This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

Centralized UL/DL Resource Allocation for flexible TDD Systems with Interference Cancellation

Anna Łukowa, Venkatkumar Venkatasubramanian

***Abstract*—Flexible time division duplex (TDD) is expected to be one of the key technologies for 5G flexible air interface. The main benefit of flexible TDD is to allow for better radio resource utilisation based on instantaneous user traffic demands. To boost network capacity, one could thus utilise flexible TDD along with a dense deployment of small-cells. Flexible TDD systems can be designed to benefit from strong-interference cancellation of the cross-link interference along with advanced inter-cell coordination. In this paper, we propose a centralized coordination scheme with joint UL/DL user scheduling, rate and MIMO rank adaptation. To reduce the complexity of coordinated scheduling, we propose a cluster-based scheduling scheme where joint UL/DL selection and joint user scheduling is performed within a cluster. MIMO rank and rate adaptation is performed across the clusters using inter-cluster rate and rank adaptation.**

Simulation based results are evaluated in an outdoor urban micro scenario with bursty traffic and assuming realistic channel estimation. It is observed that the proposed flexible TDD scheme with joint UL/DL scheduling, interference cancellation and coordination provides over 120% gain in the 5^{th} percentile and over 29% gain in the average end-to-end uplink throughput as compared to flexible TDD standalone scheduling without coordination. Our results show that successive interference cancellation receivers provide over 20% uplink throughput gain for flexible TDD systems as compared to IRC receivers. We also find that fully flexible TDD with interference cancellation significantly outperforms fully-synchronised TDD for different cell sizes.

***Keywords*—5G, cross-link interference, interference cancellation, flexible TDD, RRM, rank coordination**

I. Introduction

The continuing growth of the number of mobile subscribers, growth in traffic volume and diverse applications [1] requires a flexible and forward compatible design of cellular systems. Thus, new generation of 5G mobile communication system are envisioned to incorporate a flexible air interface which can be adapted to meet the quality of service requirements for all users. Flexible time division duplex (TDD) is a key 5G technology for flexible air interface, which allows TDD switching in short durations. This duplexing scheme allows for better radio resource management by flexible spectrum utilisation of uplink (UL) and downlink (DL) data transmission depending on the current user's needs.

To further increase network capacity, a dense deployment of small and heterogeneous cells [2] can be used with the flexible air interface. With densification, the number of users served by a single cell is reduced which results in increased amount of radio resources per user. However, reducing the cell sizes can lead to instantaneous uplink/downlink traffic asymmetries between neighbouring cells and thus result in more severe inter-cell cross-link interference situations as downlink-to-uplink interference or uplink-to-downlink interference.

Thus, from a system perspective, flexible TDD based systems need inter-cell interference management techniques as part of radio resource management (RRM).

In LTE, interference mitigation was mainly based on inter cell interference coordination (ICIC) scheme, which was introduced in 3GPP Release 8 [3]. It allows to coordinate radio resource management using fractional frequency reuse based on exchange of information between the cells. Enhanced ICIC in Release 10 [4] further introduced the concept of almost blank subframe (ABS). This allows the cells to coordinate muting in both frequency domain and in time domain. For flexible TDD systems, cross-link interference mitigation based on muting may not fully utilise the RRM gains of flexible switching in certain deployment scenarios. Therefore, coordinated scheduling between the cells by user scheduling and multiple-input multiple-output (MIMO) transmission coordination can be used for advanced inter-cell coordination in addition to resource muting as shown in [5]. At the same time, advanced receivers which are capable of cancelling inter-cell interference can be used for effective interference mitigation. The performance of advanced receivers can be further enhanced by transmitter-side coordination using coordinated MIMO adaptation, modulation and coding scheme (MCS) coordination and scheduling coordination between the cells. In this article, we consider joint interference cancellation and transmitter coordination in a flexible TDD system. Thus while most of the conventional interference management techniques for TDD tend to avoid strong interference [6], [7], in our work we propose a novel approach where strong cross-link interference is scheduled on purpose so that it can be cancelled using advanced receivers. Such an approach intuitively makes use of the strong interference regime [8] where a strongly interfering communication link in certain conditions can achieve rate a as high as would be achievable without the interference.

II. Related work and main contribution

Interference cancellation and inter-cell coordination has been considered in prior works as follows.

Interference suppression with Interference Rejection Combining (IRC) receiver was recently looked in [9], where IRC is enhanced by using cross-link interference measurement. In [10], the authors compare different types of receivers for handling strong interference in flexible TDD systems. They compare the performance of interference suppressing linear IRC receiver to IS-SIC (inter-stream successive interference cancellation). The IS-SIC performs inter-stream interference cancellation without any inter-cell interference cancellation.

A. Łukowa is with Nokia, Wrocław, Poland, and also with the Department of Electronics and Telecommunications, Poznań University of Technology, Poznań, Poland (e-mail:anna.lukowa@nokia.com)

V. Venkatasubramanian is with Nokia, Wrocław, Poland (e-mail: venkatkumar.venkatasubramanian@nokia.com)

Manuscript received 16.01.2018; revised 17.12.2018.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

Furthermore, the authors propose to use the interference taxation principle for victim aware rank adaptation as in [11]. The studies show throughput gains of 20% in the case of IS-SIC receivers as compared to IRC receivers. Taxation based rank adaptation shows 5% gains in case of low network load. However, the scheme does not make use of inter-cell interference cancellation as well as inter-cell coordinated scheduling, thus limiting the overall systems level gains of flexible TDD and interference cancellation.

Inter-cell interference cancellation is considered in [12] for standalone systems and performance evaluation is presented. The authors assume no coordination between cells and derive interference cancellation probability through system level simulation. It is shown that the performance of SIC receivers highly depends on the number of strong interferes and their transmission rates. This approach however does not consider transmission coordination between cells that can be beneficial for enhancing the performance of SIC receivers.

MIMO rank coordination is exploited with SIC receivers in [13] where the authors use interference cancellation in the context of centralized cell association. Here, they perform user association to a secondary cell, and then handle interference from the primary cell with the use of network assisted interference cancellation (NAICS) receivers. However, this work assumes only the cancellation of a single stream from only one dominant interferer using symbol level interference cancellation without modulation and coding scheme coordination. The results are shown in a 2x2 MIMO Frequency Division Duplex (FDD) system in a bursty traffic scenario without strong cross-link interference. Results show that centralized UE and cell association together with NAICS receivers gives gains in the order of 15-80% for 5^{th} percentile users depending on the offered cell load. Rank coordination is observed to further increase the 5^{th} percentile NAICS gains by 4-17%.

In [14], similar to [15], the author proposes a rate coordination scheme between a victim and its dominant interferer in a SISO system. The dominant interferer can adjust its transmission rate and transmit power so that it can be cancelled at the interfered receiver. The results shown for an LTE FDD SISO system demonstrate around 11% gain in the cell edge throughput as compared to the standalone scheduling. One main shortcoming of the proposed solution is that the UE selection is done separately by each cell which can be suboptimal for flexible UL/DL switching in TDD system. Furthermore, rate coordination, muting and scheduling are not jointly performed across the cells. In this article, we address some of the limitations of the aforementioned work in the literature and consider strong interference cancellation along with transmission rate and rank coordination in a MIMO flexible TDD system. One main challenge at the system level is that a victim receiver can be affected by multiple multi-stream strong interferers in a dense deployment scenario. We thus consider the problem of multiple interferers and propose joint clustering, and joint UL/DL scheduling and rank coordination which further extends our previous work in [16], [5].

In a system with a large number of cells, finding the optimal transmission parameters across cells has a high computational complexity. We thus propose a central coordinated scheme which utilises sub-steps of clusters formation, intra-cluster coordination as well as inter-cluster coordination. Cooperative communications based on dynamic clustering has been addressed in prior works [17]-[22]. Cluster-based resource allocation is considered in [18]-[22], where typically clustering is done to allocate the interfering small cells to non-overlapping sub-bands. In [21], [22], a greedy dynamic clustering scheme is applied for the purpose of reducing the complexity of the joint processing of the uplink. Recently [23] considers cluster-specific TDD switching for dynamic TDD. In contrast to the aforementioned approaches, in this paper we consider dynamic clustering for the purpose of dynamic TDD resource allocation based on interference-cancellation and rank coordination.

In our proposed solution, intra-cluster coordination performs joint UL/DL scheduling and cell muting for a sub-set of cooperating cells. Inter-cluster coordination performs MIMO rank and rate coordination across all the cells in multiple clusters. The clusters which are sub-sets of cooperating cells are formed dynamically based on the short term traffic and interference conditions in the network. For MIMO rank coordination, we utilise a low-complexity heuristic called successively paired rank allocation (SPARK) [5].

The main contributions of this paper are as follows:
1) Proposing a centralized clustering based coordination scheme for flexible TDD coordination with interference cancellation, 2) Evaluating the performance of the proposed centralized approach in a system with a large number of cells, 3) Comparison to centralized and time flexible fully-synchronised TDD to standalone flexible TDD without coordination.
Thus through this paper we try to answer the question: To what extent can we benefit from advanced receivers and inter-cell coordination for flexible TDD in a setup with a large number of cells? We further evaluate the extent of benefit that can be obtained with flexible TDD for different densification scenarios using our proposed solution. Simulations are performed in a realistic setting assuming imperfect channel estimation and bursty traffic (finite buffer).

III. System Model

We consider a system based on flexible TDD and multiple-input multiple-output orthogonal frequency division multiple access (MIMO-OFDMA) for 5G [24]. In flexible TDD, each subframe of 0.5 ms carries both uplink and downlink control information (with e.g. scheduling decisions or transmission format indications) and demodulation reference symbols (DMRS) [25]. In this paper we assume that flexible uplink and downlink resource switching is made on scheduling slots of 0.5 ms duration. A scheduling slot is flexibly switched to either the downlink or the uplink independently per cell. The MIMO system comprises M antennas at all transmitters and receivers. The unit norm MIMO channel fading matrix between the transmitter of a link k and a receiver of a link i is represented by a matrix $\mathbf{G}_{ki}$ of dimension $M \times M$. Correspondingly, the channel matrix between the transmitter and the receiver of the same link k is represented by $\mathbf{G}_{kk}$. With M antennas, the transmitter of a link k can transmit $\tilde{M}_k \leq M$ spatial data streams.
The link throughput at the MIMO receivers is maximised

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

by the use of signal precoding before the transmission. For channel precoding we use the Singular Value Decomposition (SVD) of the channel matrix $\mathbf{G}_{kk}$. Thus, the effective channel $\mathbf{H}_{kk}$ of a link k after precoding becomes:

$$\mathbf{H}_{kk}^{M\times\tilde{M}_k} = \mathbf{G}_{kk}^{M\times M}\mathbf{V}_{kk}^{M\times\tilde{M}_k}, \quad (1)$$

where $\mathbf{V}_{kk}^{M\times\tilde{M}_k}$ is the precoder matrix, where the number of columns reflects the number of transmitted streams $\tilde{M}_k$. In a TDD system with L cells, the signal at each receiver can be corrupted by the interference from $L-1$ interfering links which can be either uplink or downlink. The set of active interferers is then represented by $\mathbf{I}$, with the number of interferers $N_I \leq L-1$. The received signal $\mathbf{y}_i$ at a receiver of link i can be expressed as:

$$\mathbf{y}_i = \sqrt{P_i\alpha_{ii}}\mathbf{H}_{ii}^{M\times\tilde{M}_i}\mathbf{x}_i + \sum_{k\in\mathbf{I}}\sqrt{P_k\alpha_{ki}}\mathbf{H}_{ki}^{M\times\tilde{M}_k}\mathbf{x}_k + \mathbf{n}, \quad (2)$$

where $\mathbf{H}_{ii}$ is a channel matrix of a serving link i of dimension $M\times\tilde{M}_i$, $\mathbf{n}$ denotes the noise vector with per-subcarrier thermal noise power of N_o and vector $\mathbf{x}_i$ denotes the modulated symbols of transmitter of link i, which is of dimension $\tilde{M}_i\times 1$. α_{ki} is the pathloss between the k^{th} transmitter and i^{th} receiver. P_i denotes the per-subcarrier transmit power of the i^{th} transmitter. The downlink uses a constant power per subcarrier to all the users, whereas the uplink is based on LTE-like open-loop power control. In the following section, we describe the RRM approaches for the flexible TDD system.

IV. RRM APPROACHES

To handle strong interference in flexible TDD system, advanced receivers together with inter-cell transmission coordination techniques can be used. The RRM approaches compared in this work are as follows:

A. Uncoordinated, per-cell flexible TDD scheduling

In uncoordinated, per-cell flexible TDD scheduling, each cell autonomously performs uplink/downlink switching and user scheduling. The scheduling is performed in a selfish manner, i.e. each cell tries to maximise its own performance based on interference awareness. In order to schedule users, each cell has to make interference estimation to determine the signal to interference noise ratio (SINR). The interference estimation and thus SINR estimation can be based on the knowledge from the past transmission slots. One other variant is to enable dedicated interference estimation opportunities between the neighbouring cells. Thus, with the interference estimation, the scheduling SINR calculation can be based on interference cancellation receiver. However, the interference generated to the neighbour cells is not taken into account and thus per-cell flexible TDD switching decision can be suboptimal from the system point of view. Moreover, because there is no scheduling or muting coordination between the cells, it can lead to suboptimal use of IC receivers especially in the case of multiple interferer scenarios.

B. Coordinated, fully-synchronised TDD scheduling

In coordinated fully-synchronised TDD the transmission direction can be fast adapted to either uplink or downlink for all of the coordinating cells. The fast synchronised UL/DL adaptation can be done on a scheduling slot basis (of e.g 0.5 ms). This scheme completely avoids UL-DL cross interference between neighbour cells. The transmission direction for all of the cells can be optimally switched in every scheduling slot based on a system level objective such as packets delays and buffer status information of the users. The cells can further coordinate user scheduling, MIMO rank and muting to mitigate UL-to-UL and DL-to-DL interference. In addition, IC receivers can be used to mitigate the dominant interference.

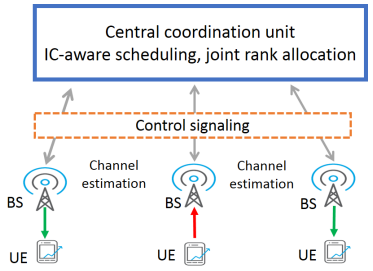

Fig. 1. Small cell flexible TDD coordination.

C. Coordinated, per-cell flexible TDD scheduling

Coordinated, per-cell flexible TDD scheduling makes cell specific UL/DL decisions based on the instantaneous uplink and downlink traffic priority and demand in each cell. One extreme case of coordinated scheme is centralized scheduling, where all the scheduling and channel state information is available at a central unit (CU). The central entity can perform joint user scheduling, flexible UL/DL switching decisions as well as transmission MIMO rank and MCS selection. The CU makes its decisions based on the packet delays and buffer status scheduling information available from the cooperating set of cells as shown in Fig.1.

1) Centralized scheduling assumptions: We consider a fully centralized system which performs joint scheduling decisions every scheduling slot for a cluster of small cells (Fig.1). The user equipments send and receive data from the small cells, while the scheduling, rate and rank allocation is done by a central unit. We assume that the central entity obtains full knowledge about packet delays as well as MIMO channel matrices (using TDD reciprocity or feedback) of all the links via the small cells.

The channel state information $\hat{\mathbf{G}}_{ki}^{M\times M}$ at the central scheduler is modeled as:

$$\hat{\mathbf{G}}_{ki}^{M\times M}[t] = \frac{1}{\sqrt{1+\sigma_E^2}}(\mathbf{G}_{ki}^{M\times M}[t-\delta] + \mathbf{E}). \quad (3)$$

Here, we have modeled that the actual channel knowledge at time instance $[t]$, $\mathbf{G}_{ki}^{M\times M}$, can be outdated by a delay $[\delta]$ because of non-ideal backhaul to the scheduler and corrupted by noise matrix $\mathbf{E}$.

The entries of noise matrix e_{ij} are complex-normal values with mean power σ_E^2: $e_{ij} \sim CN(0,\sigma_E^2)$ [26]. $\sqrt{1+\sigma_E^2}$ is an average power normalization scaling factor and σ_E^2 is the error variance reflecting the channel quality estimation in a function of the received pilot power. We represent channel estimation error variance in (3) with a LTE-like model [26] :

$$\sigma_E^2 = \frac{c_E}{(\sigma_{ii}^2)}(N_o + \sum_{k\in I}\sigma_{ki}^2), \quad (4)$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

where σ_{ii}^2 is the pilot signal power, N_o is the thermal noise power, and σ_{ki}^2 is the inter-cell interference power, and c_E is a model parameter found to be 0.0544 with a LMMSE estimator in [26] for pedestrian scenarios. We further assume that a receiver performs multi-cell channel estimation and estimates the weaker pilot signals using pilot IC. The channel estimation of stronger pilot is thus imperfect because of noise and interference. Accordingly, the channel estimation error variance of weaker pilots is modeled based on the residual interference power after cancelling the stronger pilots.

The central entity performs joint UL/DL scheduling, user selection, rank and rate allocation with an awareness of interference cancellation (IC-aware) capability at the receivers. At the scheduler, the possibility of interference cancellation is decided based on achievable SINR estimation using the MIMO channel state information. In practice, the interference cancellation receiver performance will be affected by the mismatch between the estimated and current channel realization at the receiver which will result in imperfect cancellation. Thus, to capture the effect of channel estimation imperfection on IC, we generate a 'noise perturbed version' of the channel available at the scheduler as:

$$\tilde{\mathbf{G}}_{ki}^{M\times M} = \frac{1}{\sqrt{1+\sigma_{\Delta_{ki}}^2}}(\hat{\mathbf{G}}_{ki}^{M\times M} + \mathbf{\Delta}_{ki}), \tag{5}$$

where $\tilde{\mathbf{G}}_{ki}^{M\times M}$ is the noise perturbed version of the channel and $\mathbf{\Delta}_{ki}$ represents the perturbation noise. The perturbation noise power corresponds to the channel estimation noise error as seen at each receiver. The precoders $\hat{\mathbf{V}}_{kk}^{M\times \tilde{M}_k}$ are obtained using singular value decomposition of $\hat{\mathbf{G}}_{ki}^{M\times M}$ as described in (5). Thus, the scheduler effective channel matrix (with artificial error) is $\tilde{\mathbf{H}}_{ki}^{M\times \tilde{M}_k} = \tilde{\mathbf{G}}_{ki}^{M\times M}\hat{\mathbf{V}}_{kk}^{M\times \tilde{M}_k}$.

V. IC-AWARE RESOURCE ALLOCATION PROBLEM

A. The problem definition

The centralized resource allocation problem at the central unit is to perform joint UL/DL user scheduling, rate and rank allocation for each resource block and scheduling slot:

$$\begin{aligned} &\underset{\mathbf{R},\mathbf{b},\tilde{\mathbf{M}}}{\operatorname{argmax}} \sum_{c\in\mathbf{C}}\sum_{k\in\mathbf{L_c}} w_k b_k \sum_{m=1}^{\tilde{M}_k} r_{km}(\mathbf{b},\tilde{\mathbf{M}}) \\ &s.t.\ \ \tilde{M}_k \in [1,..,M],\ b_k\in[0,1],\ \sum_{k\in\mathbf{L_c}} b_k \le 1,\ \forall_{c\in\mathbf{C}}. \end{aligned} \tag{6}$$

In (6), $\mathbf{C}$ is the set of all cells which are coordinated by a central entity. In our notation, link k is defined as a unidirectional link (either uplink or downlink) between a transmitter and a receiver, while the links are also sequentially indexed across the cells. The set $\mathbf{L}$ contains all unidirectional links from all of the $\mathbf{C}$ cells in coordinated systems. $\mathbf{L}_c$ is a subset of links in $\mathbf{L}$, belonging to a given cell c, thus $\mathbf{L} = \mathbf{L}_1 \cup \mathbf{L}_2 \cup ... \mathbf{L}_c,\ \forall_{c\in\mathbf{C}}$. The objective in (6) emulates the modified-largest delay weighted first (M-LDWF) discipline [27]-[29] for a multi-cell system. The rationale behind the M-LWDF approach, which is also called as Maxdelay in [30] is to perform scheduling to keep the user queues stable, which is done by using the head of line packet delay w_k and the channel capacity r_{km}. (6) emulates the modification of M-LDWF as in [30] by summing over the set of coordination base stations on each resource block.

TABLE I. TABLE OF CONSTANTS

Constant	Value
$\mathbf{C}, c$	$\mathbf{C}$ is the set of active cooperating cells and c is a single cell.
$\mathbf{L}$	The set of all active links within cooperating cell set $\mathbf{C}$, where $\mathbf{L} = \mathbf{L}_1 \cup \mathbf{L}_2 ... \mathbf{L}_C$ with $\mathbf{L_c}$ being the set of active links in a cell c.
M, m	M is maximum number of spatial streams and m is the m^{th} spatial stream
w_k	Maximum packet delay of link $k \in \mathbf{L}$ with $w_k = w_{kk}$.

TABLE II. TABLE OF VARIABLES

Variable	Value
$\tilde{M}_k$	Variable capturing rank selection of k^{th} link. $\tilde{\mathbf{M}}$ is a rank vector with entries $\tilde{M}_k \in [1,..,M]$.
b_k	Binary variable capturing scheduling decisions for a link $k \in \mathbf{L}$ with $b_k = b_{kk}$. $\mathbf{b}$ is a scheduling vector with entries b_k of a dimension $1\times\|\mathbf{L}\|$.
$\mathbf{C}_q$	The set of active cooperating cells in q^{th} cluster, $\mathbf{C}_q \subset \mathbf{C}$
$\mathbf{D}_{q-1}$	Set of already scheduled links in clusters $1\ to\ q-1$. The scheduling decisions are captured in vector $\mathbf{b}$, where $b_k = 1\ \forall k \in \mathbf{D}_{q-1}$
r_{kim}	Achievable rate for m^{th} stream between transmitter of link $k \in \mathbf{L}$ and the receiver of link $i \in \mathbf{L}$. The serving link rate r_{kkm} is denoted as r_{km}. The transmission rates of all of the links in a system are stored in matrix $\mathbf{R}$ which is of dimension $\|\mathbf{L}\| \times M$ with entries r_{km}.
$\bar{r}_{ki}$	The pathloss-based estimated rate between the transmitter of a link $k \in \mathbf{L}$ and the receiver of a link $i \in \mathbf{L}$. $\bar{r}_{ki}^{IC}$ represents post-IC rate

The constant w_k thus captures the packet delay of a link $k \in \mathbf{L}_c,\ \forall_{c\in\mathbf{C}_q}$. The role of weights w_k is to soft-prioritize the cells which currently experience higher packet delays without fully sacrificing spectral efficiency. It can be noted that when the weights are the same for all the links, the metric in (6) becomes a sum rate metric. We also further impose the constraints that on a resource block only one link can be scheduled per each cell, while the entire scheduling slot is either fully downlink, uplink or muted in each cell.

In (6), the scheduling decisions made for all links within the set of cooperating cells is captured by a vector $\mathbf{b}$. The vector $\tilde{\mathbf{M}}$ captures MIMO rank decisions for all of the links with $\tilde{M}_k$ representing the number of spatial streams of link k, with $1 \le \tilde{M}_k \le M$. Thus, in order to maximise the system objective in (6) we allocate a joint set of MIMO ranks across all cells, which is also called joint rank allocation. To determine the achievable spectral efficiency in each link, we assume successive interference cancellation (SIC) receivers. We assume that the receiver cancels strong inter-cell interference based on the decoding of the interference whenever possible. The decoding is done based on a priori knowledge of the interferer's modulation and coding scheme and includes residual interference because of imperfect channel estimation.

The achievable spectral efficiency between the receiver of link i and the transmitter of link k on the spatial stream m is denoted by r_{kim}. This corresponds to the interferers' quantized [31] modulation and coding scheme that can be decoded by the SIC receiver. The scheduled spectral efficiency of a serving link k is denoted as r_{km}, which depends on the interference experienced by the receiver. Thus, we denote the spectral efficiency $r_{km}(\mathbf{b},\tilde{\mathbf{M}})$ as a function of rank as well as the function of scheduled links. The $|L| \times M$ matrix $\mathbf{R}$ with entries r_{km} tabulates all of the coordinated rates.

We focus on the strong interference situations where an interference level could be several decibels above that of the intended signal. Thus, for the purpose of strong interference cancellation, we define a victim link set $\mathbf{V}$ which comprises

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

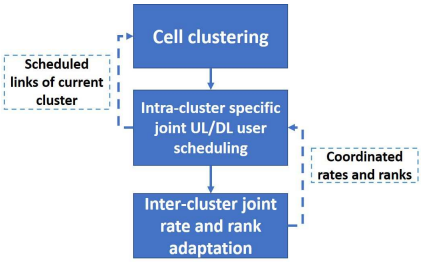

Fig. 2. Proposed decomposition of cell clustering and centralized UL/DL scheduling performed in each scheduling slot

links which suffer interference a certain threshold above the serving signal. The link is defined as a victim whenever any of its MIMO spatial streams has SINR lower than 0 dB because of the interference. For each victim $v \in \mathbf{V}$, we identify its set of strong interferers $\mathbf{S}_v$. A given victim uses a certain decoding order based on the received power of the interfering signals. Tables I and II contain a short description of all variables and constants used in (6).

VI. Proposed IC-aware solution

Finding the optimal set of scheduling decisions and transmission parameters to solve (6) is burdened with a high computational complexity because the optimal solution requires a brute-force search across all cells. The full brute force search becomes impractical for large number of cells. Thus, we propose an iterative cluster and scheduling approach, where the next cluster is formed after performing joint scheduling of current cluster. Using this approach we decouple the problem in (6) into three sub-problems:
- Finding the cell sets to form the current cluster,
- Intra-cluster joint UL/DL user scheduling,
- Inter-cluster joint rank and rate allocation, where inter-cluster corresponds to joint coordination between current cluster as well as the previous clusters.

A block diagram showing the interactions between the sub-problems is presented in Fig.2, where it is shown that the cluster formation and joint user scheduling is done in iterations. We thus modify the optimization problem from (6) as:

$$\underset{\mathbf{R},\mathbf{b},\tilde{\mathbf{M}}}{\operatorname{argmax}} \Bigg(\underbrace{\sum_{c\in\mathbf{C}_q}\sum_{k\in\mathbf{L_c}} w_k b_k \sum_{m=1}^{\tilde{M}_k} r_{km}(\mathbf{b},\tilde{\mathbf{M}})}_{q^{th}\text{ cluster UE scheduling, rate and rank allocation}} + \underbrace{\sum_{k\in\mathbf{D}_{q-1}} w_k \sum_{m=1}^{\tilde{M}_k} r_{km}(\mathbf{b},\tilde{\mathbf{M}})}_{\text{Scheduled links rate and rank allocation for previous clusters}} \Bigg) \quad (7)$$

$$s.t.\ \ \tilde{M}_k \in [1,..,M],\ b_k \in [0,1],\ \sum_{k\in\mathbf{L_c}} b_k \leq 1,\ \forall_{c\in\mathbf{C}_q}.$$

As opposed to the metric in (6), in (7) coordinated joint user scheduling is performed only for the cells within the current cluster $\mathbf{C}_q$ which is shown as the first term in (7). The scheduling hypothesis of a k^{th} link in the current q^{th} cluster is captured by b_k. $\mathbf{D}_{q-1}$ corresponds to the links scheduled in previous clusters, where $b_k = 1, \forall k \in \mathbf{D}_{q-1}$. Thus in (7) we differentiate three subproblems as follows:

1) First subproblem: IC-aware cell clustering

Clustering is performed by gathering sub-set of cells into a group. The main goal of clustering is to form a cell-set so that we can benefit from joint UL/DL scheduling within the cluster. We perform clustering which maximises the weighted sum rate, which is a generalization of the approach in [21],[22]. Furthermore, while [21],[22] perform clustering after user scheduling, here we perform clustering for the purpose of user scheduling with lower complexity.

Thus in our clustering approach, the head of line packet delays and interference cancellation capabilities are taken into account for identifying the cells in a cluster. Moreover, the clusters are formed in iterative manner, where each cluster formation is followed by user scheduling within that cluster, and the scheduling information from previous clusters is used for the current cluster formation. The proposed clustering is thus dynamic and does not depend only on the physical location of the cells. This means that cells of the same cluster do not have to be close neighbors. The detailed description of cell clustering is provided in the *Algorithm 1*.

2) Second subproblem: Intra-cluster specific joint UL/DL user scheduling

In this step, we search for the optimal UL/DL and user scheduling decision for the cells in a current cluster (*Algorithm 2*). The user scheduling decisions made for previous clusters are used as a priori information and is kept unchanged while performing scheduling decisions for current cluster.

3) Third subproblem: Inter-cluster joint rate and rank allocation

Inter-cluster joint rank and rate allocation is done within each intra-cluster joint UL/DL user scheduling iteration above. The inter-cluster coordination algorithms are described in *Algorithm 3, 4*. In inter-cluster coordination, the rate and rank allocation is optimized across the already scheduled links in previous clusters as well as for scheduling hypotheses of the current cluster. Thus inter-cluster rank and rate coordination is invoked for each joint UL/DL user scheduling search in the intra-cluster scheduling step. Once the intra-cluster user scheduling decision is made, the solution then proceeds to form the next cluster. The next cluster formation now takes into account the link scheduling decisions from the previous clusters. The rate and rank allocation output of all the scheduled links is thus available after the scheduling of the final cluster.

VII. IC-aware cell clustering (Algorithm 1)

The main objective of clustering is to benefit from joint IC-aware UL/DL scheduling within the cell cluster. Once the first cell cluster set is formed, the scheduler performs joint UL/DL switching, user scheduling, rate and rank allocation for the cells within that cluster. Subsequently, to form the q^{th} cluster the output of intra-cluster scheduling (Alg. 2) for clusters 1...q-1 is taken as input. The set of all scheduled links from previous q-1 cluster set iterations is denoted as $\mathbf{D}_{q-1}$. For the purpose of cluster specific scheduling we define the set of cells forming q^{th} cluster as $\mathbf{C}_q$. The set of scheduled links in all q cluster specific scheduling iterations are gathered in the set $\mathbf{D}_q \subset \mathbf{L}$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

Step 1: Clustering hypotheses

Let us assume that the set of remaining of unclustered cells as input to a clustering step is $\mathbf{C}'$, where for the first cluster formation $\mathbf{C}' = \mathbf{C}$. The cells from previous clusters are excluded to form the active set $\mathbf{C}'$. The maximum cluster size is $Q <= \mathbf{C}$ cells and is the same for all of the formed clusters. The number of cluster cell hypotheses that is needed to form a cluster of size Q is then $A_Q = \frac{|\mathbf{C}'|!}{Q!(\lfloor \mathbf{C}'| - Q)!}$. The set of selected links combination within a given cluster cell hypothesis is denoted as $\mathbf{L}_f^a \subset \mathbf{L}$, with a representing the a^{th} cluster cell hypothesis $a = 1..A_Q$ and $f = 1..F_a$ representing link scheduling hypothesis for a^{th} cell hypothesis. F_a is the maximal number of link scheduling hypotheses within a given cluster cell hypothesis.

As mentioned above, the main goal of clustering is to identify and form a cluster such that we can benefit from joint IC-aware UL/DL scheduling. For a cluster size of Q cells, we thus wish to identify the best Q cells which might maximise the objective in (7) during the intra-cluster scheduling phase. For the purpose of clustering, we use an interference graph of the network where two transmission nodes are defined to be connected if there is an interference relation (weak or strong) between them. Thus, the interference degree of a node is the number of edges at the node in the graph. Moreover, the objective in (7) is influenced by the packet delays in the cells as well as the achievable spectral efficiencies. To estimate the achievable spectral efficiency in IC-aware system, interference-cancellation receivers are modeled.

For the sake of the complexity, we estimate the achievable spectral efficiencies for the clustering phase based on single input single output (SISO) pathloss information as follows.

Step 2: Cluster rate estimation without IC

In this step we estimate the rates of each link for a given link scheduling hypothesis and cluster cell hypothesis. The achievable rate estimation is done for each link from the set L_f^a as well as links from $\mathbf{D}_{q-1}$ which were already scheduled in previous iterations. For cell clustering without interference cancellation we choose between frequency reuse (FR) approach and full frequency reuse (FFR). In the frequency reuse approach, we calculate achievable rates without interference (based on SNR) and then scale the achievable rates with respect to the number of the interferers. For simplicity reasons, we estimate the SNR based on the pathloss of link i as: $\rho_{ii}^{FR} = \frac{P_i \alpha_{ii}}{N_0}$, where P_i represents the per-subcarrier transmit power of interfering link i, α_{ii} is a pathloss between transmitter and a receiver of a link i and N_0 is per-subcarrier thermal noise power. No interference cancellation is assumed.

The rate calculated for each link $i \in (\mathbf{D}_{q-1} \cup \mathbf{L}_f^a)$ is then:

$$\bar{r}_{ii}^{FR} = \frac{1}{\beta_f^a} \min(\log_2(1 + \rho_{ii}^{FR}), \gamma), \tag{8}$$

where γ is the maximal spectral efficiency. β_f^a is used as the frequency reuse weight which corresponds to the maximum interference degree among all the nodes in a given link scheduling hypothesis f. The estimation of β_f^a is done based on the links in the current scheduling hypothesis L_f^a as well as links from previous intra-cluster scheduling $\mathbf{D}_{q-1}$.

For the case of full frequency reuse we calculate SINR taking into account all possible interfering links :

$$\rho_{ii}^{FFR} = \frac{P_i \alpha_{ii}}{\Sigma_{k \in (\mathbf{D}_{q-1})/i} P_k \alpha_{ki} + N_0}. \tag{9}$$

Similar to (8) we calculate the rates of each link in case of full frequency reuse:

$$\bar{r}_{ii}^{FFR} = \min(\log_2(1 + \rho_{ii}^{FFR}), \gamma). \tag{10}$$

The output of this step is a rate vector $\bar{\mathbf{r}}$ with entries $\bar{r}_i = \bar{r}_{ii},\ \forall i \in (\mathbf{D}_{q-1} \cup \mathbf{L}_f^a)$ such that:

$$\bar{\mathbf{r}} = \underset{\{\bar{\mathbf{r}}^{FR}, \bar{\mathbf{r}}^{FFR}\}}{\operatorname{argmax}} \Big(\sum_{i \in (\mathbf{D}_{q-1} \cup \mathbf{L}_f^a)} \bar{\mathbf{r}}_{ii}^{FR} w_i, \sum_{i \in (\mathbf{D}_{q-1} \cup \mathbf{L}_f^a)} \bar{\mathbf{r}}_{ii}^{FFR} w_i \Big), \tag{11}$$

where w_i is the maximum packet delay of a link i.

Step 3: Cluster rate estimation with IC

In this step we estimate the post-IC rates of interfered and interfering links and store them in a vector $\bar{\mathbf{r}}^{IC}$. Links from the previous cluster iterations $\mathbf{D}_{q-1}$ are also taken into account. For a given set of considered links $\mathbf{D}_{q-1} \cup \mathbf{L}_f^a$ within the clustering hypothesis, we check if there is any victim link, which is a link with interference above the serving signal. A victim links set $\mathbf{V}$ comprises links which suffer strong interference. For each victim $v \in \mathbf{V}$ we identify its set of strong interferers $\mathbf{S}_v$.

The estimation of achievable rate of a victim link v and allocation of the rates of strong interferers is based on interference cancellation performance at the victim receiver. The set of achievable rates can now be estimated using the Multiple Access Channel (MAC) rate region [14] of a victim receiver. A victim receiver can be interfered by multiple strong interferers. Thus, the achievable post-IC SINR between the transmitter of a link $s \in \mathbf{S}_v$ and the victim receiver v is estimated using successive interference cancellation as:

$$\rho_{sv}^{IC} = \frac{P_s \alpha_{sv}}{\Sigma_{k \in (\mathbf{D}_{q-1} \cup \mathbf{L}_f^a)/(\mathbf{S}_v' \cup s)} P_k \alpha_{kv} + N_0}, \tag{12}$$

where $\mathbf{S}_v'$ is the set of interferers cancelled in previous IC iterations. The achievable rate $\bar{r}_{sv}^{IC}$ between the transmitter of a link s and the receiver of a link v is calculated using (8). To allocate the rate of a strong interferer s we take a minimum between its achievable rate at v^{th} victim receiver and its own receiver: $\bar{r}_{ss}^{IC} = \min(\bar{r}_{sv}^{IC}, \bar{r}_{ss})$. In general two or more victims can share the same strong interferer s. The rate allocated to that interferer is then taken as a minimum across all the rates allocated to that interferer by all its victims: $\min_{v \in \mathbf{V}}(\bar{r}_{sv}^{IC})$.

The achievable post-IC SINR of a victim link after interference cancellation of its strong interferers is then estimated as:

$$\rho_{vv}^{IC} = \frac{P_v \alpha_{vv}}{\Sigma_{k \in (\mathbf{D}_{q-1} \cup \mathbf{L}_f^a)/(\mathbf{S}_v \cup v)} P_k \alpha_{kv} + N_0}, \tag{13}$$

with $\mathbf{S}_v$ being a set of cancelled strong interferers. The achievable rate $\bar{r}_{vv}^{IC}$ is then calculated as in (8).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

We thus obtain a rate vector $\bar{\mathbf{r}}^{IC}$ with entries $\bar{\mathbf{r}}_i^{IC}=\bar{r}_{ii}^{IC}$, $\forall i \in (\mathbf{V} \cup (\mathbf{S}_v,\ \forall_{v\in\mathbf{V}}))$.

Step 4: Cluster cell selection

In this step, the best set of Q cells are selected to form the q^{th} cluster. For a given link scheduling hypothesis f in the a^{th} cluster cell hypothesis, we obtain

$$M_f^a = \max\Big(\sum_{i\in(\mathbf{D}_{q-1}\cup\mathbf{L}_f^a)} w_i\bar{r}_i, \beta^a w_i \sum_{i\in(\mathbf{D}_{q-1}\cup\mathbf{L}_f^a)} \bar{r}_i^{IC}\Big). \quad (14)$$

In the above equation $\beta^a = \max_f(\beta_f^a)$, which is the maximum interference degree obtained by taking the maximum over all link hypotheses for a given cluster cell hypothesis. It can be recalled that we wish to obtain maximum benefit from joint UL/DL scheduling search in the next step. Thus we now average the metric across all link scheduling hypotheses to obtain the best cluster cell hypothesis. The intention is that the best link scheduling hypothesis can be selected in the next intra-cluster joint scheduling step based on fast fading and MIMO equalizer estimation. The best cluster cell hypothesis is thus selected as:

$$a^* = \text{argmax}(\mathbf{E}[M^a]), \quad (15)$$

where $\mathbf{E}[M^a] = \frac{1}{F_a}\sum_{f=1}^{F_a} M_f^a$. The corresponding set of cells in the a^{*th} cluster cell hypothesis for the current q^{th} cluster formation is denoted as $\mathbf{C}_q$. The set of selected cells $\mathbf{C}_q$ is then used for cluster specific UL/DL joint scheduling, rate and rank allocation as described in the *Algorithm 2*.

VIII. Intra-cluster specific joint UL/DL user scheduling (Algorithm 2)

Joint UL/DL scheduling across all the cooperating cells in a system can be computationally exhaustive. Thus, in our proposed solution joint UL/DL user scheduling is performed only for the cells within a cluster using IC or IRC receivers. Moreover, user scheduling information from previously scheduled clusters 1...q-1 is used as side information to perform joint UL/DL and user scheduling in the current cluster q.

For the cells in the current q^{th} cluster we define F_q as the number of UL/DL user scheduling (link) hypotheses. In each of the f^{th} link scheduling hypothesis only one link is selected per cell. The set of links in each scheduling hypothesis f is then denoted as $\mathbf{L}_f \subset \mathbf{L}$. The scheduling and muting decisions for f^{th} link hypothesis is then captured by a vector $\mathbf{b}^{\mathbf{f}}$. Thus, for each of the f^{th} link hypothesis we find the optimal scheduling vector $\mathbf{b}^{*f}$:

$$\mathbf{b}^{*f} = \underset{\mathbf{R},\mathbf{b},\tilde{\mathbf{M}}}{\text{argmax}} \Big(\sum_{k\in\mathbf{L_f}} w_k b_k \sum_{m=1}^{\tilde{M}_k} r_{km}(\mathbf{b},\tilde{\mathbf{M}}) + \sum_{k\in\mathbf{D}_{q-1}} w_k \sum_{m=1}^{\tilde{M}_k} r_{km}(\mathbf{b},\tilde{\mathbf{M}})\Big) \quad (16)$$

$$s.t.\ \ \tilde{M}_k \in [1,..,M],$$

$$b_k \in [0,1]\,\forall k \in L_f, b_k = 1\,\forall k \in \mathbf{D}_{q-1}.$$

Note that in (16) the links scheduled in previous clusters are also taken into account by setting $b_k = 1, \forall k \in \mathbf{D}_{q-1}$. Furthermore, for efficient interference management MIMO rank adaptation and MCS selection is jointly performed across all the scheduled links from previous clusters $\mathbf{D}_{q-1}$ as well as for all of the link hypotheses in the current q^{th} cluster. The detailed description of inter-cluster joint rate and rank allocation is further discussed below in *Algorithm 3* and *Algorithm 4* respectively.

The best scheduling vector among all the $f = 1..F_a$ link hypotheses is then denoted as $\mathbf{b}^{*f^*}$

The output of cluster specific scheduling is thus a set of scheduled links $\tilde{\mathbf{D}}_q$ in the q^{th} cluster. The scheduling decision is captured by a vector $\mathbf{b} = \mathbf{b}^{*\mathbf{f}^*}$ updated in each cluster specific scheduling iteration. We update the set of scheduled links across the clusters as $\mathbf{D}_q = \mathbf{D}_{q-1} \cup \tilde{\mathbf{D}}_q$.

In (16), the spectral efficiencies r_{km} of the m^{th} spatial stream for the k^{th} link has to be estimated. The value of r_{km} for IC-receivers in turn depend on the rank selection of the other links $\mathbf{L}_f$ and $\mathbf{D}_{q-1}$. Thus inter-cluster coordinated rank selection and MCS selection is performed to efficiently support interference cancellation. In the scheduling stage rate estimation is done using fast fading MIMO channel estimates (section IX-A). The joint MCS selection (also called as joint rate adaptation) and MIMO rank selection steps are now described in detailed in the following sections.

IX. Joint rate adaptation for SIC receiver

The efficiency of interference cancellation in a victim receiver highly depends on the interference to signal noise ratio (ISNR), interferer's modulation and coding scheme and also the number of interferers [12]. Therefore full decoding SIC receivers can benefit from coordinated rate adaptation of the interferer. In the this section, we describe the algorithm for coordinated rate adaptation in the case of MIMO-SIC receivers. For this purpose we first describe the MIMO-SIC receiver model which is used to estimate the achievable spectral efficiencies of the interferers at a SIC receiver.

A. SIC receiver model

Assume a victim link v realises a set of strong interferers $s \in \mathbf{S}_v$. For low complexity implementation, the interferers are sorted in decreasing order of received interference power. The SINR estimation of an interferer s at the victim is made by using interference covariance estimation of other interferers. The IRC filter matrix $\tilde{\mathbf{W}}_{sv}$ [32] is computed from the 'noise perturbed and outdated channel' $\tilde{\mathbf{H}}_{sv}$as:

$$\tilde{\mathbf{W}}_{sv} = (\sqrt{P_s\alpha_{sv}}\tilde{\mathbf{H}}_{sv}^{M\times\tilde{M}_s})^H \Big(P_s\alpha_{sv}\tilde{\mathbf{H}}_{sv}^{M\times\tilde{M}_s}(\tilde{\mathbf{H}}_{sv}^{M\times\tilde{M}_s})^H + \tilde{\mathbf{Z}}_{sv}\Big)^{-1}. \quad (17)$$

Here $\tilde{\mathbf{H}}_{sv}^{M\times\tilde{M}_s}$ represents the precoded channel matrix of MIMO rank M_s, while M is the number of receives antennas. $\tilde{\mathbf{Z}}_{sv}$ is the received interference covariance matrix at victim v which includes the main signal $\tilde{\mathbf{H}}_{vv}$ of the victim v and AWG noise. However, it does not capture interference covariance of interferers stronger than s, who are cancelled based on codeword decoding. The received covariance matrix is estimated based on multi-cell channel estimation using pilots signals as in [33].

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

The SINR calculation for the m^{th} spatial stream of link s at receiver v is given by:

$$\hat{SINR}_{svm} = \frac{P_s \alpha_{sv} \mid \tilde{\mathbf{w}}_{svm} \hat{\mathbf{h}}_{svm} \mid^2}{\tilde{\mathbf{w}}_{svm}(\hat{\mathbf{Z}}_{sv} + \sum_{j=1}^{s-1} \mathbf{K}_{jv}(\mathbf{K}_{jv})^H + \hat{\mathbf{Z}}_{str(sv)})(\tilde{\mathbf{w}}_{svm})^H}. \quad (18)$$

Here $\tilde{\mathbf{w}}_{svm}$ of dimension $1 \times M$ is the m^{th} row of the IRC matrix $\tilde{\mathbf{W}}_{sv}$. $\hat{\mathbf{h}}_{svm}$ represents the m^{th} column of the channel $\hat{\mathbf{H}}_{sv}$. Inter-stream interference is represented by: $\hat{\mathbf{Z}}_{str(sv)} = \sum_{m' \neq m} P_s \alpha_{sv} \hat{\mathbf{h}}_{svm'} (\hat{\mathbf{h}}_{svm'})^H$, while $\hat{\mathbf{Z}}_{sv}$ represents all inter-cell interference which were not cancelled.

Equation (18) captures the channel and filter mismatch because of the channel estimation noise, as well as imperfect interference cancellation because of channel estimation errors. The term $\mathbf{K}_{jv}$ models the leakage caused by imperfect interference cancellation of j^{th} interferer because of errors in channel estimation at the receiver and is expressed by $\mathbf{K}_{jv} = \sqrt{P_j \alpha_{jv}}(\hat{\mathbf{H}}_{jv}^{M \times \tilde{M}_j} - \tilde{\mathbf{H}}_{jv}^{M \times \tilde{M}_j})$, where $\hat{\mathbf{H}}_{jv}^{M \times \tilde{M}_j}$ is actual scheduler knowledge about channel between transmitter of a link j and the receiver of a link v.

The achievable rate of m^{th} stream between s^{th} transmitter and receiver v is then estimated based on Shannon capacity mapping as: r_{svm}=$Quant(\min(\log_2(1 + \hat{SINR}_{svm}/1.333), \gamma))[bits/s/Hz]$, where γ is the maximal spectral efficiency and $Quant$ is quantization function which obtains quantized modulation and coding scheme (MCS) rates as in [31]. Therefore if an interferer s uses the rate of r_{svm} it then can be cancelled at victim receiver v using codeword decoding. Finally, the post-IC rate of the victim, r_{vvm} is computed after cancelling all the strong interferers and modeling the leakage because of imperfect channel estimation. The estimated rates (spectral efficiencies) are now used for the joint rate adaptation algorithm below.

B. Joint rate adaptation algorithm (Algorithm 3)

The rate adaptation is modeled using interference cancellation receiver for links which realise strong interference with ISNR above 0 dB in any of the spatial streams. For all other links, a linear IRC receiver based rate adaptation is used. Rate adaptation in the case of IC can be based on rate coordination with the interferer or without coordination.

Uncoordinated IC rate region

In this case, each victim link would try to perform interference-cancellation of strong interferers without rate and rank coordination with the interferer. If the victim v succeeds with cancellation, the rate estimate on its m^{th} spatial stream $r_{vm}^{uncoord}$ is obtained using post-IC SINR estimate as in (17-18). If the victim is not able to cancel a strong interferer, the rate of a victim is estimated based only using IRC receiver. Thus the interferers of a victim do not adjust their rates to improve IC at the victim receiver.

Coordinated rate region

Here the strong interferers $s \in \mathbf{S}_v \ \forall_{v \in \mathbf{V}}$ coordinate their rates to fit IC capabilities of receivers of the victim links $v \in \mathbf{V}$. The rate of a strong interferer r_{svm}^{coord} is then calculated using SINR estimated as in (17-18). We determine the common rate at which a strong interferer s would transmit so that it is decodable by all its victim receivers:

$$r_{sm}^{coord} = \min_{v \in V}(\min(r_{svm}^{coord}, r_{sm}^{uncoord})), \ \forall s \in \mathbf{S}_v. \quad (19)$$

The victims' rate r_{vm}^{coord} are computed after interference cancellation. The links that are neither strong interferers nor a victim are assigned their uncoordinated rates.

Transmission rate selection

From the sets of the uncoordinated and coordinated rates we choose either the coordinated or uncoordinated rate tuples as:

$$\underset{u}{\operatorname{argmax}}(u \sum_{k \in \mathbf{L}} \sum_{m=1}^{\tilde{M}_k} b_k r_{km}^{uncoord} w_k, (u-1) \sum_{k \in \mathbf{L}} \sum_{m=1}^{\tilde{M}_k} b_k r_{km}^{coord} w_k), \quad (20)$$

where $u \in [0, 1]$ is a binary variable capturing the selection of optimal rate (coordinated or uncoordinated). The selected output rate for each link k is then:

$$r_{km} = \max(u r_{km}^{uncoord}, (1-u) r_{km}^{coord}) \ \forall k, m. \quad (21)$$

The output of the rate adaptation step is rate matrix $\mathbf{R}$ with the dimension of $|\mathbf{L}| \times M$.

The achievable spectral efficiency of the interfering links and the serving links depends on the MIMO transmission ranks of the links. Therefore, a coordinated rank adaptation algorithm is needed to find the best MIMO ranks of the links. In the following section we describe a centralized heuristic algorithm for joint rank adaptation.

X. Joint rank adaptation

In joint MIMO rank adaptation, the number of streams are allocated to each link to maximise the system performance in the intra-cluster scheduling stage. This is because the transmission rank of an interfering link impacts the SINR of another link through the interference covariance matrix at a receiver. Because of interference relations, the users who are to be scheduled in a current cluster can change the interference covariance matrices at the receivers of already scheduled links in previous clusters. Thus, in each cluster specific scheduling iteration, we perform inter-cluster rank adaptation across all current cluster link hypotheses as well as already scheduled links from previous clusters. Coordinated rank adaptation makes use of coordinated rate adaptation as described above. Joint rank allocation via a rank hypotheses using brute force search within each link hypothesis will however incur a high computational complexity for a large number of clusters. Thus, we introduce a low complexity scheme below.

A. Low complexity rank search - SPARK (Successively PAired RanK allocation - Algorithm 4)

The main idea of the heuristic is to perform joint rank allocation via a grouping mechanism as described in [5]. For this purpose, we search through the interferers in the same order which is used for successive decoding of interferers at a victim receiver. Thus, for example in order to jointly allocate the ranks of interferer A, B, C and a victim v, we first search for the optimal ranks of the strong interferer pair A and B, and then jointly choose the rank of interferer C and the

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

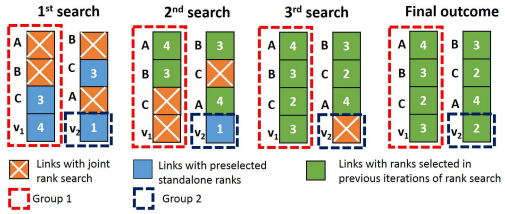

Fig. 3. An illustration of SPARK rank search for two victim links (v_1, v_2) and three strong interferers (A, B, C).

victim v (Fig.3). While this approach can in general be used for any metric, we give an implementation below for the delay-weighted metric in (7). The following steps are performed for each resource block.

Step 1 Victim set: In the first step we identify the victim links based on computing SINR for the links based on path-loss. Here we set a SINR threshold of 10dB, thus also taking into account weak interferers. All the links are also now pre-assigned their so-called standalone ranks, which maximise the capacity of the individual links without IC. The set of victim links within the coordinated cluster is denoted as $\mathbf{V}$. We sort the victims $v \in \mathbf{V}$ in descending order of the delay w_v (7). We now consider the victims sequentially in outer iterations. There are thus at most $|\mathbf{V}|$ outer iterations. The victim and interfering links which have been considered until the current iteration are stored as $\mathbf{V}_{prev}$ and $\mathbf{S}_{prev}$ respectively. $\mathbf{V}_{prev}$ and $\mathbf{S}_{prev}$ as initialized as null sets.

Step 2 Victim prioritization: In the current iteration, the victim with the highest delay and thus the largest weight w_{ck} is chosen from $\mathbf{V} = \mathbf{V} \setminus \mathbf{V}_{prev}$, where $\mathbf{X} \setminus \mathbf{Y}$ indicates a set difference between $\mathbf{X}$ and $\mathbf{Y}$.

Step 3 Rank groups: The set of strong interferers of a victim v which are not in the set $\mathbf{S}_{prev}$ is defined as $\mathbf{S}_v$. We gather together the chosen victim v and its strong interferers $\mathbf{S}_v$ into one set, which we call as a rank group: $\mathbf{C}_v = v \cup \mathbf{S}_v$.

Step 4 Interferer pairing: In this step, the interferers are considered based on the interference power which also corresponds to the order for successive interference cancellation. The interfering links already considered in previous iterations $\mathbf{S}_{prev}$ are excluded from the rank group. Note that a victim v could have been a strong interferer to a previous victim (in the previous iterations). Those victims are thus in the $\mathbf{S}_{prev}$. To avoid cycles a victim v' which belongs to the set $\mathbf{S}_{prev}$ is excluded in the rank search within the rank group of v'. Thus we denote by set $\mathbf{V}'$ the set of all victims which were excluded from the search within their own rank groups.

Step 5 Rank search: For a given pair in a rank group, we go through all possible rank combinations. In doing so, we assume all other links in that rank group and in other rank groups are set to their standalone ranks. However, the victim in that rank group, which is not to be excluded, is set to the highest rank in order to fully utilise MIMO multiplexing. While making a rank search for a pair within a rank group, the impact to all other links also in other rank groups is also taken into account according to the sum-delay-weighted metric (7). For this purpose, for a given rank combination the SINR estimation is made for all the links, which also enables us to jointly allocate the rates to all interferers until the current iteration.

1. Joint scheduling, rate and rank adaptation

$\mathbf{C} \leftarrow$ *active cell set*;
$\mathbf{L_c} \leftarrow$ *active links in cell* c;
$\mathbf{C}' \leftarrow \mathbf{C}$;
$\mathbf{D_0} \leftarrow [\,]$;
for *q=1 to* $\frac{|\mathbf{C}|}{Q}$ **do**
  IC-aware cell clustering (Alg.1) $\leftarrow \mathbf{C}', \mathbf{D}_{q-1}$;
  Output Alg.1: $\mathbf{C}_q$;
  for *all* $c \in \mathbf{C}_q$ **do**
    Intra - cluster joint UL/DL user scheduling(Alg.2) $\leftarrow \mathbf{D}_{q-1}, \mathbf{C}_q$;
    for *all scheduling hypotheses* **do**
      Inter-cluster joint rank allocation - SPARK (Alg.4);
      for *all rank hypotheses* **do**
        Inter-cluster rate adaptation (Alg. 3);
        Output Alg.3: $\mathbf{R}$ *(rate)*;
      end
      Output Alg.4: $\tilde{\mathbf{M}}, \mathbf{R}$ *(rank)*;
    end
    Output Alg.2: $\mathbf{D}_q, \mathbf{b}, \tilde{\mathbf{M}}, \mathbf{R}$ *(intra-cluster UL/DL)*;
  end
  $\mathbf{C}' \leftarrow \mathbf{C}'/\mathbf{C}_q$;
end
Output : $\mathbf{D}, \mathbf{b}^*, \tilde{\mathbf{M}}^*, \mathbf{R}^*$;

The rate computation is done as explained in Section IX-B.

Step 6 Rank decision: The ranks which maximise the metric in (7) are then allocated to a given pair of links in the rank group and not changed in succeeding rank searches. After going through all pairs within the rank group we then go to step 1, proceed to the next victim, update the $\mathbf{V}_{prev}$ and $\mathbf{S}_{prev}$.

The end-to-end algorithm comprising of cluster formation, intra-cluster coordination along with inter-cluster joint rate and rank adaptation is shown in Pseudo-code 1.

B. Throughput calculation

In the throughput calculation, we recompute the SINR using the channel realization in the current time slot which is corrupted by an independent noise with variance modeled using (4) based on scheduled links. The SINR can be degraded because of a) IRC filter mismatch to current actual channel realization and b) imperfect interference cancellation. Based on the SINR re-computation, we check whether the MCS and rank decided by the scheduler can now be supported in the current channel realization. If the achievable SINR is not sufficient for the MCS decision of the scheduler, we assume that zero throughput is achieved.

XI. Discussion on Complexity

The optimal resource allocation decisions can be found using a brute force search through all possible UL/DL user scheduling and MIMO rank combinations. Assuming a system with $|\mathbf{C}|$ coordinated cells and $\frac{|\mathbf{L}|}{|\mathbf{C}|}$ links in each cell, a brute force search complexity will result in a total complexity of:

$$\left(\frac{|\mathbf{L}|}{|\mathbf{C}|}M + 1\right)^{|\mathbf{C}|}, \tag{22}$$

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

where M is the maximum number of spatial streams which can be transmitted. Thus, even for small number of users per cell, the problem of full search through all possible user scheduling, muting and rank decisions incurs very high computational cost which is exponential in $|\mathbf{C}|$.

In our proposed solution we reduce the complexity in the scheduling search as well as the search for the optimal rank. The complexity of the scheduling search is reduced by dividing the set of all cells into clusters. With the maximum cluster size of Q cells the joint UL/DL user scheduling complexity becomes $(\frac{|\mathbf{L}|}{|\mathbf{C}|}+1)^Q$, which is now exponential in cluster size.

The MIMO rank search complexity is reduced by the use of SPARK algorithm, where the rank search is performed for pairs of links as described in section X-A. The complexity of intra-cluster joint UL/DL user scheduling and inter-cluster SPARK rank coordination is:

$$\underbrace{\Big(\frac{|\mathbf{L}|}{|\mathbf{C}|}+1\Big)^{Q}}_{\substack{\text{intra-cluster}\\ \text{scheduling options}}} \sum_{q=1}^{\frac{|\mathbf{C}|}{Q}} \underbrace{\frac{qQ}{2}}_{\substack{\text{SPARK}\\ \text{rank search}}} M^2 = \frac{|\mathbf{C}|}{Q}\Big(\frac{|\mathbf{L}|}{|\mathbf{C}|}+1\Big)^{Q}\frac{|\mathbf{C}|+Q}{4}M^2. \tag{23}$$

For example, consider a system with 15 cells and 4 users per each cell (possible 4 UL links and 4 DL in each cell). The number of receive and transmit antennas is M=4. The optimal UL/DL user scheduling and rank search complexity of the brute force approach in (22) will be 2.8624e+18 search options. The complexity of proposed scheme (23) with the cluster size of Q=3 cells is now 4.500e+3 options.

XII. Performance Evaluation

We study system performance of TDD schemes in an outdoor urban micro-cell scenario with multiple cells as in Fig. 4. Each base station (BS) and user equipment (UE) is equipped with four co-polarised omni-directional antennas. The deployment scenario consists of 64 users who are randomly dropped in a 600m x 700m area. We investigate three different base station (BS) deployment scenarios: 4x4 (16 cells), 3x3 (9 cells) and 2x3 (6 cells). For each of the BS deployment scenarios we use the same 600m x 700m area with the same position of the 64 users in the grid. The base stations in the same row or column are in the line of sight (LOS BS-to-BS interference), while for all other base stations we assume non-line of sight (NLOS BS-to-BS interference) as shown for 16 cells deployment scenario in Fig. 4.

TABLE III. Table of Simulation Parameter

Parameter	Default value
Carrier frequency	3.5 GHz
Pathloss model	3GPP LOS-NLOS probability based model [34]. Shadow fading standard deviation of 3dB, Ricean K-factor 9dB
Channel model	3GPP Spatial Channel model for MIMO [35]
UE noise figure	9 dB
DL transmit power	34 dBm (16 BSs), 37 dBm (9 BSs), 40 dBm (6 BSs)
UL transmit power	Uplink SNR target of 14 dB. Pathloss mode slow power control is applied. Maximum transmit power 23 dBm
Antenna configuration	4X4 MIMO. Omni-directional. Co-polarised antennas spaced 4λ at BS and 1λ at UE
Number of BSs	16,9,6
UE deployment	64 users dropped randomly in the scene. UE speed 3 Kmph
Number of cells in a cluster	Q=3
Traffic	Poisson with 1 sec inter-arrival time and DL/UL packet size ratio of 4:1. Fixed packet size of UL=2.1 Mbytes & DL=8.4 Mbytes (16 BSs scenario)

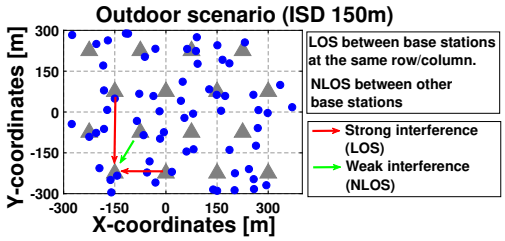

Fig. 4. Simulation scenario with 16 cells and 64 users.

A short description of simulation parameters is presented in Table III. For complexity reasons, the system bandwidth of 100 MHz is modeled into 16 bandwidth parts with the traffic equally split among them. Simulation is conducted in the 100 MHz bandwidth. The traffic is modeled independently for the uplink (UL) and the downlink (DL) according to a Poisson distribution per user with fixed file sizes and 1 second mean inter-arrival time. The uplink packet size is 2.1 Mbytes and downlink packet size is 8.4 Mbytes per user per 100 MHz bandwidth. To study the upper bound performance of the schemes, we assume resource block-specific precoding, rank selection and MCS selection. A channel outdating delay of 4 ms is assumed at the scheduler while channel estimation noise is modeled as in (4) for all the schemes based on instantaneous SINR. The following TDD schemes are evaluated and summarised in Table IV.

A. Evaluated schemes

1) Uncoordinated, per-cell flexible TDD scheduling: In this scheme each cell makes independent scheduling decisions taking into account the interference knowledge from the previous scheduling slot as in [6]. In the case of downlink the users of a small cell are assumed to feedback the accurate SINR per resource block based on the estimate from the previous scheduling slot. In the case of uplink, the base station makes the interference estimation. The interference knowledge from previous scheduling slot is then used to estimate the achievable SINR in a current transmission slot. Based on SINR estimation each cell makes the UL/DL user scheduling and MCS decisions to maximise its own performance without considering the impact to the neighbour cells.

IC-aware: The SINR calculation is based on interference cancellation receiver, where the interferers set and MCS knowledge is taken from the previous scheduling slot. Thus, each cell can perform its own IC-aware rank adaptation based on outdated interference knowledge. However, no rank or rate coordination is assumed between cells which can leads to suboptimal use of IC receivers especially in the case of multiple interferers.

2) Coordinated, fully-synchronised TDD scheduling: In this scheme a common UL/DL switching decision is made for all coordinated cells by the central entity. The transmission direction (uplink or downlink) is chosen dynamically on a scheduling slot basis (0.5 ms) and based on the metric of sum

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

TABLE IV. EVALUATED SCHEMES

SCHEME	**TDD switching**	**User scheduling decisions**	**Interference information**	**Interference scenario**
Uncoordinated, per-cell flexible TDD	Flexible per-cell	Standalone user scheduling rate and rank allocation	Knowledge about incoming interference from previous transmission slot.	UL-to-UL, DL-to-DL, UL-to-DL, DL-to-UL
Coordinated, fully-synchronised TDD	Synchronised across cells	Centralized joint user scheduling, coordinated rate and rank allocation	Knowledge about incoming and outgoing interference between cooperating cells at centralized scheduler.	UL-to-UL, DL-to-DL
Coordinated, per-cell flexible TDD w/o IC	Flexible per-cell	Centralized joint user scheduling, coordinated rate and rank allocation	Knowledge about incoming and outgoing interference between cooperating cells at centralized scheduler.	UL-to-UL, DL-to-DL, UL-to-DL, DL-to-UL
Coordinated, per-cell flexible TDD w/ IC	Flexible per-cell	IC-aware centralized joint user scheduling, coordinated rate and rank allocation	Knowledge about incoming and outgoing interference between cooperating cells at centralized scheduler.	UL-to-UL, DL-to-DL, UL-to-DL, DL-to-UL

of the packets delays in all the cells. Clusters are formed as in section VII by taking into account only the synchronised link direction for all the cells. Thus in this scheme, intra-cluster joint user scheduling and inter-cluster rate and rank allocation are performed for the metric in (7) only to mitigate UL-to-UL and DL-to-DL interference. To obtain a good upper bound performance of the fully-synchronised scheme, we always assume interference cancellation using SIC receiver as in Section IX-A. A link is chosen as a victim link for IC when SINR of any spatial stream is less than 0 dB. For the SPARK rank coordination algorithm (Section X-A) we set victim SINR threshold to 10 dB, thus also taking weak interferers into account for rank coordination.

3) Coordinated, per-cell flexible TDD: This is the proposed scheme where cell-specific UL/DL decision is made for each scheduling slot. Coordinated intra-cluster joint UL/DL selection and user scheduling is performed within the clusters. Joint MIMO rank and rate is also performed to optimise UL/DL switching as described in *Algorithm 1- Algorithm 4*.

Without IC: In this scheme cell clustering, intra-cluster scheduling and inter-cluster rank coordination algorithms are performed without utilising interference cancellation receivers. Linear IRC receivers are used for interference suppression and thus the step with rate estimation with IC is omitted.

The rank coordination in this scheme uses a brute force search after pre-selecting four strongest interferers per victim. We then choose the best performing set of ranks among all of the victims and its interferers. This scheme thus upper bounds the centralized scheme in [11] using full channel knowledge, and computing instantaneous SINR after equalization.

IC-aware: This is our proposed solution where joint resource allocation is performed based on inter-cell interference cancellation capabilities of the receivers (IC-aware scheduling). We use IC-aware clustering as in VII, and IC-aware intra-cluster scheduling as described in section VIII. The transmission parameters (MCS, MIMO rank) and muting decisions are made to benefit from IC-capable receivers as in IX-B and X-A. Rank coordination is performed with SPARK algorithm where SINR threshold for rank coordination is set to 10dB.

Genie aided upper bound: In this scheme, we model a genie aided upper bounded, where a genie provides all interference data to be cancelled via the backhaul for the uplink, while non ideal channel estimation is still modeled. Clustering algorithm VII, and intra-cluster scheduling in VIII, and SPARK rank coordination X-A are employed only for the downlink.

XIII. SIMULATION RESULTS

Simulations are conducted for the 64 user scenario in Fig. 4 for over 15 seconds. The E2E per user packet delay was measured as a difference of the packet arrival time at the transmitter and the time when the packet was fully delivered to the target receiver. We first discuss the performance for the dense 16 BS scenario. The cumulative distribution function (CDF) of the packet delays are plotted in Fig. 5. We then calculate the end to end per user throughput as the ratio between transmitted packet size and the end-to-end delay of that packet. The uplink and downlink throughput is calculated for each user and the numerical results are summarised in Table V. The average cell load in terms of resource utilisation is found to be around 40 % for the 16 BS scenario.

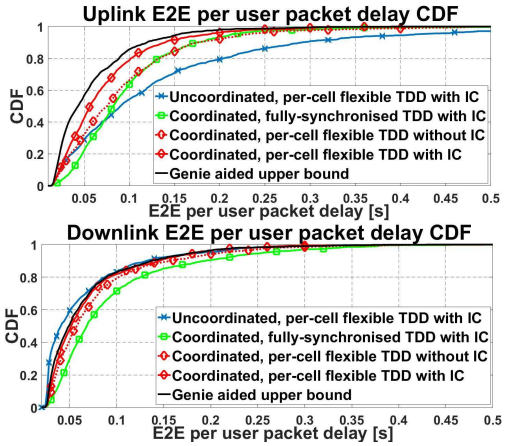

Fig. 5. Uplink and downlink E2E per user packet delay.

A. Performance of E2E per user throughput

1) Performance of per-cell flexible TDD: We evaluate the performance of the proposed coordinated per-cell flexible TDD scheme with IC and first compare it to uncoordinated per-cell flexible TDD scheduling with IC and then to coordinated fully-synchronised TDD scheduling with IC.

From Table V it can be seen that coordinated per-cell flexible TDD scheduling with IC achieves significant E2E uplink throughput gains as compared to the uncoordinated scheme with IC. The 5^{th} percentile E2E throughput in uplink is increased from 40.6 Mbps to 90.3 Mbps (222.4% gain) with the proposed coordination. The average uplink throughput gain is also significant with over 29% increase with coordinated scheduling (375.6 Mbps) as compared to the uncoordinated scheme (290.8 Mbps).

The 5^{th} percentile throughput corresponds to the users who experience low SINR for a certain time period because of interference. As a result of the metric in (7), the traffic of those users waits in transmission buffer until the SINR conditions improve. The uplink performance improvement is achieved thanks to inter-cell coordination where the user packet delay is used as a weighting factor in scheduling metric (7). However, the average downlink throughput reduces from 1667 Mbps in

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

TABLE V. E2E UPLINK AND DOWNLINK PER USER THROUGHPUT [MBPS] NUMERICAL RESULTS

SCHEME	Uplink [Mbps]		Downlink [Mbps]	
	$5^{th}\%$	Avg	$5^{th}\%$	Avg
Uncoordinated, per-cell flexible TDD scheduling without IC	**38.7**	**291**	**339.4**	**1614**
Uncoordinated, per-cell flexible TDD scheduling with IC	**40.6** (baseline)	**290.8** (baseline)	**353.7** (baseline)	**1667** (baseline)
Coordinated, fully-synchronised TDD scheduling with IC	**79.2** (95.1%)	**256.5** (-11.8%)	**259.5** (-26.6%)	**1092** (-34.5%)
Coordinated, per-cell flexible TDD scheduling w/o IC	**73.8** (81.8%)	**321.3** (10.5%)	**316.9** (-10.4%)	**1293** (-22.4%)
Coordinated, per-cell flexible TDD scheduling with IC	**90.3** (122.4%)	**375.6** (29.2%)	**365.2** (3.25%)	**1398** (-16.1%)
Genie aided upper bound	**109.6** (169.9%)	**475.8** (63.6%)	**366.5** (3.6%)	**1474** (-11.6%)

uncoordinated scheduling to 1398 Mbps in coordinated scheme (16% loss). This is because the proposed coordination scheme improves the achieved UL/DL fairness.

The upper bound results of IC are seen with the genie aided upper bound of coordinated per-cell flexible TDD scheduling. The results show that genie aided IC at the victim receiver can bring uplink gain of 169,9% and 63.6% in the 5^{th} percentile and average respectively as compared to uncoordinated scheduling. The genie-aided upper bound however still undergoes a downlink performance penalty to improve the UL/DL fairness using flexible UL/DL switching.

We now compare the performances of coordinated per-cell flexible TDD with IC and coordinated fully-synchronised TDD with IC. The results are shown in Table V. We see that coordinated per-cell flexible TDD scheme with IC significantly outperforms coordinated fully-synchronised TDD with IC in average E2E per user throughput, for both the uplink and the downlink. The average E2E per user throughput is increased from 256.5 Mbps to 375.6 Mbps in the uplink and from 1092 Mbps to 1398 Mbps in the downlink, which gives 46% gain and over 28% gain respectively. The gains are also significant in the 5^{th} percentile of E2E per user throughput results, where coordinated per-cell flexible TDD scheme shows over 14% gain in the uplink and 41% gain in the downlink as compared to fully-synchronised TDD scheme. This is because per-cell flexible TDD can better adapt to the instantaneous traffic demand of each cell and can also take advantage of strong interference cancellation. On the other hand fully-synchronised coordination avoids harmful cross-link interference and does not benefit from strong interference cancellation. Moreover, the requirement of synchronised UL/DL switching leads to the wastage of radio resources.

From the results we also observe that the coordinated fully-synchronised TDD is also suboptimal to the uncoordinated per-cell flexible TDD scheme in the average uplink throughput, with around 12% loss as compared to uncoordinated scheduling, However fully-synchronised TDD achieves satisfactory $5^{th}\%$ uplink throughput performance as compared to the uncoordinated per-cell flexible TDD scheme in this scenario.

2) Performance of interference cancellation: The performance gain of interference cancellation for coordinated per-cell flexible TDD is now evaluated. The results in Table V show that IC increases the uplink 5^{th} percentile throughput of coordinated per-cell flexible TDD scheme by around 22%.

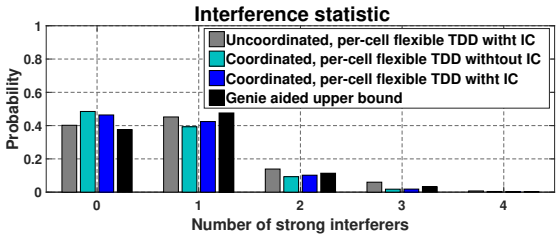

Fig. 6. Statistics of strong interference during uplink slots.

The uplink average throughput also improves by around 17% with IC. Furthermore, the gains are visible also in the downlink with 15% and 8% in the 5^{th} percentile and average respectively as compared to the scheme without IC. The downlink gains are partially realised because the downlink interferers are not highly penalized in terms of MIMO rank because of IC receivers in the uplink. The other reason for downlink gains is that the uplink is more spectral efficient in the case of IC. It thus needs less resources to be served which frees more resources to the downlink with flexible UL/DL switching.

It is also observed that in the case of uncoordinated per-cell flexible TDD scheduling the IC gains with SIC receivers without inter-cell coordination are limited. There is no visible gain in the average uplink throughput and around 5% gain in the $5^{th}\%$ with the scheme with IC.

Thus, without inter-cell coordination the SIC receivers are not able to effectively cancel the strong interferers. This is especially the case when there are multiple strong interferers. To utilise the capabilities of SIC receivers appropriate advanced coordination techniques are needed, as employed in the proposed coordinated per-cell flexible UL/DL scheduling with IC.

B. MIMO rank statistics

We now present the interference and MIMO rank statistics of UL/DL strong interference for the four per-cell flexible TDD schemes. Fig. 6 shows the statistics of downlink to uplink strong interference from neighbour LOS base station during the uplink transmissions.

From Fig. 6 it can be observed that in the case of per-cell flexible TDD uncoordinated scheduling, the uplink receiver sees a downlink interference almost 60% of the time. The coordinated per-cell flexible TDD with IC also realises a downlink interference almost 55% of the time. Moreover, one can observe that coordinated per-cell flexible TDD scheduling with IC sees more interference because the scheduler is IC-aware and strong interferers are scheduled on purpose so that they can be cancelled at the victim receiver. The statistics also show that for all of the schemes the uplink receiver sees two or more interferers around 20% of the time. We now further investigate the MIMO rank statistics of the strong interferers for different schemes. In Fig. 7 we show the distribution of the sum of number of MIMO spatial streams for uplink and its strongest downlink interferer. In the case of coordinated per-cell flexible TDD scheduling without IC, sum uplink and downlink MIMO rank less than 5 is chosen around 80% of time, which matches to the number of receive antennas. In the case of coordinated schemes with IC, the sum uplink and downlink MIMO rank higher than 4 is chosen around 53% of the time. Thus, our proposed coordinated IC-aware per-cell

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

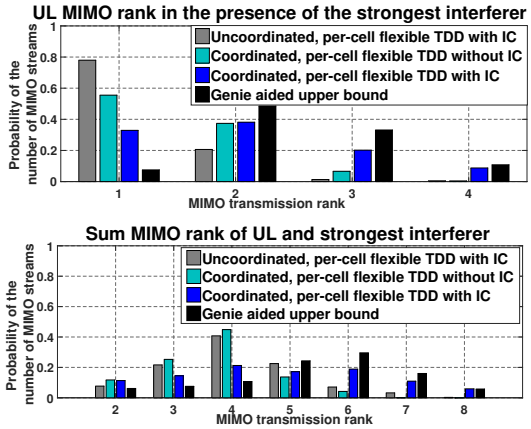

Fig. 7. MIMO rank statistics in the presence of the strongest interferer.

flexible TDD solution schedules strong cross-link interference more often while also employing a higher MIMO rank for the interference and the victim link. We observe that IC-aware schemes can benefit from high SINR on the cross-link and rate coordination for interference cancellation and thus realise a higher MIMO rank for interferer.

We now look at the uplink MIMO rank distribution when there the strongest interferer is present. It can be observed that the proposed coordinated per-cell flexible UL/DL scheduling with IC utilises single-user MIMO more effectively for the uplink with a 65% chance of having a rank 2 or more. Whereas in the case of the uncoordinated scheme only 25% of uplink transmission are done with more than one stream. Even in the presence of the strongest interference uplink can still use 4 MIMO streams around 8% of transmission and 3 streams in almost 20% of the time by using interference cancellation.

C. Clustering

In Table VI we compare the performance of per-cell flexible TDD scheduling with IC w.r.t. cell clustering schemes. We compare our proposed dynamic IC-aware clustering with clustering, where all clusters are formed before scheduling for the active cells. Thus in the comparative schemes, the scheduling and rank decisions are decided separately in series for each cluster, and the next cluster uses the decisions from previous cluster as side information. We evaluate the following three clustering schemes for the case of centralized scheduling with interference-cancellation.

Random clusters: Random cells are gathered to form a cluster.

K-means clustering: The clusters are formed such that the within-cluster sum of pathlosses between active BSs is minimized [17].

Affinity propagation (AP) clustering: where clusters are formed based on the pathlosses between the pairs of BSs [36].

The number of cells in clusters in random and k-means clustering is constant and equal to Q=3, whereas AP clustering the size of the clusters can be dynamic, however can not exceed Q=3.

From the results in Table VI it can be seen that our proposed dynamic IC-aware clustering achieves around 5%,21%,10% higher performance in the uplink 5^{th} percentile as compared to K-means, AP and random clustering respectively. The

TABLE VI. E2E UL AND DL PER USER THROUGHPUT [MBPS] W.R.T. CLUSTERING MODE

Clustering mode	Uplink [Mbps]		Downlink [Mbps]	
	$5^{\text{th}}\%$	Avg	$5^{\text{th}}\%$	Avg
Dynamic IC-aware clustering	90.3	375.6	365.2	1398
K-means clustering	85.9	344.1	332.7	1372
AP clustering	74.5	376.9	321.4	1333
Random-cell clustering	81.8	332.1	332.0	1381
Cluster-specific dynamic TDD	78.1	306.1	286.9	1048

average uplink performance and 5^{th} percentile downlink is also improved by around 10% as compared to K-means and random clustering. IC-aware clustering achieves better performance because of iterative cluster formation and inter-cluster rank coordination. AP-clustering undergoes a performance penalty because of variable cluster sizes.

We further compare our IC-aware clustering with flexible TDD with cluster-specific dynamic TDD switching recently proposed in [23]. In cluster-specific dynamic TDD switching we gather the sets of strongly interfering active cells (i.e. cells with LOS between BSs) into separate clusters. Based on the packets delays a common TDD direction is decided for all of the cells in a given cluster. Thus, strong cross-link interference between cells with LOS are avoided. However, separate clusters can have different UL/DL switching pattern, which can results in a weak cross-link interference between NLOS BSs. The results in Table VI show that our proposed IC-aware scheme with flexible TDD outperforms cluster-specific dynamic TDD mostly because of better adaptation to the instantaneous traffic demands of individual cells as well as effective utilisation of IC receivers in case of strong cross-link interference.

D. Cell densification scenarios

In this section we now study three different base stations densification scenarios. We consider the following base station deployments: 2 x 3 (6 BSs), 3 x 3 (9 BSs), 4 x 4 (16 BSs) covering the same area of 600m x 700m as in Fig.4. The BSs are deployed equally spaced, and thus cell radii are 180m, 141m, 106m respectively. To compensate for the different cell sizes we apply 3 base station transmit powers of 40 dBm, 37 dBm and 34 dBm respectively. The number of users (64) and their location is the same for all of the BS deployments. The densification scenarios have different characteristics as follows.

- The number of users per each BS is higher for the larger cells. This means there are around 11 users per base station in the 6 BS scenario. The same per user packet arrival rate is maintained with the result that traffic arrives more often per BS.
- The 5^{th} percentile and average received signal to noise ratio is lower for the larger cells even after applying higher transmit power because of higher NLOS probability from BS to UEs

To keep the offered traffic per base station roughly constant in the three scenarios, we scale the sizes of generated packets per user. For this purpose we use the 16 BS deployment scenario as baseline and apply the scaling below:

$$packet\_size(\eta) = \frac{\eta}{16} * packet\_size\_baseline. \quad (24)$$

Here η is the number of BSs in the deployment and $packet\_size(\eta)$ corresponds to the packet size in that deploy-

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TVT.2019.2893061, IEEE Transactions on Vehicular Technology

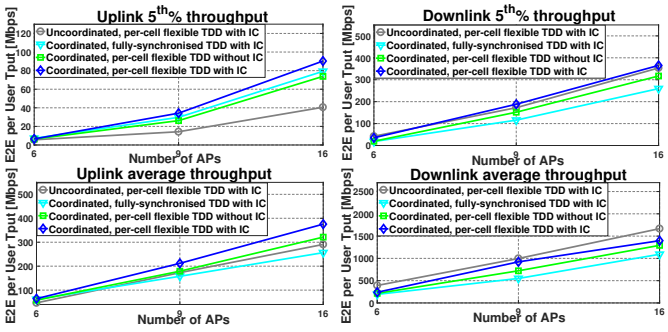

Fig. 8. E2E uplink and downlink per user throughput results

TABLE VII. TRANSMISSION STATISTICS

16 BS deployment scenario					
Scheduling scheme	Transmission statistics			Avg spectral [bits/s/Hz]	
	UL	DL	Muting	UL	DL
Uncoord. per-cell flex. TDD w/ IC	55.6%	44.4%	0%	4.1	16.6
Coord. fully-synch. TDD w/ IC	17.8%	38.8%	43.4%	9.2	17.8
Coord. per-cell flex. TDD w/o IC	43%	44.7%	12.3%	6.0	18.0
Coord. per-cell flex. TDD w/ IC	40.8%	48.6%	10.6%	6.6	17.8
6 BS deployment scenario					
Scheduling scheme	Transmission statistics			Avg spectral eff. [bits/s/Hz]	
	UL	DL	Muting	UL	DL
Uncoord. per-cell flex. TDD w/ IC	47%	53%	0%	2.0	7.1
Coord. fully-synch. TDD w/ IC	23%	54%	24.4%	2.8	7.2
Coord. per-cell flex. TDD w/o IC	43%	50%	11.9%	2.3	8.2
Coord. per-cell flex. TDD w/ IC	40%	53%	12.8%	2.6	8.0

ment. The DL packet sizes are thus 4.72 Mbytes for 9 BSs scenario and 3.15 Mbytes for 6 BSs scenario. The DL/UL packet size ratio is again 4:1. The resulting average resource utilisation (load) in case of uncoordinated scheduling for the 6 BS, 9 BS, and 16 BS scenarios is found to be 82 %, 57 % and 40 % respectively.

1) Throughput performance with cell densification: Fig. 8 shows the UL and DL E2E per user throughput results with respect to the number of BSs. Results show that increasing the number of BSs steadily increases the 5^{th} percentile and average downlink throughput for the per-cell flexible TDD schemes. This is because the SINR steadily improves with densification on account of higher LOS probability from BS to UEs. Increasing the number of BSs also improves uplink throughput for the per-cell flexible TDD schemes with coordination. The uncoordinated scheme achieves the best downlink performance at the cost of the uplink as we add more cells. Uncoordinated scheduling is seen to achieve satisfactory $5^{th}\%$ uplink throughput as compared to coordinated schemes in the 6 AP scenario. This is because the MIMO rank tends to be lower with larger cell and we also have less number of interferers. As a result, the uplink receiver is able to effectively mitigate the interference in the 6 BS case even without coordination.

2) IC performance: From Fig. 8 it can be also observed that the proposed coordinated per-cell flexible TDD scheduling with IC outperforms coordinated per-cell flexible TDD scheduling without IC in each of the BS densification scenarios. It is also seen to achieve better UL/DL fairness as compared to uncoordinated scheme without a significant penalty on the downlink $5^{th}\%$throughput. The IC gain in uplink $5^{th}\%$ throughput is 9%, 30%, 22% and in average throughput is 10%, 18%, 17% for 6, 9 and 16 BSs respectively.
In the case of larger cells we still see good uplink gains with the proposed IC scheme by improving the MIMO rank of cell center users. This in turn provides more resources for serving the $5^{th}\%$ users and thus reduces the worst case packet delays.

3) Performance gain of per-cell flexible TDD over fully-synchronised TDD: From Fig. 8 we see that coordinated per-cell flexible TDD with and without IC significantly outperforms fully-synchronised scheme in all the three densification scenarios. The uncoordinated scheme also outperforms fully-synchronised TDD scheme except for the uplink 5^{th} percentile throughput. The average throughput growth with densification is also limited for the fully-synchronised TDD scheme. The fully-synchronised scheme shows the lowest average uplink and downlink throughput increase w.r.t. the number of cells.

4) Transmission statistics: The UL/DL switching and muting results are presented in Table VII by considering only the slots with active traffic (inactive periods are filtered out). From the results, we make the following observations:

- We observe that coordinated schemes realise higher downlink spectral efficiency as compared to the uncoordinated scheme. Interestingly, we observe that coordinated per-cell flexible scheme achieves higher average downlink spectral efficiency as compared to the fully-synchronised scheme. This is because the downlink of coordinated per-cell flexible TDD may benefit from a lower level of inter-cell UE-to-UE interference instead of neighbour cell downlink.
- In terms of uplink spectral efficiency the fully-synchronised scheme outperforms the per-cell flexible schemes by avoiding cross-link interference. However, the higher spectral efficiency does not compensate for less dynamic resource utilisation (i.e. the sub-optimal UL/DL switching per-cell).

XIV. CONCLUSIONS

This paper discussed diverse RRM approaches for tackling strong interference in a flexible TDD system. A centralized per-cell flexible TDD solution which performs clustering, joint UL/DL user scheduling and coordinated MIMO rank and rate adaptation was presented. The coordinated scheme performs joint UL/DL scheduling within clusters by using interference-cancellation-aware flexible TDD switching. A low-complexity heuristic algorithm called SPARK is used for inter-cluster rank coordination. Performance evaluation was made in an outdoor urban micro scenario assuming many users, bursty traffic and realistic channel estimation. The system performance was studied in terms of end to end user throughput for three different cell densification scenarios. The results show that our proposed coordinated per-cell flexible TDD scheme with IC achieves the best user throughput improvement with densification. The results further show that fully-synchronised TDD which completely avoids cross-link interference is suboptimal for different cell sizes. Coordinated per-cell flexible TDD with or without IC outperforms fully-synchronised TDD scheme in both the uplink and the downlink. Uncoordinated per-cell flexible TDD also outperforms fully-synchronised TDD in average throughput.

We further find that advanced inter-cell coordination brings significant gains to SIC receivers. Our proposed multi-cell

coordinated scheduling improves $5^{th}\%$ E2E uplink per user throughput by around 120% as compared to fully-flexible standalone scheduling with IC.

REFERENCES

[1] ICT-317669 METIS/D2.1, "Requirement analysis and design approaches for 5G air interface," August 2013.

[2] N. Bhushan, J. Li, D. Malladi, R. Gilmore, "Network Densification: The Dominant Theme for Wireless Evolution into 5G," *IEEE Communications Magazine, vol.52, issue:2*, February 2014.

[3] 3GPP TS 36.420 V14.0.1, "E-UTRAN; X2 general aspects and principles", March 2017.

[4] 3GPP TS 36.300 V14.3.0, "E-UTRA and E-UTRAN; Overall description," June 2017.

[5] A.Lukowa, V.Venkatasubramanian, "On the Value of MIMO Rank Coordination for Interference Cancellation-based 5G Flexible TDD Systems," *IEEE CSCN 2016*, November 2016.

[6] V. Venkatasubramanian, M. Hesse, P. Marsch, M. Maternia, "On the performance gain of flexible UL/DL TDD with centralized and decentralized resource allocation in dense 5G deployments," *IEEE PIMRC 2014, pp.1840-1845*, September 2014.

[7] M. Ding, D. Lpez-Prez, R. Xue, A. V. Vasilakos, W. Chen, "On Dynamic Time-Division-Duplex Transmissions for Small-Cell Networks," *IEEE Transactions on Vehicular Technology, vol. 65, issue 11*, November 2016.

[8] A. Carleial, "A case where interference does not reduce capacity," *IEEE Trans. on Information Theory, vol.21, pp.569-570*, September 1975.

[9] S. Guo, X. Hou, H. Wang, "Dynamic TDD and Interference Management towards 5G," *IEEE WCNC 2018*, June 2018.

[10] F. M. L. Tavares, G. Berardinelli, D.Catania, T. B. Sorensen, P. Mogensen, "Managing inter-cell interference with advanced receivers and rank adaptation in 5G small cells," *European Wireless 2015*, July 2015.

[11] N.H. Mahmood, K.Pedersen, P. Mogensen, "Interference aware inter-cell rank coordination for 5G systems," *IEEE Access, vol.5*, Feb. 2017.

[12] N.H. Mahmood, L.G.U. Garcia, P. Popovski, P. Mogensen, "On the performance of successive interference cancellation in 5G small cell networks," *IEEE WCNC 2014*, April 2014.

[13] V. Fernandez-Lopez, K.I. Pedersen, B.Soret, J.Steiner, P.Mogensen, "Improving Dense Network Performance Through Centralized Scheduling and Interference Coordination," *IEEE Transactions on Vehicular Technology, vol. 66, issue 5*, May 2017.

[14] B. Natarajan "Coordinated Scheduling for Advanced UE Receivers using Belief Propagation," *IEEE VTC Spring 2015*, May 2015.

[15] V. Abdrashitov, W. Nam, D. Bai "Rate and UE Selection Algorithms for Interference-Aware Receivers," *IEEE VTC Spring 2014*, May 2014.

[16] A.Lukowa, V.Venkatasubramanian, "Performance of strong interference cancellation in flexible UL/DL TDD systems using coordinated muting, scheduling and rate allocation," *IEEE WCNC 2016*, April 2016.

[17] K. Hosseini, W.Yu, R.S.Adve, "Cluster Based Coordinated Beamforming and Power Allocation for MIMO Heterogeneous Networks,"*13th Canadian Workshop on Information Theory*, October 2013.

[18] J. Gong, S.Zhou, Z.Niu, L.Geng, M.Zheng, "Joint Scheduling and Dynamic Clustering in Downlink Cellular Networks,"*IEEE Globecom 2011*, January 2012.

[19] W. Li, W. Zheng, X.Wen, T.Su, "Dynamic Clustering based Sub-band Allocation in Dense Femtocell Environments,"*IEEE VTC Spring 2012*, July 2012.

[20] A. Abdelnasser, E.Hossain, D.I.Kim, "Clustering and Resource Allocation for Dense Femtocells in a Two-Tier Cellular OFDMA Network,"*IEEE Transactions on Wireless Communications, vol. 13, issue 3*, January 2014.

[21] A. Papadogiannis, D.Gesbert, E. Hardouin, "A Dynamic Clustering Approach in Wireless Networks with Multi-Cell Cooperative Processing," *IEEE ICC 2008*, May 2008.

[22] M.M.U. Rahman, H. Ghauch, S. Imtiaz, J. Gross, "RRH clustering and transmit precoding for interference-limited 5G CRAN downlink," *IEEE Globecom Workshops 2015*, February 2016.

[23] J.Rached, R.Nasri, L. Decreusefond, "Interference Analysis in Dynamic TDD System Combined or not With Cell Clustering Scheme," *IEEE VTC Spring 2018*. June 2018.

[24] P. Mogensen et al., "5G small cell optimized radio design," *IEEE Globecom 2013 Workshops*, December 2013.

[25] R1-165027, 3GPP TSG-RAN WG 1, "Basic frame structure principles for 5G," May 2016.

[26] J.C. Ikuno, S. Pendl, M. Simko, M. Rupp, "Accurate SINR estimation model for system level simulation of LTE network," *IEEE ICC 2012*, June 2012.

[27] M. Andrews et al., "Providing Quality of Service over a Shared Wireless Link," *IEEE Communications Magazine*, February 2001.

[28] M. Andrews, K.Kumaran, K.Ramanan, A. Stolyar, "Scheduling in a Queueing System with Asynchronously Varying Service Rates," *Probability in the Engineering and Informational Sciences*, pp. 191217, April 2004.

[29] K. Kim, I. Koo, S. Sung, K.Kim, "Multiple QoS Support using M-LWDF in OFDMA Adaptive Resource Allocation," *IEEE LANMAN 2004*, October 2004.

[30] M. Andrews, L. Zhang, "Scheduling Algorithms for Multicarrier Wireless Data Systems," *IEEE/ACM Transactions on Networking, vol.19, issue 2*, April 2011.

[31] 3GPP TS 36.213 V14.3.1, "E-UTRA; Physical layer procedures," June 2017.

[32] F. Tavares, G. Berardinelli, N. Mahmood, T. Sorensen, P. Mogensen, "On the potential of interference rejection combining in B4G networks," *IEEE VTC Fall 2013*, September 2013.

[33] Y. Ohwatari, N. Miki, T. Asai, T. Abe, H. Taoka, "Performance of Advanced Receiver Employing Interference Rejection Combining to Suppress Inter-cell Interference in LTE-Advanced Downlink," *IEEE VTC Fall 2011*, September 2011.

[34] 3GPP TR 36.814 V9.2.0 ,"Further advancements for E-UTRA physical layer aspects," March 2017.

[35] 3GPP TR 38.901 V14.0.1, "Study on channel model for frequencies from 0.5 to 100 GHz," July 2017.

[36] B. J. Frey, D. Dueck, "Clustering by passing messages between data points," *Science, vol. 315, no. 5814*, 2007.

Anna Łukowa received her Bachelor of Science degree and Master of Science degree from Information and Communication Technology from Poznań University of Technology (PUT), Poland in 2013 and 2014 respectively. Since 2014, Anna has been a PhD student at Nokia, Poland and PUT. Her research interests are in radio resource management, interference coordination with advanced receivers, modeling and performance analysis of 5G systems. She is currently a L1 Algorithm Architect at Nokia, Poland.

Venkatkumar Venkatasubramanian is currently employed as Senior Systems Engineer in Nokia, Poland. He has been a Senior Research Engineer with Nokia, Poland since January 2013 and has represented Nokia in various EU projects while at Nokia-Bell Labs. Prior to joining Nokia, Venkatkumar was a research associate at Fraunhofer Heinrich Hertz Institute in Berlin, specializing in MIMO-OFDM and relay wireless systems. Venkatkumar completed his PhD in 2011 from Victoria University in Melbourne, Australia. His research interests are in 5G Physical layer, interference coordination and radio resource management of 5G cellular systems.