Coded CSI Reference Signals for 5G - exploiting sparcity of 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
University of Uppsalaät Uppsala, Sweden
Email: Mikael.Sternad@signal.uu.se

***Abstract*—Future 5G systems are expected to provide by factors higher performance partly unleashed by massive MIMO as well as tight cooperation like joint transmission CoMP. For the valuable below 6 GHz RF-frequency bands frequency as well as time division duplex (FDD/TDD) have to be supported. Large cooperation areas over several cells together with massive MIMO are therefore a real challenge with respect to estimation of the channel state information (CSI) per antenna element and the according uplink reporting in case of FDD. Pilot contamination has been extensively discussed in the literature 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. Analysis of channel statistics for urban macro scenarios applying massive MIMO and strong UE beamforming reveal a sparse nature of the typical channel matrices. We propose a coded allocation of CSI RSs, inherently exploiting this sparse nature and allowing accurate CSI estimation of UE individual subsets of relevant channel components despite a very low reference signal overhead of 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 up to high energy efficiency. From the beginning pilot contamination i.e. the mutual interference between multiple reference signals - has been identified as one of the main challenges eventually limiting massive MIMO performance [1].

To allow the estimation of the massive number of antenna elements often time division duplex (TDD) including channel reciprocity is being assumed. UEs will 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 use these CSI information from all UEs for a proper MU MIMO precoding.

High spectral or energy efficiency can be achieved by strong multi stream or multi user MIMO (MU MIMO) transmission so that multiple UEs have to transmit their reference signals simultaneously. Mutual crosstalk can be avoided by use of orthogonal CSI reference signals, but in case of massive number of streams the according overhead will become large. In cellular networks the situation is even worse due to the unavoidable inter cell interference demanding for orthogonal resources for SRS over multiple cells, i.e. some form of frequency reuse. Many proposals exist 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 get an overview about currently discussed options the interested reader is referred to [3]. Here the focus will be on a slightly different setup, i.e. 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. For large massive MIMO arrays the overhead for orthogonal reference signals would be 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 and any performance gains in these bands will be appreciated by mobile network operators (MNO). But, below 6 GHz many of the RF bands are paired FDD bands, motivated by higher coverage compared to TDD.

FDD based constantly transmitted DL CSI RSs have 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, which improves CSI estimation accuracy and even more is a prerequisite for a proper channel prediction. Note, channel prediction is seen as one of the main enablers for future 5G system relying on joint transmission cooperative multi point (JT CoMP) for interference mitigation [4].

Despite the mentioned benefits special care has to be taken with respect to the overall overhead for CSI RSs and 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 eight virtually precoded beams per cell [5]. Even for this GoB or DFT precoder concept the overall number of effective channel components becomes large, as enlarged cooperation areas might comprise up to nine cells a eight beams each or equivalently 72 overall beams. To avoid inter cell interference - or in other words pilot contamination - one might require hundreds of orthogonal CSI reference signals.

Here we propose to use so called coded CSI reference signals, which allow UEs to estimate a limited set of for example 2 N = 20 relevant out of potentially N^2 = 100 potential channel components with e.g. about 2 times N = 20 resource elements (RE) for CSI RSs. This concept exploits the inherent sparseness of massive MIMO channel matrices. According to theory it is often claimed that asymptotically there will be only a single relevant interference free channel component [1]. Due to the fixed GoB and reflections and diffractions in real world non line of sight (NLOS) scenarios even with massive MIMO the UEs receive multiple relevant channel components. But, for large number of antenna elements the half power beam width of the beams will be small (few degree) and together with UE Rx beamformers the channel matrix will be sparse, i.e. each UE will receive only a limited set of relevant channel components. Channel components are defined between a certain Tx-beam to the UE Rx-beam and relevant CCs are those being within a certain power window with respect to the strongest channel component of that UE.

For coded CSI reference signals each effective channel will transmit on a set of pre-allocated CSI resources elements individually precoded - or coded - CSI reference signals. UEs reconstruct their set of relevant channel components by a properly calculated Moore Penrose Pseudo inverse of the precoding matrix. The important point is that each UE can reconstruct a different individual set of relevant channel components, as long as the number of CCs for this UE is sparse, i.e. less than 2 N. For that purpose a full rank and good condition of any potential sub matrix has to be ensured, which led to a Vandermonde like coding matrix.

The sparse radio channel conditions are discussed in Section II. Section III explains the coded CSI concept being analyzed in detailed in Section IV, while Section V provides the conclusions.

II. Sparse 5G Radio Channels

Massive MIMO - potentially in combination with JT CoMP covering many sites or cells - leads to 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 - reported to the eNB. For a suitable system design - as being sketched in the next 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, which 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 [6].

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 [7]. Due to this decoupling the performance for a single cooperation area provides already a quite accurate estimate for the performance of a full cellular network. Each cell - i.e. each sector of the three macro sites - is 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 effective beams (GoB), i.e. downscale the 16 x 32 = 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 144 APs. For the typical cross polarized antenna elements the number of APs will increase further to over 288. For a 3GPP LTE system it would mean - under the assumption of a 5 ms periodicity - an overhead for CSI RSs of $>$ 35 percent. The goal for the coded CSI RSs is to reduce this overhead to more reasonable numbers like five percent or less without sacrificing - or even improving - the achievable accuracy for the channel estimation.

Relevant Channel Components

An important prerequisite for the coded CSI concept - as being explained further below - is to assume 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. Theoretic investigation predict asymptotically even single CCs per UE overcoming any inter cell interference [xxx]. For real world scenarios limited beamforming gains, reflections in non line of sight (NLOS) scenarios, inter beam correlations, etc. lead to a more complex channel structure, especially under assumption of a fixed GoB solution. Raytracing as well as system level simulations based on the Quadriga channel model [8] 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. 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 with a massive MIMO antenna array forming a regular GoB consisting of ten beams in this case with a half power beam width (HPBW) of only 2 degree. Vertically the beam patterns are that 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 this strong coverage within the center of the cooperation area, where almost everywhere a Rx power of at least -80 dBm is being achieved. Note, the figure provides only 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 as a 3D plot, which is for better visibility color coded according to the legend. Many 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, which 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 some 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.

Similarly for 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 remains 16 x 9 = 151 transmitted effective CCs (= Tx beams) from which the UEs receive on average only about 20 relevant CCs or 20/151 $\approx$ 15 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 [9].

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 is 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 is for 5G in the range of about 20 to 100 MHz.

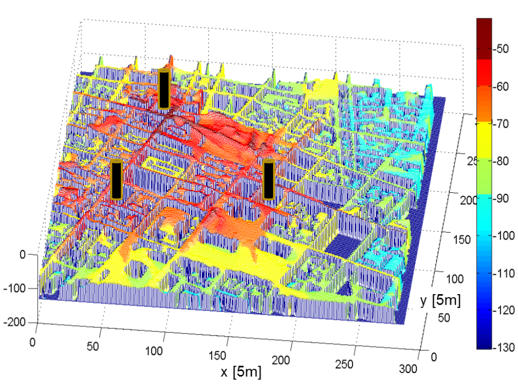

Fig. 1: Coverage of single CA in Munich Schwabing

III. Coded CSI Reference Signal Concept

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.

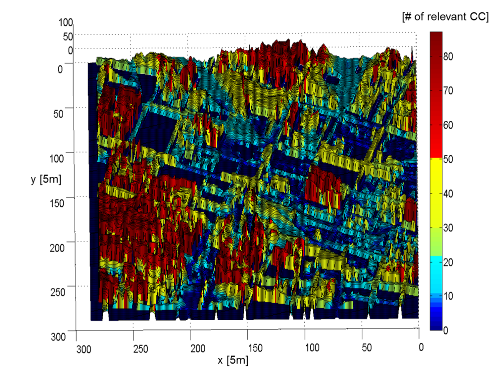

Fig. 2: Number of relevant CCs over location

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 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.

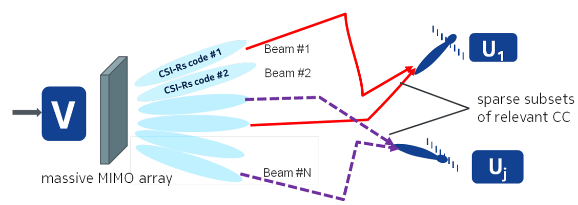

Fig. 3: UE individual subsets of relevant CCs

For illustration of the basic idea of coded CSI RSs lets consider the most simple case of a system having nine APs and a maximum of six UE individual different sub-sets of relevant CCs. In Figure 4 left 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 identification of errored elements. For example AP 5 can be identified by the combination of CSI 2 and 5 and similarly AP 6 by CSI 3 and 5, as visible from the circles in the figure.

Each magenta colored element indicates that the according AP (x-axis) transmits the according CSI RS (y-axis), i.e. for example AP 1 transmits CSI RS 1 and 4. Note, so far all active CSI RSs are transmitting the same signal, i.e. a '1'.

UEs receive the $\mathbf{C}^{Kx1}$ vector $\mathbf{y} = \mathbf{C}^{33\_base} * \mathbf{h}$, where $\mathbf{h}$ is the $\mathbf{C}^{Ix1}$ vector comprising all CCs. $\mathbf{y}_k$ is the sum over all CCs for which CSI RS k is active (transmits a '1'). For 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. For that reason 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 gives a linear equation system of three equations for two variables, which can be easily solved. From another perspective one can note that each CC is 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 doing the More Penrose matrix inversion of $C_{rel} = C^{33\_base}_{1..K,\mathbf{i}_{rel}}$, i.e. $\hat{\mathbf{h}}_{\mathbf{i}_{rel}} = pinv(C_{rel}) * \mathbf{y}$, with $\mathbf{i}_{rel}$ being the indices to the relevant CCs or APs. In our case $\mathbf{i}_{rel}$ = {5,6}.

To generate C_{rel} the UEs have to identify i_{rel} based on e.g. RSRP measurements as known from 3GPP LTE. In addition the CSI reporting will be done only for the relevant CCs i_{rel} so that the eNBs have to be informed by the UEs about the actual relevant CCs. As pathloss is a large scale parameter (LSP) the list of relevant CCs will have to be updated only quite seldom, e.g. every 50 0ms with accordingly low extra overhead.

In principle the goal would be to allow channel estimation for any subset of relevant CCs with cardinality $|\mathbf{h}_{\mathbf{i}_{rel}}|$ = 6 and the motivation is that we have six linear equations per CSI RS. Unfortunately $\mathbf{C}^{33\_base}$ as it is does not fulfill this target as can be easily concluded from Figure 4 middle. By selection of the subset of rel CCs AP1 to 6 the CSI RS 6 will be zero (does not transmit) for all 6 APs and the rank of the according coding matrix C_{rel} is therefore at maximum 5. The coding matrix C_{33} as depicted 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.

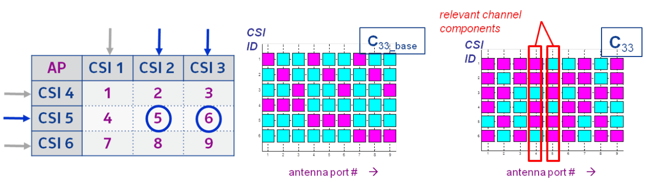

Fig. 4: Basic concept of coded CSI reference signals; magenta/blue stands for '1'/'0' code elements

Vandermonde like coding matrix

For larger coding matrices as needed for the target scenario with some tenth 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. To ensure inherently full rank for any subset

$$\mathbf{i}_{rel}$$

with cardinality

$$\mathcal{K} \ll I$$

out of overall I APs it is proposed to use a Vandermonde like coding matrix $\mathbf{C}_V \in \mathcal{C}^{K \times I}$ as defined in equation (1) with $k \in 1...K$ being 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 single phase parameter ϕ_v defines the whole matrix $\mathbf{C}_V$ and for suitably chosen values one can ensure linear independency between any set of columns. An 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 easily get large and for channel estimation 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 as defined for example since Release 10 for 3GPP LTE systems, which provides 40 REs for CSI RSs per physical resource block (PRB). 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 certain UE UE_j receives the Rx vector $\mathbf{y}^{\mathbf{j}}$ 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). 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.

$$\mathbf{y}^j \quad = \quad \mathbf{C}_V \mathbf{h}^j \; + \; \mathbf{n} \tag{2}$$

Equation 2 can be rewritten as done in equation (3), 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}. 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 wanted 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. complementing $\mathbf{i}_{rel}$ - 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.

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

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

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

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

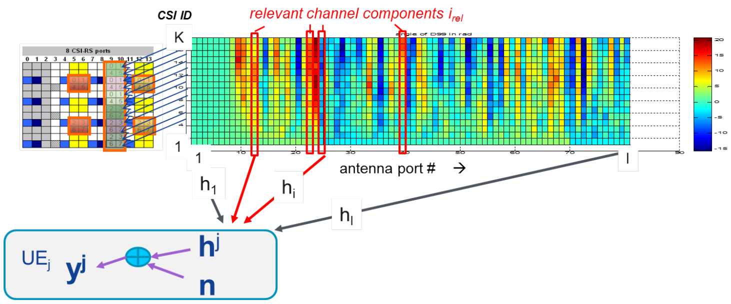

Fig. 5: Vandermonde like allocation of phase values (color coded acc. to legend) to CSI RSs and APs with the coding matrix $\mathbf{C}_V$

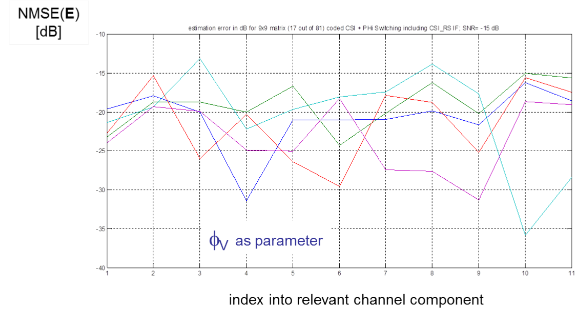

Fig. 7: NMSE [dB] for different ϕ_V values

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 comprising three sites 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 be only a limited number of relevant CCs. Most CCs have small to very small Rx-power.

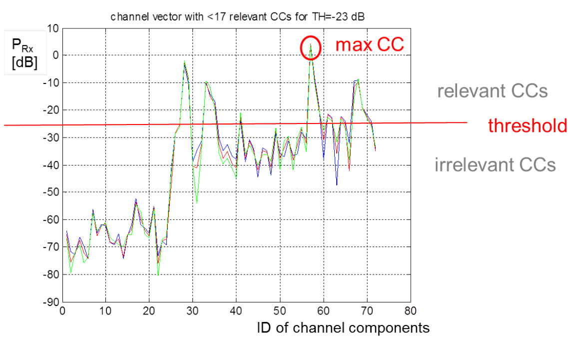

Fig. 6: Typical Rx power levels for a single cooperation area with 72 CCs

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 5 the achievable normalized mean square error (NMSE) of the error $\mathbf{E}$ - being the expectation of $\mathbf{E}$ normalized to the power of $\|\mathbf{h}\|_2$ - is illustrated for an example of 11 relevant out of overall 81 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.

A second observation can be found in figure 8 where on the right side the condition 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 κ approaches K, which is in this example 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.

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.

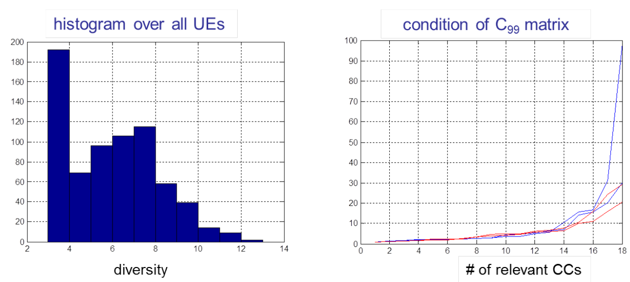

Fig. 8: histogram of code diversity (left) and of condition of $\mathbf{C}_V$ for different sets of relevant CCs

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 be typically 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 well, even so 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 the 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 $\mathbf{h}_0^j(i)$.

$$
\begin{aligned}
\mathbf{h}^{j[Ix1]} \longrightarrow \mathbf{H}^{j[KxI]} &= \mathbf{H}_0^j + \Delta\mathbf{H}^j; \\
\mathbf{H}_0^j(1..K, i) &= \mathbf{h}^j(i)\mathbf{1}^{\mathbf{Kx1}}; \\
\Delta\mathbf{H}^j(k, i) &= \mathbf{h}^j(k, i) - \mathbf{h}_0^j(i);
\end{aligned} \tag{6}
$$

Here $\mathbf{1}^{\mathbf{Kx1}}$ is the all one vector with length K. With $\Delta\mathbf{H}^j$ one can calculate the additional error term for the received signal $\mathbf{y}_{XT}^j$ including the inter code crosstalk according to equation (7) which is achieved 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 in equation (4) $\mathbf{y}^j$ by $\mathbf{y}_{XT}^j$ and again a low condition 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 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 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 just from the beginning per PRB two or more orthogonal sets of REs 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 this 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 of CCs I. For 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 ZF for the REs carrying CSI RSs, i.e. based on a first estimation the crosstalk is being reduced by an according ZF leading to an improved CSI estimation and so on. 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 per CSI RE of $\mathbf{n} = -20dB$, a crosstalk error between REs of $\leq -25dB$ and a threshold for selection of relevant CCs $\mathbf{i}_{rel}$ of TH = 23dB, 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 of $\mathbf{n} = -20dB$ per RE under the assumption of one orthogonal RE per CC. The black curve follows the typical Rayleigh 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 or even better compared to the conventional schemes. 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 for 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.

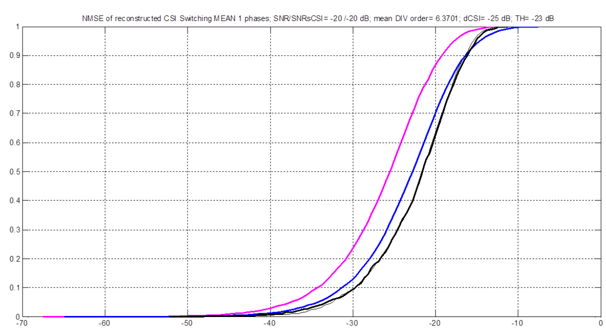

Fig. 9: CDF of NMSE for conventional (black), and coded CSI with (magenta) and without (blue) reestimation

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 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 [10] - 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 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. The 5G frame structure is not known yet, but the relative overhead should be in a similar order.

A typical IFM-A framework including massive MIMO with eight horizontal times two vertical beams times two polarizations per cell for a cooperation area of nine cells would lead for a conventional estimation with one orthogonal CSI RS per CC to an overhead of 34 percent. In case of integration of small cells as proposed e.g. in [5] 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.

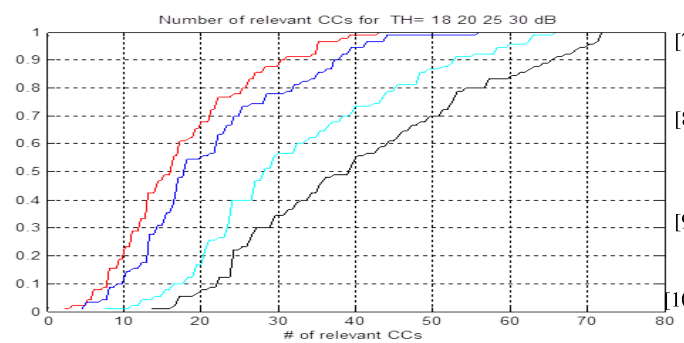

Fig. 10: CDF of number of relevant CCs for different threshold values TH

V. Conclusion

The coded CSI concept exploits inherently the spatial channel structure, which can be sparse for a suitable designed overall IMF-A framework including CoMP as well as massive MIMO. As main benefit the number of resource elements for channel estimation in case of the coded CSI concept increases linearly with respect to the relevant - instead of to the overall - CCs. The concept allows to integrate massive MIMO into FDD systems with a very low overhead of about 5 percent and very low NMSE estimation error of about -25 dB on average. It provides a practical solution, obviously solving inherently 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 $\mathbf{C}_V$. Theoretical interesting will be a comparison to solutions esteeming from latest research in compressed sensing.

Acknowledgment

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

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

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

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

[7] 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.

[8] 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.

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

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