Joint Covariance Matrix Estimation and Pilot Allocation in Massive MIMO Systems

David Neumann, Kamel Shibli, Michael Joham, and Wolfgang Utschick
Associate Institute for Signal Processing, Technische Universität München, 80290 Munich, Germany
{d.neumann,kamel.shibli,joham,utschick}@tum.de

***Abstract*—Pilot contamination is a throughput limiting factor in cellular massive MIMO systems. Previous work has shown that the impact of pilot-contamination can be reduced by exploiting structural information in form of channel covariance matrices. Additionally, significant gains can be obtained through coordinated user assignment. In this paper, we extend these approaches to a realistic scenario with imperfect knowledge of the channel and its distribution at the base station. We formulate an optimization problem for assigning users to the available pilot sequences, which at the same time takes the estimation of the covariance matrices into account and propose a suboptimal greedy algorithm for efficient implementation. Simulation results with established physical channel models demonstrate the significant performance gains of the proposed method in the case of imperfect knowledge of the covariance matrices at the base station.**

I. Introduction

Massive MIMO is a promising technology for enhancing the spectral efficiency of fifth generation cellular networks. Asymptotically, it has been shown that the array gain increases proportionally with the number of available antennas at the base station. Moreover, under certain conditions on the distributions of the channel vectors, the channels of different users are asymptotically orthogonal making it possible to use very simple signal processing techniques [1].

However, a major challenge of massive MIMO is the limited number of available pilot sequences in the training phase. To fully exploit the potential of a massive MIMO system, several users have to share the same pilot sequence leading to interference during the training phase, so called pilot-contamination. Different methods have been proposed in the literature to tackle this effect by exploiting the second order statistics of the channel [2]–[5].

Additional gains can be achieved, if users and pilots are assigned properly [2], [3] [6] [7]. These approaches require knowledge about the covariance matrices at the base station prior to resource allocation and beamformer design.

A novel approach for solving the covariance matrix estimation problem in the presence of pilot-contamination has been proposed in [8] based on a systematic allocation of pilot sequences to users in consecutive coherence intervals. We adapt this approach in our simulation for estimating the covariance matrices, and afterwards we use this information to adaptively assign users to groups based on an efficient greedy algorithm. In contrast to state-of-the-art algorithms, our approach takes into account that the covariance matrix estimation requires different pilot allocations in consecutive channel coherence intervals. Specifically, we optimize the assignment problem over several coherence intervals.

II. System Model

We consider a single cell scenario with one base station having M antennas and communicating with K single-antenna users in time division duplex transmission mode.

We assume a block fading model, i.e., the channel vectors $\boldsymbol{h}_k[t] \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{h}_k})$ between user k and the base station are constant within a coherence interval t of T channel accesses. In each coherence interval, we have an uplink training phase where T_{tr} channel accesses are used to gather observations on the channel vectors. We assume that T_{tr} is smaller than the number of users K leading to pilot contamination.

In every coherence interval, each user k is assigned one of T_{tr} orthogonal pilot sequences. Assuming a coherent and synchronized reception at the base station, we can simply use a least squares (LS) estimator and correlate the signal received during the training phase with the different pilot sequences yielding T_{tr} observations

$$\boldsymbol{y}_p[t] = \sum_{k \in \Omega_p[t]} \boldsymbol{h}_k[t] + \boldsymbol{n}_p[t] \in \mathbb{C}^M \tag{1}$$

where $\Omega_p[t]$ is the set of users sharing the pilot sequence p in coherence interval t with $p = 1, \ldots, T_{\text{tr}}$, and $\boldsymbol{n}_p \sim \mathcal{N}_{\mathbb{C}}(\mathbf{0}, \boldsymbol{C}_{\boldsymbol{n}})$ is additive white Gaussian noise.

From (1), we obtain the covariance matrix of the LS channel estimate

$$\boldsymbol{C}_{\boldsymbol{y}_p[t]} = \sum_{k \in \Omega_p[t]} \boldsymbol{C}_{\boldsymbol{h}_k} + \boldsymbol{C}_{\boldsymbol{n}} \in \mathbb{C}^{M \times M} \tag{2}$$

where we assumed that the channel vectors of the individual users are uncorrelated with each other, and uncorrelated with the noise as well.

For a more convenient notation, we introduce the pilot allocation matrix $\boldsymbol{\Pi}[t] = \{0,1\}^{K \times T_{\text{tr}}}$ with

$$[\boldsymbol{\Pi}[t]]_{k,p} = \begin{cases} 1 & \text{for } k \in \Omega_p[t] \\ 0 & \text{otherwise} \end{cases}.$$

We then collect the vectorized covariance matrices in

$$\boldsymbol{B}_{\boldsymbol{Y}[t]} = [\text{vec}(\boldsymbol{C}_{\boldsymbol{y}_1[t]}), \ldots, \text{vec}(\boldsymbol{C}_{\boldsymbol{y}_{T_{\text{tr}}}[t]})] \in \mathbb{C}^{M^2 \times T_{\text{tr}}}$$

and

$$\boldsymbol{B}_{\boldsymbol{H}} = [\text{vec}(\boldsymbol{C}_{\boldsymbol{h}_1}), \ldots, \text{vec}(\boldsymbol{C}_{\boldsymbol{h}_K})] \in \mathbb{C}^{M^2 \times K}.$$

which allows us to rewrite (2) as

$$\boldsymbol{B}_{\boldsymbol{Y}[t]} = \boldsymbol{B}_{\boldsymbol{H}}\boldsymbol{\Pi}[t] + \mathrm{vec}(\boldsymbol{C}_{\boldsymbol{n}})\mathbf{1}^{\mathrm{T}} \tag{3}$$

where $\mathbf{1}$ is the all-ones vector.

To minimize the channel estimation error and reduce the impact of pilot-contamination, the MMSE estimator is used [cf. [3]] which gives the enhanced channel estimate of user k employing the pilot sequence p in coherence interval t

$$\hat{\boldsymbol{h}}_k[t] = \boldsymbol{C}_{\boldsymbol{h}_k}\boldsymbol{C}_{\boldsymbol{y}_p[t]}^{-1}\boldsymbol{y}_p[t]. \tag{4}$$

The additional gain of the MMSE estimator compared to the conventional LSE is substantial in scenarios where the ratio of the number of antennas to the number of multipaths is sufficiently large such that the channel covariance matrices are rank deficient, i.e., the interference mitigation is enhanced if users sharing the same training sequence have less structure in common [3], [6].

III. Estimation of Covariance Matrices

The MMSE channel estimation presented in the previous section requires perfect knowledge of the covariance matrices $\boldsymbol{C}_{\boldsymbol{h}_k}$ and $\boldsymbol{C}_{\boldsymbol{y}_p[t]}$. In practice, both are not known beforehand at the base station and have to be estimated. Hence, the MMSE estimate is actually calculated as

$$\hat{\boldsymbol{h}}_k[t] = \hat{\boldsymbol{C}}_{\boldsymbol{h}_k}\hat{\boldsymbol{C}}_{\boldsymbol{y}_p[t]}^{-1}\boldsymbol{y}_p[t] \tag{5}$$

where $\hat{\boldsymbol{C}}_{\boldsymbol{h}_k}$ and $\hat{\boldsymbol{C}}_{\boldsymbol{y}_p[t]}$ are now the estimates of $\boldsymbol{C}_{\boldsymbol{h}_k}$ and $\boldsymbol{C}_{\boldsymbol{y}_p}$, respectively.

If we assume that the same pilot allocation $\boldsymbol{\Pi}[t] = \boldsymbol{\Pi}$ is used in each channel coherence interval, the estimation of $\boldsymbol{C}_{\boldsymbol{y}_p}$, which then does not depend on t, is straightforward and can be computed from the least squares estimates $\boldsymbol{y}_p[t]$ using rank-one updates in each channel coherence interval t

$$\hat{\boldsymbol{C}}_{\boldsymbol{y}_p} \leftarrow (1-\alpha)\hat{\boldsymbol{C}}_{\boldsymbol{y}_p} + \alpha\boldsymbol{y}_p[t]\boldsymbol{y}_p[t]^{\mathrm{H}} \tag{6}$$

with $\alpha \in (0,1)$. For an efficient and implementation of the covariance matrix estimation, we refer to the works in [9]–[11].

In a subsequent step, we can estimate the channel covariance matrices $\hat{\boldsymbol{C}}_{\boldsymbol{h}_k}$. From (3), we have

$$\boldsymbol{B}_{\boldsymbol{H}}\boldsymbol{\Pi} = \boldsymbol{B}_{\boldsymbol{Y}} - \mathrm{vec}(\boldsymbol{C}_{\boldsymbol{n}})\mathbf{1}^{\mathrm{T}}. \tag{7}$$

Even if we assume that the noise covariance matrix is known we cannot uniquely reconstruct the channel covariance matrices since $\boldsymbol{\Pi} \in \{0,1\}^{K \times T_{\mathrm{tr}}}$ is a tall matrix (by assumption, $K > T_{\mathrm{tr}}$).

To be able to estimate the channel covariance matrices, we adopt the approach proposed in [8] which requires a dynamic assignment of pilots to users in consecutive channel coherence intervals. Instead of using the same pilot allocation $\boldsymbol{\Pi}$ in each coherence interval, we use N different pilot allocations repeatedly. That is, we have different allocation matrices $\boldsymbol{\Pi}_n$, $n = 1, \ldots, N$, and in coherence interval t we use the allocation $n(t) = t \bmod N$. Thus, we obtain in total NT_{tr} different combinations of the channel covariance matrices. Collecting all least squares covariance matrices together and extending (7), we get the linear equation system

$$\boldsymbol{B}_{\boldsymbol{H}}[\boldsymbol{\Pi}_1, \ldots, \boldsymbol{\Pi}_N] = [\boldsymbol{B}_{\boldsymbol{Y}}^{(1)}, \ldots, \boldsymbol{B}_{\boldsymbol{Y}}^{(N)}] - \mathrm{vec}(\boldsymbol{C}_{\boldsymbol{n}})\mathbf{1}^{\mathrm{T}} \tag{8}$$

with

$$\boldsymbol{B}_{\boldsymbol{Y}}^{(n)} = [\mathrm{vec}(\boldsymbol{C}_{\boldsymbol{y}_1}^{(n)}, \ldots, \mathrm{vec}(\boldsymbol{C}_{\boldsymbol{y}_{T_{\mathrm{tr}}}}^{(n)})]$$

and $\boldsymbol{C}_{\boldsymbol{y}_p}^{(n)}$ is the covariance matrix of the LS estimate that results from correlation with pilot sequence p in the coherence intervals t in which $\boldsymbol{\Pi}_n$ is used. For a full-rank matrix

$$\tilde{\boldsymbol{\Pi}} = [\boldsymbol{\Pi}_1, \ldots, \boldsymbol{\Pi}_N] \in \mathbb{C}^{K \times NT_{\mathrm{tr}}}$$

with K independent rows, we can now apply the pseudo-inverse to reconstruct the user covariance matrices

$$\boldsymbol{B}_{\boldsymbol{H}} = \left([\boldsymbol{B}_{\boldsymbol{Y}}^{(1)}, \ldots, \boldsymbol{B}_{\boldsymbol{Y}}^{(N)}] - \mathrm{vec}(\boldsymbol{C}_{\boldsymbol{n}})\mathbf{1}^{\mathrm{T}}\right)\tilde{\boldsymbol{\Pi}}^{+}. \tag{9}$$

The estimation of the noise covariance matrix could be included in the estimation procedure but for simplicity we assume that the noise covariance matrix is known.

IV. User Assignment

As mentioned in Section II, previous work indicates that the performance of the MMSE channel estimation improves if users sharing the same pilot sequence have less structure in common. This insight is exploited in [3] [6] [7] by assigning pilots to users depending on the properties of their covariance matrices. Several greedy algorithms have been proposed to solve the assignment problem. However, these papers assume perfect knowledge of the covariance matrices at the base station and hence the optimal assignment is identical in each training phase. In our scenario, the covariance matrices are not known at the base station and have to be estimated. Therefore, the previously developed approaches do not apply, since the pilot allocation matrix $\boldsymbol{\Pi}$ has to be full rank (cf. (8)).

A. Network Utility Maximization Problem

For perfect channel distribution information (CDI), i.e., perfect knowledge of the covariance matrices, the same pilot allocation can be used in each coherence interval. Consequently, given a utility function $U(\boldsymbol{\Pi}, \boldsymbol{B}_{\boldsymbol{H}})$ which returns the utility achieved in one coherence interval for the pilot allocation $\boldsymbol{\Pi}$ and the channel covariance matrices $\boldsymbol{B}_{\boldsymbol{H}}$ we try to find the optimal allocation

$$\{\Omega_1^{\star}, \ldots, \Omega_{T_{\mathrm{tr}}}^{\star}\} = \underset{\Omega_1,\ldots,\Omega_{T_{\mathrm{tr}}}}{\operatorname{argmax}}\, U\left(\boldsymbol{\Pi}(\Omega_1, \ldots, \Omega_{T_{\mathrm{tr}}}), \boldsymbol{B}_{\boldsymbol{H}}\right). \tag{10}$$

Users which are in different sets Ω_p do not interfere with each other during the training phase. Thus a systematic allocation helps to avoid unfavorable interference conditions.

The extension to the case where the channel covariance matrices have to be estimated is not straightforward. Say we have a method which, for a given number of users K and available pilot sequences T_{tr}, generates the desired N allocation matrices $\boldsymbol{\Pi}_n$, $n = 1, \ldots, N$, such that accurate estimation of the channel covariance matrices is possible. How can we use this allocation matrices while still having some degrees of freedom to optimize the allocation for a given network utility function?

We define a certain number of groups $G < T_{\text{tr}}$ for which, analogously to the case with perfect CDI, we guarantee that users in different groups do not interfere during the training phase. To this end, we divide the available pilot sequences in G disjoint subsets, one subset for each of the groups. For notational simplicity, suppose that both the number of users and the number of available training sequences is divisible by G and we have $\tilde{K} = K/G$ users in each group which use $\tilde{T}_{\text{tr}} = T_{\text{tr}}/G$ of the available pilot sequences.

Now we simply generate the desired schedule of pilot allocations $\boldsymbol{\Pi}_n^g \in \{0,1\}^{\tilde{K}\times\tilde{T}_{\text{tr}}}$ with $n = 1, \ldots, N$, for each group $g = 1, \ldots, G$, and perform the channel covariance matrix estimation separately.

For example, if we have $K = 8$ users and $T_{\text{tr}} = 4$ pilot sequences, we can divide the users into $G = 2$ groups with four users each. In group 1, we have the users $\Omega_1 = \{1, 2, 3, 4\}$ which use pilot sequences 1 and 2. In group 2, we have the users $\Omega_2 = \{5, 6, 7, 8\}$ which use pilot sequences 3 and 4. The schedule of allocation matrices for $\tilde{K} = 4$ and $\tilde{T}_{\text{tr}} = 2$ is given by

$$\boldsymbol{\Pi}_1^g = \begin{bmatrix} 1 & 0 \\ 1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix} \quad \boldsymbol{\Pi}_2^g = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{bmatrix} \quad \boldsymbol{\Pi}_3^g = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \tag{11}$$

which can be verified to lead to a full-rank matrix $[\boldsymbol{\Pi}_1^g, \boldsymbol{\Pi}_2^g, \boldsymbol{\Pi}_3^g]$. The full assignment matrices are then given by

$$\boldsymbol{\Pi}_n = \begin{bmatrix} \boldsymbol{\Pi}_n^1 & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Pi}_n^2 \end{bmatrix}, n = 1, \ldots, 3$$

i.e., we get block-diagonal matrices with the allocation matrices of the sub-groups on the diagonal. Clearly, users in different groups by design never use the same pilot sequence and thus never interfere with each other during training.

With the proposed approach to design the pilot allocations $\boldsymbol{\Pi}_n$, we are able to accurately estimate the covariance matrices irrespective of how the users are divided into the groups. On the other hand, we can now optimize the assignment of the users into the groups to avoid unfavorable interference conditions. Note that we need $\tilde{T}_{\text{tr}} \geq 2$ to be able to design the pilot allocation schedule. Thus, the number of groups cannot exceed $G_{\text{max}} = T_{\text{tr}}/2$. This is in contrast to the case of perfect CDI where we basically have T_{tr} groups and only have to decide which user uses which pilot sequences.

Since the design of the pilot allocations $\boldsymbol{\Pi}_n$ is fixed given a partition of the users $\{\Omega_1, \ldots, \Omega_G\}$ into the G available groups, we can now formulate the utility maximization problem for imperfect CDI

$$\max_{\Omega_1,\ldots,\Omega_G} \sum_{n=1}^{N} U\left(\boldsymbol{\Pi}_n\left(\Omega_1, \ldots, \Omega_G\right), \hat{\boldsymbol{B}}_{\boldsymbol{H}}\right). \tag{12}$$

Note that we now use the estimate of the channel covariance matrices $\hat{\boldsymbol{B}}_{\boldsymbol{H}}$ since the actual covariance matrices are not available. We also have to combine the utility of the N different pilot allocations $\boldsymbol{\Pi}_n$ which are necessary to estimate the covariance matrices.

The necessary number of different allocations N depends on the number of groups G. From [8], we know that in general the assignment interval N, i.e., the number of different assignments $\boldsymbol{\Pi}_n$ has to follow

$$N \geq \frac{K-1}{T_{\text{tr}}-1} \tag{13}$$

to allow for unique identification of the channel covariance matrices. With the proposed grouping strategy, we only have $\tilde{K} = K/G$ users in each group which use $\tilde{T}_{\text{tr}} = T_{\text{tr}}/G$ pilots. Since we treat the covariance matrix estimation separately in each group we thus have

$$N \geq \frac{K/G-1}{T_{\text{tr}}/G-1} = \frac{K-G}{T_{\text{tr}}-G} \geq \frac{K-1}{T_{\text{tr}}-1} \tag{14}$$

which indicates that for a larger number of groups G, we need a longer coherence interval of the covariance matrices to obtain covariance matrix estimates with a similar accuracy.

B. Greedy Algorithm

In the following, we propose an algorithm for the assignment problem of K users to G groups. Similar work can be found e.g. in [3] for perfect CDI or in [12] under perfect knowledge of the channel at the transmitter. Here, we aim to find a close-to-optimal choice for the mapping of the users to groups based on their estimated covariance matrices.

Since the optimization problem in (12) is still combinatorial, we use the typical greedy approach to find a suboptimal solution. We start with empty sets Ω_g and add one user after the other, greedily optimizing the current utility. The optimal greedy assignment in each iteration can be written as

$$(g^*, k^*) = \underset{g\in\bar{\Gamma}, k\in\bar{\Theta}}{\arg\max}\, U\left(\Omega_1, \ldots, \Omega_g \cup \{k\}, \ldots, \Omega_G, \hat{\boldsymbol{B}}_{\boldsymbol{H}}\right) \tag{15}$$

where $\bar{\Gamma}$ and $\bar{\Theta}$ denote the set of still available groups and users respectively.

Optimally, we would like to put users with mutually orthogonal covariance matrices in the same group. In this case, the MMSE channel estimation can completely filter out pilot-contamination. To determine the spatial correlation between two users, we consider the performance metric of [6], i.e.,

$$\Delta_{kk'} = \frac{\operatorname{tr}\left[\hat{\boldsymbol{C}}_{\boldsymbol{h}_k}\hat{\boldsymbol{C}}_{\boldsymbol{h}_{k'}}\right]}{\left\|\hat{\boldsymbol{C}}_{\boldsymbol{h}_k}\right\|_F \left\|\hat{\boldsymbol{C}}_{\boldsymbol{h}_{k'}}\right\|_F} \tag{16}$$

with $\|.\|_F$ denoting the Frobenius norm. Based on the spatial correlation, the utility function is expressed as

$$U\left(\Omega_1, \ldots, \Omega_G\right) = -\sum_{g=1}^{G} \sum_{k,k'\in\Omega_g} \Delta_{k,k'}. \tag{17}$$

The complete greedy algorithm is given in Algorithm 1. The first step is an initialization step that fills the groups with users by maximizing their sum correlation using (16). The second step is the iterative greedy assignment of users to groups.

Algorithm 1 Greedy Algorithm

Require: $\hat{\boldsymbol{C}}_{\boldsymbol{h}_k}$ with $k \in \{1, 2, \dots, K\}$

Initial set of all users

$$\Theta \leftarrow \{1, 2, \dots, K\}$$

Initial non-empty groups

$$\bar{\Gamma} \leftarrow \{1, 2, \dots, G\}$$

Assign one initial user to each group. The users of different groups should have high correlation, that is,

$$(k_1^\star, k_2^\star) \leftarrow \underset{k_1, k_2 \in \Theta, k_1 \neq k_2}{\operatorname{argmax}} \Delta_{k_1 k_2}$$

$$\Omega_1 \leftarrow \{k_1^\star\} \quad \Omega_2 \leftarrow \{k_2^\star\} \quad \bar{\Theta} \leftarrow \bar{\Theta} \setminus \{k_1^\star, k_2^\star\}$$

for $g = 3$ to G **do**

$$k_g^\star \leftarrow \underset{k_g, \in \Theta}{\operatorname{argmax}} \sum_{i=1}^{g-1} \Delta_{k_i^\star k_g}$$

$$\Omega_g \leftarrow \{k_g^\star\} \quad \bar{\Theta} \leftarrow \bar{\Theta} \setminus \{k_g^\star\}$$

Greedily assign remaining users

while $\bar{\Theta} \neq \varnothing$ **do**

$$(g^\star, k^\star) = \underset{g \in \bar{\Gamma}, k \in \bar{\Theta}}{\operatorname{argmax}} U\left(\Omega_1, \dots, \Omega_g \cup \{k\}, \dots, \Omega_G, \hat{\boldsymbol{B}}_{\boldsymbol{H}}\right)$$

$$\Omega_{g^\star} \leftarrow \Omega_{g^\star} \cup k^\star \quad \bar{\Theta} \leftarrow \bar{\Theta} \setminus \{k^\star\}$$

if $|\Omega_{g^\star}| = K/G$ **then**

$$\bar{\Gamma} \leftarrow \bar{\Gamma} \setminus \{g^\star\}$$

V. Results

We consider a single-cell configuration with a uniform linear array (ULA) at the base station. The channel coherence interval is assumed to be $T = 96$ channel accesses of which T_{tr} slots are used for training. We focus on the steady state case of the covariance matrix estimation, i.e., the covariance matrices are constant over the simulated time-span and thus the adaptivity constant α determines the accuracy of the estimation. The covariance matrices are generated based on the urban macro cell model in [13]. Due to the large ULA at the base station, the covariance matrices have Toeplitz structure and are approximately diagonalized by the DFT matrix. To increase the performance of the numerical operations and the accuracy of the covariance matrix estimation, we assume that the DFT matrix perfectly diagonalizes the covariance matrices (cf. [8]).

For our joint covariance matrix estimation and pilot allocation approach, the K users are divided into G groups. For simplicity, we suppose that each group serves K/G users and uses T_{tr}/G training slots. Finally, the performance of our approach for covariance matrix estimation and pilot allocation in massive MIMO is measured based on the achieved data rate in the uplink, where we use linear MMSE filters based on the MMSE channel estimates and the channel estimation error covariance matrices [14]. As a comparison, we depict the performance of a regularized zero-forcing filter which directly uses the LS estimates $\boldsymbol{y}_p[t]$ and the performance of the MMSE filter with perfect knowledge of the channel vectors.

A. Performance with Respect to the Number of Users

We investigate the behaviour of the system for a fixed amount of training slots and different numbers of simultaneously served users. Figure 1 shows the results for $T_{\text{tr}} = 20$, $G = 10$ and $M = 400$. For all curves which exploit the CDI, the rate improves first by increasing the number of users due to the multiplexing gains. For large numbers of users, the additional interference and pilot-contamination dominates and the rates start to decrease. The optimal amount of users in the case of imperfect CDI (ICDI) is around $K = 60$. For $K = 60$ users, our method for pilot-allocation leads to a gain in total throughput of 22%.

As can be inferred, the LS based filter has the worst performance due to pilot contamination. A substantial improvement is achieved using the MMSE filter based on the MMSE estimates, and additional gain can be observed by an appropriate user assignment. For the case of perfect CDI we could have, as mentioned earlier, used the same pilot allocation in each channel coherence interval, i.e., chose the allocation which allows for the most effective mitigation of pilot-contamination. However, to get a better comparison between the CDI and the ICDI case, we use the greedy algorithm in Algorithm 1 also for perfect CDI. That is, the gap between perfect and imperfect CDI depends only on the accuracy of channel estimation and the choice of α. For the simulation, we chose $\alpha = 0.01$ which should provide a reasonable trade-off between adaptivity and accuracy in practice.

B. Performance with Respect to Adaptivity

To investigate the impact of the adaptivity of the covariance matrix estimation on the user rate, simulations were performed for different values of α. The results are shown with respect to $1/\alpha$ which can be interpreted as the implicitly assumed coherence interval of the covariance matrices. Of course, for a longer coherence interval, we expect a more accurate estimation and thus better performance in the case of imperfect CDI. In Figure 2, the curves for full channel state information, LS based filter and perfect CDI are obviously independent of α. On the contrary, imperfect (coordinated) CDI strongly depends on the value of α, since the estimation of the covariance matrices improves for smaller values. The gap between perfect and imperfect CDI can be further reduced, e.g., by extending the estimator introduced in Section III to maximum-likelihood estimation [8].

C. Performance with Different Group Sizes

In this scenario, we analyze the effect of G on the data rate. It can be clearly seen from Figure 3 that the rate improves

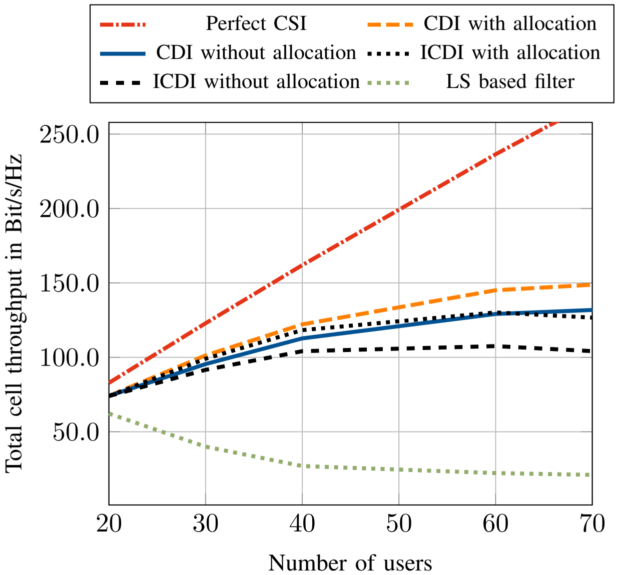

Fig. 1. Cell throughput for different numbers of users K in a single-cell with $T_{\text{tr}} = 20$, $G = 10$ and $M = 400$. We compare previous results with perfect knowledge of the covariance matrices (CDI) with our novel approach for pilot allocation based on estimated covariance matrices (ICDI).

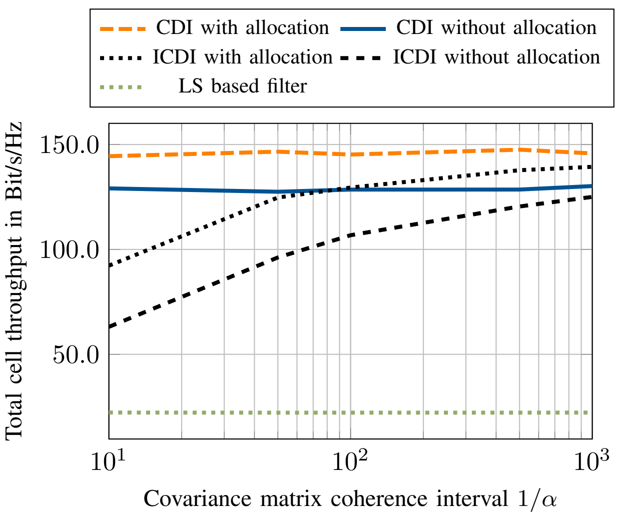

Fig. 2. User rate for different values of α with $M = 400$, $K = 60$, $T_{\text{tr}} = 20$ and $G = 10$.

with larger G for coordinated pilot assignment with perfect and imperfect CDI at the base station, since more degrees of freedom are available for mapping the users to groups. In an uncoordinated scenario, the assignment is done at random and hence the performance is independent of G. Note that, as mentioned before, the choice of G is a trade-off between accuracy of the covariance matrix estimation and degrees of freedom for the user allocation. However, for practical system parameters, the accuracy of the covariance matrix estimation is of minor concern and the number of groups should be chosen as large as possible, i.e., $G = T_{\text{tr}}/2$.

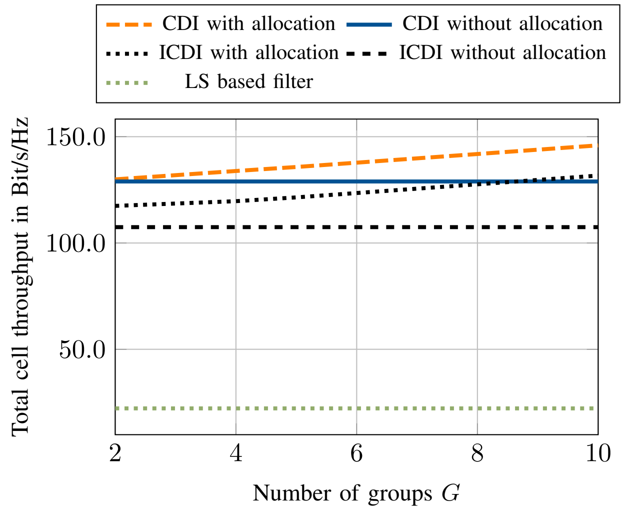

Fig. 3. Cell throughput for different values of G with $M = 400$, $K = 60$, $T_{\text{tr}} = 20$ and $\alpha = 0.01$.

VI. Conclusion

This paper presents a new approach for allocation of users to pilot sequences in a practical massive MIMO systems with imperfect knowledge of the channel covariance matrices at the base station. We formulated an optimization problem with respect to the pilot allocation which considers the dynamic pilot allocation necessary for the estimation of the covariance matrices. We showed that the allocation problem can be reduced to assigning users to groups, and proposed a greedy algorithm for efficient implementation. Simulation results demonstrated an additional gain with greedy user assignment compared to the uncoordinated scenario.

References

[1] F. Rusek, D. Persson, B. K. Lau, E. Larsson, T. Marzetta, O. Edfors, and F. Tufvesson, "Scaling up MIMO: Opportunities and challenges with very large arrays," *Signal Processing Magazine, IEEE*, vol. 30, no. 1, pp. 40–60, Jan 2013.

[2] J. Nam, A. Adhikary, J.-Y. Ahn, and G. Caire, "Joint spatial division and multiplexing: Opportunistic beamforming, user grouping and simplified downlink scheduling," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, pp. 876–890, Oct. 2014.

[3] H. Yin, D. Gesbert, M. Filippou, and Y. Liu, "A coordinated approach to channel estimation in large-scale multiple-antenna systems," *IEEE Journal on Selected Areas in Communications*, vol. 31, no. 2, pp. 264–273, February 2013.

[4] D. Neumann, M. Joham, and W. Utschick, "CDI precoding for massive MIMO," in *Proceedings of 10th International ITG Conference on Systems, Communications and Coding (SCC)*, Feb 2015, pp. 1–6.

[5] D. Neumann, A. Grundinger, M. Joham, and W. Utschick, "Rate-balancing in massive MIMO using statistical precoding," in *Signal Processing Advances in Wireless Communications (SPAWC), 2015 IEEE 16th International Workshop on*, June 2015, pp. 226–230.

[6] M. Li, S. Jin, and X. Gao, "Spatial orthogonality-based pilot reuse for multi-cell massive MIMO transmission," in *International Conference on Wireless Communications Signal Processing (WCSP)*, Oct 2013, pp. 1–6.

[7] V. Saxena, G. Fodor, and E. Karipidis, "Mitigating pilot contamination by pilot reuse and power control schemes for massive MIMO systems," in *IEEE 81st Vehicular Technology Conference (VTC Spring)*, May 2015, pp. 1–6.

[8] D. Neumann, M. Joham, and W. Utschick, "Covariance matrix estimation in massive MIMO systems," 2016, in preparation.

[9] N. Shariati, E. Bjrnson, M. Bengtsson, and M. Debbah, "Low-complexity polynomial channel estimation in large-scale mimo with arbitrary statistics," *IEEE Journal of Selected Topics in Signal Processing*, vol. 8, no. 5, pp. 815–830, Oct 2014.

[10] T. L. Marzetta, G. H. Tucci, and S. H. Simon, "A random matrix-theoretic approach to handling singular covariance estimates," *IEEE Transactions on Information Theory*, vol. 57, no. 9, pp. 6256–6271, Sept 2011.

[11] D. Neumann, M. Joham, L. Weiland, and W. Utschick, "Low-complexity computation of LMMSE channel estimates in massive MIMO," in *Proceedings of 19th International ITG Workshop on Smart Antennas (WSA)*, March 2015, pp. 1–6.

[12] T. Yoo and A. Goldsmith, "Optimality of zero-forcing beamforming with multiuser diversity," in *IEEE International Conference on Communications (ICC)*, vol. 1, May 2005, pp. 542–546 Vol. 1.

[13] 3GPP, "Spatial channel model for multiple input multiple output (MIMO) simulations," 3rd Generation Partnership Project (3GPP), Tech. Rep., 2015.

[14] D. Neumann, M. Joham, and W. Utschick, "On MSE based receiver design for massive MIMO," in *Proceedings of 11th International ITG Conference on Systems, Communications and Coding (SCC)*, Feb 2016, pp. 1–6.