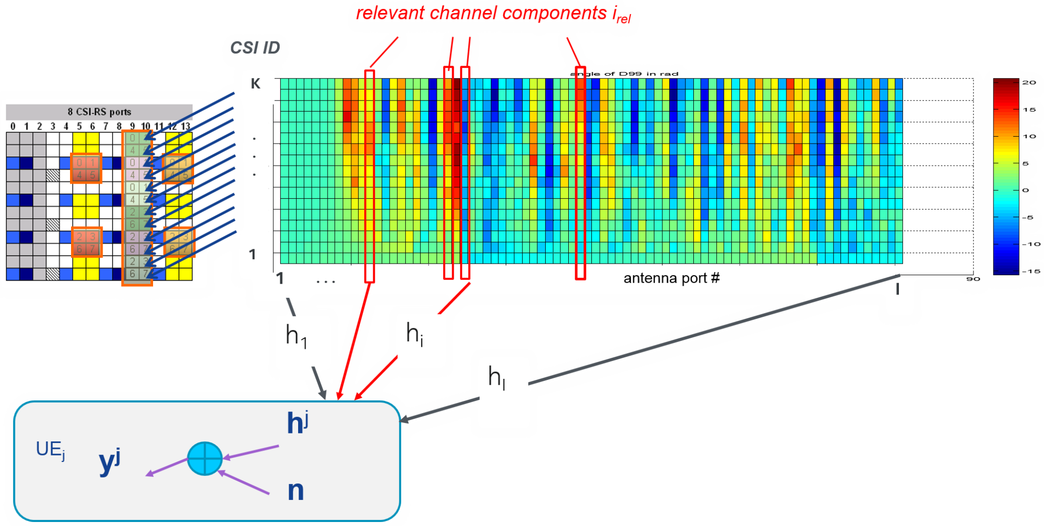

Fig. 5: Vandermonde like allocation of phase values (color coded according to the legend) to CSI RSs and APs with the coding matrix $\mathbf{C}_V$

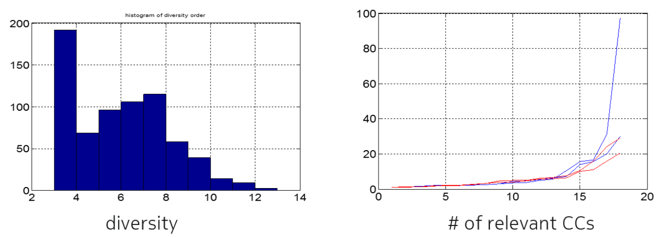

Fig. 8: histogram of code diversity (left) and of condition number of $\mathbf{C}_V$ in [dB] for different sets of relevant CCs

Figure 8 left shows a typical histogram of the code diversity order for the here investigated CoMP scenario with 72 CCs for different UEs, where the minimum DIV is two, but is quite often significantly larger than that.

From equation (3), it is obvious that the AWGN $\mathbf{n}$ will disturb $\mathbf{y}^j$ and will accordingly affect the CSI estimation quality, but there is a further more challenging aspect as can be concluded from Figure 5. So far we assumed a perfectly flat radio channel per PRB, but in reality the channel will typically be frequency selective and time varying so that there are small channel variations over the orthogonal CSI RS REs. For code division multiple access (CDMA), this leads to the well known code crosstalk issue. For coded CSI, performance will be degraded as shown below, but in a slightly different manner as the reconstruction for the channel estimation of $\hat{\mathbf{h}}^j(\mathbf{i}_{rel})$ involves a Moore-Penrose Pseudo inversion.

Equation (6) extends the channel vector $\mathbf{h}^j$ from above to a matrix $\mathbf{H}^j$ which is composed of the error free base matrix $\mathbf{H}_0^j$ and the complex delta matrix $\Delta\mathbf{H}^j \in \mathcal{C}^{KxI}$ containing the channel deviations over the resource elements k of CC i, compared to the baseline CC $\mathbf{h}_0^j(i)$.

$$
\begin{aligned}
\mathbf{h}^{j[I\times 1]} \longrightarrow \mathbf{H}^{j[K\times I]} &= \mathbf{H}_0^j + \Delta\mathbf{H}^j; \\
\mathbf{H}_0^j(1..K,i) &= \mathbf{h}^j(i)\mathbf{1}^{\mathbf{Kx1}}; \\
\Delta\mathbf{H}^j(k,i) &= \mathbf{H}^j(k,i) - \mathbf{h}_0^j(i); \\
& \quad k = 1\cdots K; \quad i = 1\cdots I.
\end{aligned} \tag{6}
$$

Here $\mathbf{1}^{\mathbf{Kx1}}$ is the all one column vector with length K. $\mathbf{H}_0^j \in \mathcal{C}^{K\times I}$ is the ideal channel matrix containing exactly the same channel value $\mathbf{h}_0^j(i)$ on all K resource elements for the i-th CC, while $\mathbf{H}^j \in \mathcal{C}^{K\times I}$ includes the non ideal time and frequency variations over K. With $\Delta\mathbf{H}^j$ one can calculate an additional error term for the received signal $\mathbf{y}_{XT}^j$ including the inter code crosstalk according to equation (7) below. It is obtained by a Hadamard or Schur product with the code matrix $\mathbf{C}_V$. For estimation of the new error vector $\mathbf{E}_{XT}$ one has to exchange $\mathbf{y}^j$ in equation (4) by $\mathbf{y}_{XT}^j$ by equation (7) below. Also here, a sufficiently high condition number of $\mathbf{C}_V(1..K, \mathbf{i}_{rel})$ will be mandatory for maintaining a high estimation quality.

$$
\mathbf{y}_{XT}^j = \mathbf{C}_V\mathbf{h}^j + \mathbf{n} + (\mathbf{C}_V \circ \Delta\mathbf{H}^j)\mathbf{1}^{\mathbf{Ix1}} \tag{7}
$$

A. Proposed Enhancements

The observations in the previous sub chapter leads naturally to some useful enhancements of the baseline *coded CSI* concept. Firstly, one should ensure by a proper system design - for example by making use from strong massive MIMO beamforming gains - that the number of relevant CCs is sufficiently small. In addition K - as the number of REs used for CSI RSs - has to be chosen with respect to the expected maximum number of relevant CCs κ so that the code diversity order DIV is always large enough. Secondly, it is proposed to do the channel estimation according to equation (4) multiple times, but for different values of ϕ_v. This can be achieved either by a re-estimation of the same relevant CCs in time or by providing two or more orthogonal sets of REs per PRB bandwidth for the CSI RSs. The accordingly multiple times higher overhead will be rewarded - after averaging or selection of the best estimation with lowest NMSE - by a significantly improved CSI estimation quality.

Another option, which might be combined with the previous solution, is to use multiple code sets with relatively small K values like for example K=9 so that the REs for these 9 CSI RSs can be allocated close to each other, e.g. in fields of REs of size 3 x 3. That minimizes the code crosstalk due to a relatively low channel variation for adjacent REs. The number of code sets depends on the number of relevant CCs, the intended code diversity order and the overall number I of CCs. For the same assumptions, a single longer code allocation over K REs will be more efficient, but lower code crosstalk as well as lower complexity might be in favor of this solution.

As further improvement one might consider iterative zero-forcing (ZF) for the REs carrying CSI RSs. Based on a first estimation the crosstalk is being reduced by an appropriate ZF filtering, leading to an improved CSI estimation. Alternatively a minimum mean square error (MMSE) filter might help to limit the noise rise.

B. Simulation Results

The coded CSI concept as described above has been simulated for the IMF-A framework with K=18 CSI RS and I=72 APs for various parameter settings. For the generation of the CDFs of the channel estimation NMSE in Figure 9, the main parameters are an AWGN power per CSI RE of -20 dB, a crosstalk error between REs of ≤ -25 dB and a threshold for selection of relevant CCs $\mathbf{i}_{rel}$ of TH = 23 dB, which led to an average code diversity order of DIV = 6.3.

The simulation is done for a single estimation (blue curve) and as average over two re-estimations with two different phase values ϕ_v defining the code matrix $\mathbf{C}_V$ as being proposed in the previous sub-chapter (magenta curve).

In addition, for comparison with conventional CSI estimation techniques, the black curve has been simulated using the same AWGN with power -20 dB per RE under the assumption of one orthogonal RE per CC. The same total RS power per AP has been used as in the case with coded CSI RS. The black curve follows the typical Rayleigh fading channel power to noise power statistic, while the blue and red curves for the novel coded CSI concept have slightly different statistics due to the different underlying error processes as derived in equation (4). Nonetheless, even for a single estimation, the NMSE is close to or even better as compared to the conventional scheme that needs to use 72 orthogonal CSI RS positions. In case of averaging over two re-estimations, the NMSE is reduced further by about 6 to 8 dB, which is about 3 to 5 dB better than the otherwise achievable 3 dB gain that would result from a conventional double estimation of the same CC.

For the averaged re-estimation solution, the average NMSE is close to -26 dB, i.e. in this example the coded CSI RS concept improved the estimation accuracy despite the very low overall overhead for CSI RSs.

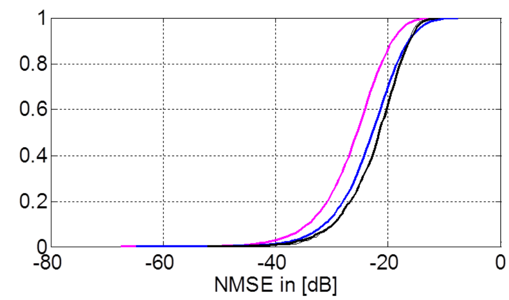

Fig. 9: CDF of NMSE for conventional (black), and coded CSI with (magenta) and without (blue) reestimation

C. Overhead

Finally we make a short analysis of the required overhead for CSI RSs one might need for a future 5G system. In Figure 10 for a threshold of TH = 20 dB, the maximum number of relevant CC is 40 and on average it is about 20. These results have been achieved for UEs with single omni directional antennas. Including UE sided spatial filters based on multiple UE antennas - potentially in combination with advanced virtual beamforming techniques as being described in [11] - the number of relevant CCs $\mathbf{i}_{rel}$ is expected to be in the range of 10 to typically less than 20.

Assuming K=20 resource elements for CSI RSs would then lead on average to a high code diversity order DIV of about 10. In case of re-estimation as described in the previous sub chapter, one has to double the number of REs to K = 2 x 20 = 40. The rate with which CSI RS are transmitted can be configured by the eNB and typical values are every five or ten ms, i.e. every fifth or tenth transmission time interval (TTI). In LTE a PRB has overall 168 REs so that the overhead for this special case would be 40/168/5 = 4.7 percent at a CSI RS repetition interval of 5 ms. The 5G frame structure is not known yet, but the relative overhead should be of a similar order.

A typical IFM-A framework includes massive MIMO, with eight horizontal times two vertical beams times two polarizations per cell for a cooperation area of nine cells. This would for a conventional estimation with one orthogonal CSI RS per CC result in an overhead of 34 percent. In case of integration of small cells, as proposed e.g. in [6], the overhead will increase further for the conventional solution, despite a relatively poor estimation quality. In contrast, the overhead of the coded CSI concept scales with the number of relevant CCs, which is expected to increase only moderately with the number of small cells due to the strong shadowing of below rooftop radio stations.

V. Conclusion

The coded CSI concept inherently exploits the spatial channel structure, which can be sparse for a suitably designed framework that includes CoMP as well as massive MIMO. As main benefit, the number of resource elements required for channel estimation using the coded CSI concept increases linearly with respect to the relevant - instead of to the overall - CCs. The concept allows us to integrate massive MIMO

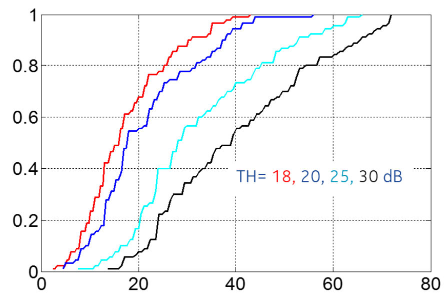

Fig. 10: CDF of number of relevant CCs for different threshold values TH

into FDD systems with a very low pilot overhead of about 5 percent. In our simulation evaluation, it has provided channel estimation with very low average NMSE of about -25 dB at 20 dB SNR for the given scenario assumptions. It provides a practical solution that solves the pilot contamination issue for TDD as well as FDD systems.

Further research is needed to verify these results for different scenarios and to find most suitable setups e.g. with respect to minimum code crosstalk, maximum code diversity or minimum overall number of resource elements for CSI RS. Further optimizations might be possible for the design of the code matrix $\mathbf{C}_V$. Comparison to solutions emerging from the latest research in compressed sensing is also of theoretical interest.

Acknowledgment

This work has been performed in the framework of the Horizon 2020 project FANTASTIC-5G (ICT-671660) receiving funds from the European Union. The authors would like to acknowledge the contributions of their colleagues in the project, although the views expressed in this contribution are those of the authors and do not necessarily represent the project.

In addition the research leading to these results has received funding from the European Commission's seventh framework programme FP7-ICT-2009 under grant agreement n° 247223 also referred to as METIS.

References

[1] E. G. Larsson, F. Tufvesson, O. Edfors, and T. L. Marzetta, "Massive MIMO for next generation wireless systems," *IEEE Communications Magazine*, vol. 52, no. 2, pp. 186–195, 2013.

[2] G. Fodor and M. Telek, "On the pilot-data trade-off in single input multiple output systems," in *Proceedings of European Wireless*, 2014.

[3] L. L. et al, "An overview of massive mimo: Benefits and challenges," *IEEE J. Selected Topics Signal Proc*, vol. 18, no. 5, 2014.

[4] Z. Jiang, A. F. Molisch, G. Caire, and Z. Niuand, "On the achievable rates of fdd massive mimo systems with spatial channel correlation," *ICCC 2014 Symposium on Selected Topics in Communications*, pp. 276–280, 2014.

[5] V. Jungnickel, K. Manolakis, W. Zirwas, B. Panzner, V. Braun, M. Lossow, M. Sternad, R. Apelfröjd, and T. Svensson, "The role of small cells, coordinated multipoint and massive mimo in 5g," *IEEE Communications Magazine, Special Issue on 5G, May 2014, pp. 44-51*, 2014.

[6] W. Zirwas, "Opportunistic comp for 5g massive mimo multilayer networks," in *Proceedings of International ITG Workshop on Smart Antennas (WSA 2015)*, 2015.

[7] ARTIST4G, "Artist deliverable 1.4 - interference avoidance techniques and system design," project report, http://ict-artist4g.eu/, Tech. Rep., June 2012.

[8] L. Thiele, M. Kurras, S. Jaekel, and W. Zirwas, "Interference floor shaping for liquid coverage zones in coordinated 5g networks," in *Proceedings of Asilomar 2015*, 2015.

[9] S. Jaekel, L. Raschkowski, K. Borner, and L. Thiele, "Quadriga: A 3-d channel model with time evolution for enabling virtual field trials," *IEEE Trans. on Antennas and Propagations*, vol. 62, no. 6, pp. 3242–3256, 2014.

[10] M. B. Amin, W. Zirwas, and M. Haardt, "Advanced channel prediction concepts for 5g radio systems," in *Proceedings of International Symposium on Wireless Communication Systems (ISWCS2015)*, 2015.

[11] W. Zirwas and M. Haardt, "Channel prediction for b4g radio systems," in *Proceedings of 77th IEEE Vehicular Technology Conf. (VTC 2013 Spring)*, 2013.