Dynamic UL/DL Mode Selection and Resource Allocation in Multi-Cell MIMO TDD Systems

Antti Arvola, *Student Member, IEEE,* Satya Joshi, *Member, IEEE,* and Antti Tölli, *Senior Member, IEEE*

Abstract—This paper considers joint uplink/downlink mode selection and resource allocation problem in multi-cell systems for flexible TDD frame structures. To combat the difficult interference characteristics of such systems, a dynamic mode selection and resource allocation scheme is derived using Lyapunov optimization. The result is a cross-layer optimization scheme, where a long-term time-average utility function of flow rates and interference penalty is optimized by solving a series of instantaneous subproblems. The proposed dynamic algorithm provides a solution for the UL/DL mode selection and resource allocation, guided by the accumulating network layer queues and virtual queues. Furthermore, a heuristic method is considered to decouple the difficult power constraint in the resource allocation subproblem.

I. Introduction

Small cell systems and dynamic time division duplexing (TDD) have been gaining traction as possible solutions to the high data rate and low latency demands of the 5G technologies. They have also been proposed as important research items in the 3GPP New Radio standardization [1]. As the instantaneous traffic demand between uplink (UL) and downlink (DL) can vary greatly among the cells in small cell systems, dynamic TDD with flexible frame structures is an obvious solution to provide low latency services with high data rates. However, varying the cells transmit UL and DL modes can induce difficult interference characteristics, for example UL-DL cross-link interference (CLI), where a user in an uplink cell is interfering with a downlink user in the adjacent cell [2].

Previous research in single-input single-output (SISO) dynamic TDD CLI mitigation via mode selection has mostly focused on various time-slot allocation algorithms, see [3]–[5] and the references therein. Cell mode selection problems have also been a focus in device-to-device communications, for example [6], [7]. However, the single-antenna schemes lack the interference mitigation benefits of beamforming.

The literature in beamforming-based interference mitigation is plentiful. Considering the interference characteristics of TDD systems with fixed UL and DL transmit modes of the cells, the cross-link power leakage can be mitigated for example via beamformer training schemes, [8], [9], where precoded pilots are applied to iteratively train optimal beamformers. This approach along with the UL/DL mode selection is considered in [10]; however, the approach proposed in [10] is myopic and does not consider the long-term statistics of the dynamic network.

We model the joint UL/DL mode selection and resource allocation problem as a time-dynamic optimization problem.

We adopt the Lyapunov drift-plus-penalty framework to derive a dynamic control algorithm, which solves a series of instantaneous subproblems, providing a solution to the original long-term time-dynamic problem [11]. The proposed dynamic algorithm decouples the UL/DL mode selection and the resource allocation, guided by the accumulating network layer and virtual queues. Furthermore, we propose a heuristic method to decouple the power allocation and beamformers in the resource allocation subproblem.

II. System Model and Problem Formulation

We consider a dynamic TDD multiple-input multiple-output (MIMO) system, where B half-duplex M-antenna base stations (BSs) serve U users (UEs) with N-antennas each. We denote the sets of all BSs and users by $\mathcal{B} = \{1, \dots, B\}$ and $\mathcal{U} = \{1, \dots, U\}$, respectively, and the set of data streams of user u by $\mathcal{L}_u = \{1, \dots, L_u\}$. We denote the BS serving user u by b_u, and the set of users served by BS b by $\mathcal{U}_b$.

Let $\mathbf{m}_{ul} \in \mathbb{C}^{M \times 1}$ and $\underline{\mathbf{m}}_{ul} \in \mathbb{C}^{N \times 1}$ denote the precoders associated with lth data stream of user u for DL and UL transmissions, respectively. Then the DL signal received by uth user at time t can be expressed as

$$\mathbf{y}_u(t) = \mathbf{H}_{b_u u}(t)\Big[\mathbf{m}_{ul}(t)d_{ul}(t) + \sum_{j \in \mathcal{L}_u \setminus \{l\}}\mathbf{m}_{uj}(t)d_{uj}(t)\Big] + \sum_{k \in \mathcal{U} \setminus \{u\}}\mathbf{H}_{b_k u}(t)$$
$$\times \sum_{j \in \mathcal{L}_k}\mathbf{m}_{kj}(t)d_{kj}(t) + \sum_{k \in \mathcal{U} \setminus \mathcal{U}_{b_u}}\tilde{\mathbf{H}}_{ku}(t)\sum_{j \in \mathcal{L}_k}\underline{\mathbf{m}}_{kj}(t)\underline{d}_{kj}(t) + \mathbf{n}_u(t),$$

where $\mathbf{H}_{b_u u}(t) \in \mathbb{C}^{N \times M}$ is the channel matrix between BS b_u and user u, $\tilde{\mathbf{H}}_{ku}(t) \in \mathbb{C}^{N \times N}$ is the channel matrix between users k and u, and $\mathbf{n}_{ul}(t)$ is complex Gaussian noise with covariance $N_0\mathbf{I}$, notations $d_{ul}(t)$ and $\underline{d}_{ul}(t)$ denote DL and UL data symbols associated with lth stream of user u, respectively. We assume that data symbols are normalized, such that $\mathbb{E}[|d_{ul}(t)|^2] = \mathbb{E}[|\underline{d}_{ul}(t)|^2] = 1, \forall u, l$.

Correspondingly, the UL received signal at BS b_u at time t can be expressed as

$$\underline{\mathbf{y}}_{b_u}(t) = \mathbf{H}_{b_u u}^{\mathrm{H}}(t)\Big[\underline{\mathbf{m}}_{ul}(t)\underline{d}_{ul}(t) + \sum_{j \in \mathcal{L}_u \setminus \{l\}}\underline{\mathbf{m}}_{uj}(t)\underline{d}_{uj}(t)\Big] + \sum_{k \in \mathcal{U} \setminus \{u\}}\mathbf{H}_{b_u k}^{\mathrm{H}}(t)$$
$$\times \sum_{j \in \mathcal{L}_k}\underline{\mathbf{m}}_{kj}(t)\underline{d}_{kj}(t) + \sum_{k \in \mathcal{U} \setminus \mathcal{U}_{b_u}}\hat{\mathbf{H}}_{b_k b_u}(t)\sum_{j \in \mathcal{L}_k}\mathbf{m}_{kj}(t)d_{kj}(t) + \underline{\mathbf{n}}_{b_u}(t),$$

where $\hat{\mathbf{H}}_{b_k b_u}(t) \in \mathbb{C}^{M \times M}$ is the channel matrix between BSs b_k and b_u, and $\mathbf{n}_{b_u}(t)$ is complex Gaussian noise with covariance $N_0\mathbf{I}$.

We assume minimum mean-square error (MMSE) estimation at the receiver [12]. Thus, the DL and UL symbols $d_{ul}(t)$ and $\underline{d}_{ul}(t)$ can be estimated as

$$\hat{d}_{ul}(t) = \mathbf{w}_{ul}^{\mathrm{H}}(t)\mathbf{y}_{ul}(t), \qquad \hat{\underline{d}}_{ul}(t) = \underline{\mathbf{w}}_{ul}^{\mathrm{H}}(t)\underline{\mathbf{y}}_{b_u}(t),$$

where $\mathbf{w}_{ul}(t) \in \mathbb{C}^{N \times 1}$ and $\underline{\mathbf{w}}_{ul}(t) \in \mathbb{C}^{M \times 1}$ are linear MMSE combiners associated with lth data stream of user u

$$\gamma_{ul}(t) = \frac{|\mathbf{w}_{ul}^{\mathrm{H}}(t)\mathbf{H}_{b_u u}(t)\mathbf{m}_{ul}(t)|^2}{\tilde{N}_0 + \sum\limits_{j \in \mathcal{L}_u \backslash \{l\}}|\mathbf{w}_{ul}^{\mathrm{H}}(t)\mathbf{H}_{b_u u}(t)\mathbf{m}_{uj}(t)|^2 + \sum\limits_{k \in \mathcal{U} \backslash \{u\}}\sum\limits_{j \in \mathcal{L}_k}|\mathbf{w}_{ul}^{\mathrm{H}}(t)\mathbf{H}_{b_k u}(t)\mathbf{m}_{kj}(t)|^2 + \sum\limits_{k \in \mathcal{U} \backslash \mathcal{U}_{b_u}}\sum\limits_{j \in \mathcal{L}_k}|\mathbf{w}_{ul}^{\mathrm{H}}(t)\tilde{\mathbf{H}}_{ku}(t)\underline{\mathbf{m}}_{kj}(t)|^2} \quad (1)$$

$$\underline{\gamma}_{ul}(t) = \frac{|\underline{\mathbf{w}}_{ul}^{\mathrm{H}}(t)\mathbf{H}_{b_u u}^{\mathrm{H}}(t)\underline{\mathbf{m}}_{ul}(t)|^2}{\underline{\tilde{N}}_0 + \sum\limits_{j \in \mathcal{L}_u \backslash \{l\}}|\underline{\mathbf{w}}_{ul}^{\mathrm{H}}(t)\mathbf{H}_{b_u u}^{\mathrm{H}}(t)\underline{\mathbf{m}}_{uj}(t)|^2 + \sum\limits_{k \in \mathcal{U}^{\mathrm{UL}} \backslash \{u\}}\sum\limits_{j \in \mathcal{L}_k}|\underline{\mathbf{w}}_{ul}^{\mathrm{H}}(t)\mathbf{H}_{b_u k}^{\mathrm{H}}(t)\underline{\mathbf{m}}_{kj}(t)|^2 + \sum\limits_{k \in \mathcal{U} \backslash \mathcal{U}_{b_u}}\sum\limits_{j \in \mathcal{L}_k}|\underline{\mathbf{w}}_{ul}^{\mathrm{H}}(t)\hat{\mathbf{H}}_{b_k b_u}(t)\mathbf{m}_{kj}(t)|^2} \quad (2)$$

in DL and UL transmissions, respectively. Thus, the stream-wise signal-to-interference-plus-noise ratio (SINR) on DL and UL transmissions can be expressed as in (1) and (2), where $\tilde{N}_0 = \mathbb{E}[|\mathbf{w}_{ul}^{\mathrm{H}}(t)\mathbf{n}_{ul}|^2]$ and $\underline{\tilde{N}}_0 = \mathbb{E}[|\underline{\mathbf{w}}_{ul}^{\mathrm{H}}(t)\underline{\mathbf{n}}_{b_u}|^2]$.

In this paper, we assume Gaussian signaling and stream-wise detection. Thus, the receiver treats all the interfering signals as additional noise. Therefore, the DL rate $r_u(t)$ and UL rate $\underline{r}_u(t)$ for user u at time t are assumed to be

$$\begin{aligned} r_u(t) &= \sum_{l \in \mathcal{L}_u} \log_2(1 + \gamma_{ul}(t)), \\ \underline{r}_u(t) &= \sum_{l \in \mathcal{L}_u} \log_2(1 + \underline{\gamma}_{ul}(t)). \end{aligned} \quad (3)$$

A. Network Queueing

We consider a network-queue model similar to [11], where the exogenously arriving data of user u at BS b_u is first admitted to a transport layer queue $D_u(t)$, then admitted to network layer queue $Q_u(t)$ according to flow control decisions made in each time slot. The dynamics of the transport layer queue $D_u(t)$ are given by

$$D_u(t+1) = \big[D_u(t) - O_b(t)a_u\big]^+ + \lambda_u(t), b \in \mathcal{B}, u \in \mathcal{U}_b, \quad (4)$$

where $a_u = A^{\max}$ is a fixed-size packet[1] outbound from the transport layer queue $D_u(t)$ of user u, $O_b(t)$ is a binary DL flow control variable for all users $u \in \mathcal{U}_b$ associated with BS b, and $\lambda_u(t)$ is the exogenously arrival rate of user u (which depends on application demand, e.g., speech, video, etc.). From the transport layer queue, the data is admitted to the network layer queue $Q_u(t)$, which updates according to

$$Q_u(t+1) = \big[Q_u(t) - r_u(t)\big]^+ + O_b(t)a_u, b \in \mathcal{B}, u \in \mathcal{U}_b, \quad (5)$$

where $r_u(t)$ is the transmission rate, defined in (3).

Similarly, at each user $u \in \mathcal{U}$, the dynamics of transport and network layer queues $\underline{D}_u(t)$ and $\underline{Q}_u(t)$ for UL data traffic can be expressed as

$$\underline{D}_u(t+1) = \big[\underline{D}_u(t) - \underline{O}_b(t)\underline{a}_u\big]^+ + \underline{\lambda}_u(t), b \in \mathcal{B}, u \in \mathcal{U}_b, \quad (6)$$

$$\underline{Q}_u(t+1) = \big[\underline{Q}_u(t) - \underline{r}_u(t)\big]^+ + \underline{O}_b(t)\underline{a}_u, b \in \mathcal{B}, u \in \mathcal{U}_b, \quad (7)$$

where $\underline{\lambda}_u(t)$ is exogenous arrival, $\underline{a}_u = \underline{A}^{\max}$ is the fixed-size data packet, and $\underline{O}_b(t)$ is a binary flow control variable in UL for all users associated with BS b.

B. Problem Formulation

Our objective is to dynamically assign the UL/DL transmission mode of the BSs (i.e., cells) based on the traffic demand, while simultaneously maximizing the utility of the dynamic TDD system. Under network stability, admitted data $O_b(t)a_u$ for all t in the internal queue is transmitted to the corresponding DL user over a finite period of time [11].

---

[1] The outbound flow is fixed unless the transport layer queue contains less data than the fixed value. Then $a_u = \min(D_u(t), A^{\max})$.

(similarly, in the case of UL transmission). Thus, we define a network utility function as

$$\phi(\overline{\mathbf{O}}, \underline{\overline{\mathbf{O}}}) = \sum_{b \in \mathcal{B}} \sum_{u \in \mathcal{U}_b} \big(w_u \log(1 + \overline{O}_b a_u) + \underline{w}_u \log(1 + \underline{\overline{O}}_b \underline{a}_u)\big), \quad (8)$$

where $w_u > 0$ and $\underline{w}_u > 0$ are the user-specific priority weights, $\overline{O}_b = \lim_{t \to \infty} \frac{1}{t}\sum_{\tau=1}^t O_b(\tau)$, and $\underline{\overline{O}}_b = \lim_{t \to \infty} \frac{1}{t}\sum_{\tau=1}^t \underline{O}_b(\tau)$.

Furthermore, we want to mitigate the complicated interference scenario of the system through UL/DL mode selection. To do this, let $\alpha_b(t)$ and $\underline{\alpha}_b(t)$ denote aggregated interference projected by BS b (in DL transmission) and the users associated with BS b (in UL transmission), respectively. We assume parameters $\alpha_b(t)$ and $\underline{\alpha}_b(t)$ can be obtained via network feedback. Let $\overline{O_b \alpha_b} = \lim_{t \to \infty} \frac{1}{t}\sum_{\tau=1}^t O_b(\tau)\alpha_b(\tau)$ and $\overline{\underline{O}_b \underline{\alpha}_b} = \lim_{t \to \infty} \frac{1}{t}\sum_{\tau=1}^t \underline{O}_b(\tau)\underline{\alpha}_b(\tau)$. Then, the long-term projected interference penalty can be expressed as

$$f(\overline{\mathbf{O}}, \underline{\overline{\mathbf{O}}}) = \sum_{b \in \mathcal{B}}(\overline{O_b \alpha_b} + \overline{\underline{O}_b \underline{\alpha}_b}). \quad (9)$$

Furthermore, let $\overline{r}_u = \lim_{t \to \infty} \frac{1}{t}\sum_{\tau=1}^t r_u(\tau)$ and $\underline{\overline{r}}_u = \lim_{t \to \infty} \frac{1}{t}\sum_{\tau=1}^t \underline{r}_u(\tau)$. Then, by using expressions (8) and (9), the dynamic UL/DL mode selection and resource allocation problem can be formulated as

maximize $\phi(\overline{\mathbf{O}}, \underline{\overline{\mathbf{O}}}) - \delta f(\overline{\mathbf{O}}, \underline{\overline{\mathbf{O}}})$

subject to $\overline{O}_{b_u} a_u \le \overline{r}_u, \quad u \in \mathcal{U}$ (10a)

$\quad\quad\quad \underline{\overline{O}}_{b_u} \underline{a}_u \le \underline{\overline{r}}_u, \quad u \in \mathcal{U}$ (10b)

$\quad\quad\quad \sum\limits_{u \in \mathcal{U}_b}\sum\limits_{l \in \mathcal{L}_u} \|\mathbf{m}_{ul}(t)\|^2 \le x_b(t)P_{\max}, b \in \mathcal{B}, \forall t$ (10c)

$\quad\quad\quad \sum\limits_{l \in \mathcal{L}_u} \|\underline{\mathbf{m}}_{ul}(t)\|^2 \le \underline{x}_b(t)P_{\max}^{\mathrm{UE}}, b \in \mathcal{B}, u \in \mathcal{U}, \forall t$ (10d)

$\quad\quad\quad \overline{x}_b = \overline{O}_b, \quad b \in \mathcal{B}$ (10e)

$\quad\quad\quad \underline{\overline{x}}_b = \underline{\overline{O}}_b, \quad b \in \mathcal{B}$ (10f)

$\quad\quad\quad O_b(t), \underline{O}_b(t) \in \{0, 1\}, \quad b \in \mathcal{B}, \forall t$ (10g)

$\quad\quad\quad O_b(t) + \underline{O}_b(t) = 1, \quad b \in \mathcal{B}, \forall t$ (10h)

$\quad\quad\quad x_b(t), \underline{x}_b(t) \in \{0, 1\}, \quad b \in \mathcal{B}, \forall t$ (10i)

$\quad\quad\quad x_b(t) + \underline{x}_b(t) = 1, \quad b \in \mathcal{B}, \forall t,$ (10j)

with variables $\{O_b(t), \underline{O}_b(t), x_b(t), \underline{x}_b(t)\}_{b \in \mathcal{B}}$ and $\{\mathbf{m}_{ul}(t), \underline{\mathbf{m}}_{ul}(t), \mathbf{w}_{ul}(t), \underline{\mathbf{w}}_{ul}(t)\}_{u \in \mathcal{U}, l \in \mathcal{L}_u}$ for all $t \in \{1, 2, \ldots\}$; where $\delta \ge 0$ is a tradeoff parameter. The constraints (10a)-(10b) ensure that the network is stable [11]. The constraints (10c) and (10d) limit the total transmit power of each BS and user, respectively. Note that the binary variables $x_b(t)$ and $\underline{x}_b(t)$ set *zero* or *maximum* power for UL/DL transmissions in constraints (10c) and (10d). We impose long-term consensus between variables $\{x_b(t), \underline{x}_b(t)\}$ and $\{O_b(t), \underline{O}_b(t)\}$ by the constraints (10e)-(10f) in order for the UL/DL mode selection to be guided by the higher layer of the network.

III. Dynamic Algorithm via Lyapunov Optimization

To solve the problem of time averages (10), we utilize Lyapunov optimization, namely the drift-plus-penalty method [11]. We modify problem (10) such that it conforms to the structure required for the drift-plus-penalty method, as follows:

$$\begin{array}{ll} \text{maximize} & \overline{\phi(\mathbf{O}(t), \underline{\mathbf{O}}(t))} - \delta f(\overline{\mathbf{O}}, \overline{\underline{\mathbf{O}}}) \\ \text{subject to} & \text{constraints } (10a) - (10j), \end{array} \quad (11)$$

with variables $\{O_b(t), \underline{O}_b(t), x_b(t), \underline{x}_b(t)\}_{b \in \mathcal{B}}$ and $\{\mathbf{m}_{ul}(t), \underline{\mathbf{m}}_{ul}(t), \mathbf{w}_{ul}(t), \underline{\mathbf{w}}_{ul}(t)\}_{u \in \mathcal{U}, l \in \mathcal{L}_u}$ for all $t \in \{1, 2, \ldots\}$; where

$$\overline{\phi(\mathbf{O}(t), \underline{\mathbf{O}}(t))} \triangleq \lim_{t \to \infty} \tfrac{1}{t} \textstyle\sum_{\tau=1}^{t} \sum_{b \in \mathcal{B}} \sum_{u \in \mathcal{U}_b} \big[w_u \log(1 + O_b(\tau)a_u) + \underline{w}_u \log(1 + \underline{O}_b(\tau)\underline{a}_u) \big].$$

Note that by using Jensen's inequality, we can easily verify that $\overline{\phi(\mathbf{O}(t), \underline{\mathbf{O}}(t))}$ is a lower bound on $\phi(\overline{\mathbf{O}}, \overline{\underline{\mathbf{O}}})$. Thus, by solving problem (11), we can find a solution for (10).

We impose the time-average inequality constraints (10a)-(10b) and equality constraints (10e)-(10f) by transforming them into a queue stability problem [11]. By following the steps in [11, Ch. 5], we can show that constraints (10a) and (10b) can be characterized by the queue dynamics (5) and (7), respectively. Furthermore, let $\{U_b(t)\}_{b \in \mathcal{B}}$ and $\{\underline{U}_b(t)\}_{b \in \mathcal{B}}$ be virtual queues associated with constraints (10e) and (10f), respectively, and updated according to

$$\begin{aligned} U_b(t+1) &= U_b(t) - O_b(t) + x_b(t), \quad b \in \mathcal{B} \\ \underline{U}_b(t+1) &= \underline{U}_b(t) - \underline{O}_b(t) + \underline{x}_b(t), \quad b \in \mathcal{B} \end{aligned} \quad (12)$$

We now define the Lyapunov function as

$$L(\Theta(t)) = \tfrac{1}{2} \Big[\sum_{u \in \mathcal{U}} \big(Q_u(t)^2 + \underline{Q}_u(t)^2 \big) + \sum_{b \in \mathcal{B}} \big(U_b(t)^2 + \underline{U}_b(t)^2 \big) \Big] \quad (13)$$

where $\Theta(t) = [Q_1(t) \ldots Q_{|\mathcal{U}|}(t), \underline{Q}_1(t) \ldots \underline{Q}_{|\mathcal{U}|}(t), U_1(t) \ldots U_{|\mathcal{B}|}(t), \underline{U}_1(t) \ldots \underline{U}_{|\mathcal{B}|}(t)]$. Moving forward, we define the Lyapunov drift, which describes the change in network congestion between time slots as [11]

$$\Delta(\Theta(t)) = \mathbb{E}[L(\Theta(t+1)) - L(\Theta(t)) | \Theta(t)], \quad (14)$$

where the expectation is taken with respect to the admission policy and random channel states. Let

$$\begin{aligned} \phi(\mathbf{O}(t), \underline{\mathbf{O}}(t)) &= \textstyle\sum_{b \in \mathcal{B}} \big(\sum_{u \in \mathcal{U}_b} w_u \log(1 + O_b(t)a_u) \\ &\quad + \underline{w}_u \log(1 + \underline{O}_b(t)\underline{a}_u) \big) \\ f(\mathbf{O}(t), \underline{\mathbf{O}}(t)) &= \textstyle\sum_{b \in \mathcal{B}} \big(O_b(t)\alpha_b(t) + \underline{O}_b(t)\underline{\alpha}_b(t) \big) \end{aligned}$$

Then we can find a solution for problem (11), by minimizing, for each time t, the following drift-plus-penalty function:

$$\Delta(\Theta(t)) - C \mathbb{E}\big[\phi(\mathbf{O}(t), \underline{\mathbf{O}}(t)) - \delta f(\mathbf{O}(t), \underline{\mathbf{O}}(t)) \big| \Theta(t) \big], \quad (15)$$

subject to constraints

$$\begin{aligned} &\textstyle\sum_{u \in \mathcal{U}_b} \sum_{l \in \mathcal{L}_u} \|\mathbf{m}_{ul}(t)\|^2 \le x_b(t) P_{\max}, \quad b \in \mathcal{B} \quad &(16a) \\ &\textstyle\sum_{l \in \mathcal{L}_u} \|\underline{\mathbf{m}}_{ul}(t)\|^2 \le \underline{x}_b(t) P_{\max}^{\text{UE}}, \quad b \in \mathcal{B}, u \in \mathcal{U} \quad &(16b) \\ &O_b(t), \underline{O}_b(t) \in \{0,1\}, \quad b \in \mathcal{B} \quad &(16c) \\ &O_b(t) + \underline{O}_b(t) = 1, \quad b \in \mathcal{B} \quad &(16d) \\ &x_b(t), \underline{x}_b(t) \in \{0,1\}, \quad b \in \mathcal{B} \quad &(16e) \\ &x_b(t) + \underline{x}_b(t) = 1, \quad b \in \mathcal{B}. \quad &(16f) \end{aligned}$$

Note that in (15), $C \ge 0$ is a design variable that facilitates a tradeoff between minimizing the drift and maximizing the objective of problem (11). To expand on the drift-plus-penalty expression, let us first derive the Lyapunov drift. Substituting (13) to (14), and using the queue dynamic expressions (5), (7) and (12), the Lyapunov drift becomes

$$\begin{aligned} \Delta(\Theta(t)) &= \mathbb{E}[L(\Theta(t+1)) - L(\Theta(t)) | \Theta(t)] \\ &\le G + \textstyle\sum_{u \in \mathcal{U}} Q_u(t) \mathbb{E}[(O_b(t)a_u) - r_u(t) | \Theta(t)] \\ &\quad + \textstyle\sum_{u \in \mathcal{U}} \underline{Q}_u(t) \mathbb{E}[\underline{O}_b(t)\underline{a}_u - \underline{r}_u(t) | \Theta(t)] \\ &\quad + \textstyle\sum_{b \in \mathcal{B}} U_b(t) \mathbb{E}[x_b(t) - O_b(t) | \Theta(t)] \\ &\quad + \textstyle\sum_{b \in \mathcal{B}} \underline{U}_b(t) \mathbb{E}[\underline{x}_b(t) - \underline{O}_b(t) | \Theta(t)] \end{aligned} \quad (17)$$

where G is a positive constant given by

$$\begin{aligned} G &\ge \tfrac{1}{2} \big(\textstyle\sum_{u \in \mathcal{U}} \mathbb{E}[(O_b(t)a_u)^2 + r_u(t)^2 | \Theta(t)] \\ &\quad + \textstyle\sum_{u \in \mathcal{U}} \mathbb{E}[(\underline{O}_b(t)\underline{a}_u)^2 + \underline{r}_u(t)^2 | \Theta(t)] \\ &\quad + \textstyle\sum_{b \in \mathcal{B}} \big[\mathbb{E}[x_b(t)^2 + O_b(t)^2 | \Theta(t)] + \mathbb{E}[\underline{x}_b(t)^2 + \underline{O}_b(t)^2 | \Theta(t)] \big] \big). \end{aligned}$$

To minimize the drift-plus-penalty (15), we utilize opportunistic minimization of expectation [11, Ch. 1.8]. Thus, (15) can be minimized by solving a series of decoupled subproblems, detailed in Algorithm 1.

Algorithm 1: Algorithm for solving the UL/DL mode selection and resource allocation problem (10)

1) Flow control: Find $O_b(t)$ and $\underline{O}_b(t)$, for each b:

$$\begin{array}{ll} \text{minimize} & \sum_{u \in \mathcal{U}_b} \Big[Q_u(t)O_b(t)a_u + \underline{Q}_u(t)\underline{O}_b(t)\underline{a}_u \Big] \\ & - \Big[U_b(t)O_b(t) + \underline{U}_b(t)\underline{O}_b(t) \Big] \\ & - C\big(\phi(\mathbf{O}(t), \underline{\mathbf{O}}(t)) - \delta f(\mathbf{O}(t), \underline{\mathbf{O}}(t)) \big) \\ \text{subject to} & O_b(t), \underline{O}_b(t) \in \{0,1\}, \\ & O_b(t) + \underline{O}_b(t) = 1, \end{array}$$

2) Resource allocation (RA):

$$\begin{array}{ll} \text{minimize} & \sum_{b \in \mathcal{B}} \Big[U_b(t)x_b(t) + \underline{U}_b(t)\underline{x}_b(t) \Big] \\ & - \sum_{u \in \mathcal{U}} \Big[Q_u(t)r_u(t) + \underline{Q}_u(t)\underline{r}_u(t) \Big] \\ \text{subject to} & \text{constraints } (16a) - (16f) \end{array}$$

3) Update queues $\{Q_u, \underline{Q}_u\}_{u \in \mathcal{U}}$, $\{U_b, \underline{U}_b\}_{b \in \mathcal{B}}$.

Algorithm 1 observes current queue backlogs $\Theta(t)$ and finds solution for (10) by solving *flow control* and *resource allocation* (RA) subproblems. Step 1 finds the flow control variables $\{O_b(t), \underline{O}_b(t)\}_{b \in \mathcal{B}}$, and is decoupled among the BSs. The solution can be determined by evaluating the two points of the constraint set, i.e., $\{O_b(t) = 1, \underline{O}_b(t) = 0\}$ and $\{O_b(t) = 0, \underline{O}_b(t) = 1\}$, $\forall b$. In step 2 the transmit and receive beamformers are optimized, and for this, an algorithm is derived in the next section.

IV. Algorithm Derivation for RA

RA problem in Algorithm 1 is a weighted sum-rate maximization (WSRMax) problem along with an additional linear terms in the objective function. The WSRMax problem is known to be NP-hard [13]. In this section we provide a fast but possibly suboptimal algorithm to solve RA problem of Algorithm 1. Our approach follows a similar procedure to [10], by utilizing successive convex approximation (SCA) in conjunction with alternating optimization. In the following, we omit the time index t to simplify the notation.

We begin by introducing auxiliary variables β_{ul} and $\underline{\beta}_{ul}$ to upper-bound the denominators of the SINR expressions (1) and (2), respectively. Then, by relaxing the SINR constraints, the RA problem can be written as

maximize $\sum_{u\in\mathcal{U}}[Q_u r_u + \underline{Q}_u \underline{r}_u] - \sum_{b\in\mathcal{B}}[U_b x_b + \underline{U}_b \underline{x}_b]$
subject to constraints $(16a), (16b), (16e), (16f)$

$$\gamma_{ul} \leq |\mathbf{w}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}\mathbf{m}_{ul}|^2/\beta_{ul}, u\in\mathcal{U}, l\in\mathcal{L}_u \quad (18a)$$

$$\underline{\gamma}_{ul} \leq |\mathbf{w}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}^{\mathrm{H}}\underline{\mathbf{m}}_{ul}|^2/\underline{\beta}_{ul}, u\in\mathcal{U}, l\in\mathcal{L}_u \quad (18b)$$

$$\beta_{ul} \geq \tilde{N}_0 + \sum_{j\in\mathcal{L}_u\setminus\{l\}}|\mathbf{w}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}\mathbf{m}_{uj}|^2$$
$$+ \sum_{k\in\mathcal{U}\setminus\{u\}}\sum_{l\in\mathcal{L}_u}|\mathbf{w}_{ul}^{\mathrm{H}}\mathbf{H}_{b_k u}\mathbf{m}_{kj}|^2 \quad (18c)$$
$$+ \sum_{k\in\mathcal{U}\setminus\mathcal{U}_b}\sum_{l\in\mathcal{L}_u}|\mathbf{w}_{ul}^{\mathrm{H}}\tilde{\mathbf{H}}_{ku}\underline{\mathbf{m}}_{kj}|^2, u\in\mathcal{U}, l\in\mathcal{L}_u$$

$$\underline{\beta}_{ul} \geq \underline{\tilde{N}}_0 + \sum_{j\in\mathcal{L}_u\setminus\{l\}}|\underline{\mathbf{w}}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}^{\mathrm{H}}\underline{\mathbf{m}}_{uj}|^2$$
$$+ \sum_{k\in\mathcal{U}\setminus\{u\}}\sum_{l\in\mathcal{L}_u}|\underline{\mathbf{w}}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u k}^{\mathrm{H}}\underline{\mathbf{m}}_{kj}|^2 \quad (18d)$$
$$+ \sum_{k\in\mathcal{U}\setminus\mathcal{U}_b}\sum_{l\in\mathcal{L}_u}|\underline{\mathbf{w}}_{ul}^{\mathrm{H}}\hat{\mathbf{H}}_{b_k b_u}\mathbf{m}_{kj}|^2, u\in\mathcal{U}, l\in\mathcal{L}_u,$$

with variables $\{\mathbf{m}_{ul}, \underline{\mathbf{m}}_{ul}, \mathbf{w}_{ul}, \underline{\mathbf{w}}_{ul}, \gamma_{ul}, \underline{\gamma}_{ul}, \beta_{ul}, \underline{\beta}_{ul}\}_{u\in\mathcal{U}, l\in\mathcal{L}_u}$ and $\{x_b, \underline{x}_b\}_{b\in\mathcal{B}}$.

In order to simplify algorithm derivation, we utilize alternating optimization to design receiver combiners $\{\mathbf{w}_{ul}, \underline{\mathbf{w}}_{ul}\}$ and transmit precoders $\{\mathbf{m}_{ul}, \underline{\mathbf{m}}_{ul}\}$. That is for fixed precoders, receiver combiners $\{\mathbf{w}_{ul}, \underline{\mathbf{w}}_{ul}\}$ are set to scaled optimal MMSE solution [14]. Then for fixed MMSE receivers, transmit precoders $\{\mathbf{m}_{ul}, \underline{\mathbf{m}}_{ul}\}$ are designed. This process is repeated until convergence.

Unfortunately, the design of transmit precoders, even for fixed receivers, is a non-convex optimization problem due to constraints (18a), (18b), and (16e). In the sequel we derive an SCA based approach to convexify the nonconvex constraint functions, and solve problem (18) for fixed $\{\mathbf{w}_{ul}, \underline{\mathbf{w}}_{ul}\}$.

Constraint functions (18a) and (18b) can be expressed as difference-of-convex (DC) functions. Thus, the best convex approximation of these constraints can be obtained by replacing their R.H.S. expressions with their first order Taylor approximations. The best convex approximation of constraint functions (18a) and (18b), around fixed local points $\{\breve{\mathbf{m}}_{ul}, \breve{\beta}_{ul}\}$ and $\{\underline{\breve{\mathbf{m}}}_{ul}, \underline{\breve{\beta}}_{ul}\}$, can be expressed as

$$\gamma_{ul} \leq \frac{|\mathbf{w}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}\breve{\mathbf{m}}_{ul}|^2}{\breve{\beta}_{ul}} - \frac{\breve{\mathbf{m}}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}^{\mathrm{H}}\mathbf{w}_{ul}\mathbf{w}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}\breve{\mathbf{m}}_{ul}}{\breve{\beta}_{ul}^2(\beta_{ul}-\breve{\beta}_{ul})}$$
$$+ 2\mathrm{Re}[\breve{\mathbf{m}}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}^{\mathrm{H}}\mathbf{w}_{ul}\mathbf{w}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}(\mathbf{m}_{ul}-\breve{\mathbf{m}}_{ul})/\breve{\beta}_{ul}] \quad (19)$$

$$\underline{\gamma}_{ul} \leq \frac{|\underline{\mathbf{w}}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}^{\mathrm{H}}\underline{\breve{\mathbf{m}}}_{ul}|^2}{\underline{\breve{\beta}}_{ul}} - \frac{\underline{\breve{\mathbf{m}}}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}\underline{\mathbf{w}}_{ul}\underline{\mathbf{w}}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}^{\mathrm{H}}\underline{\breve{\mathbf{m}}}_{ul}}{\underline{\breve{\beta}}_{ul}^2(\underline{\beta}_{ul}-\underline{\breve{\beta}}_{ul})}$$
$$+ 2\mathrm{Re}[\underline{\breve{\mathbf{m}}}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}\underline{\mathbf{w}}_{ul}\underline{\mathbf{w}}_{ul}^{\mathrm{H}}\mathbf{H}_{b_u u}^{\mathrm{H}}(\underline{\mathbf{m}}_{ul}-\underline{\breve{\mathbf{m}}}_{ul})/\underline{\breve{\beta}}_{ul}]. \quad (20)$$

To tackle the combinatorial nature of (18), we relax the hard binary constraint (16e), and employ a regularization function to enforce a binary solution for vector $\mathbf{x} = [x_1, \ldots, x_B, \underline{x}_1, \ldots, \underline{x}_B]$, given by

$$h(\mathbf{x}) = \sum_{b\in\mathcal{B}}\left(\log(x_b+\epsilon) + \log(\underline{x}_b+\epsilon)\right), \quad (21)$$

where ϵ is a small positive constant. Using the above regularization function, the new objective of (18) becomes

$$\sum_{u\in\mathcal{U}}[Q_u r_u + \underline{Q}_u \underline{r}_u] - \sum_{b\in\mathcal{B}}[U_b x_b + \underline{U}_b \underline{x}_b] - \varphi h(\mathbf{x}),$$

where $\varphi > 0$ is a regularization coefficient that emphasizes a binary solution for $\mathbf{x}$. The new objective is a nonconvex DC function. As with constraints (18a) and (18b), we can convexify it by using the first order Taylor approximation of $h(\mathbf{x})$ around a local point $\breve{\mathbf{x}} = [\breve{x}_1, \ldots, \breve{x}_B, \underline{\breve{x}}_1, \ldots, \underline{\breve{x}}_B]$

$$h(\mathbf{x};\breve{\mathbf{x}}) = \sum_{b\in\mathcal{B}}\left[\log(\breve{x}_b+\epsilon) + \log(\underline{\breve{x}}_b+\epsilon) + (x_b-\breve{x}_b)/(\breve{x}_b+\epsilon) + (\underline{x}_b-\underline{\breve{x}}_b)/(\underline{\breve{x}}_b+\epsilon)\right] \quad (22)$$

Finally, omitting the constant terms of (22), we can express the convex approximation of problem (18) as

maximize $\sum_{u\in\mathcal{U}}\left[Q_u r_u + \underline{Q}_u \underline{r}_u\right] - \sum_{b\in\mathcal{B}}\left[U_b x_b + \underline{U}_b \underline{x}_b\right]$
$\quad - \varphi\sum_{b\in\mathcal{B}}\left[x_b/(\breve{x}_b+\epsilon) + \underline{x}_b/(\underline{\breve{x}}_b+\epsilon)\right]$
subject to constraints $(16a), (16b), (18c), (18d), (19), (20)$

$$x_b, \underline{x}_b \in [0,1], b\in\mathcal{B} \quad (23a)$$
$$x_b + \underline{x}_b = 1, b\in\mathcal{B} \quad (23b)$$

A. Heuristic Decoupling of Precoders and Power Allocation

In the RA problem of Algorithm 1, UL/DL power allocation variables $\{x_b, \underline{x}_b\}_{b\in\mathcal{B}}$ and precoders/combiners $\{\mathbf{m}_{ul}, \underline{\mathbf{m}}_{ul}, \mathbf{w}_{ul}, \underline{\mathbf{w}}_{ul}\}$ are coupled due to constraints (10c) and (10d). A heuristic decoupling can be applied by setting $x_b(t) = O_b(t)$ and $\underline{x}_b(t) = \underline{O}_b(t)$ (or, by observing the long-term statistics of $O_b(t)$ and $\underline{O}_b(t)$ from the solutions of step 1 of Algorithm 1), $\forall b, t$. This heuristic can be justified by the fact that problem (18) is already guided by the transport-layer mode selection through the queues Q_u, $\underline{Q}_u$, U_b and $\underline{U}_b$, the state of which determines the power allocation variables.

The decoupling makes problem (18) a known WSRMax under a fixed UL/DL mode, which can be solved separately at the BSs using, for example, over-the-air training [8], [9]. However, decoupling the variables in this manner is a suboptimal solution, and can degrade the overall system performance.

V. NUMERICAL RESULTS

We consider a system with $B = 7$ BSs each with $M = 4$ antennas, and $U = 14$ users each with $N = 2$ antennas. Users are equally divided among the BSs (i.e., two users per BS). We consider single-stream transmission, i.e., $L_u = 1, \forall u$. The users are equidistant from their serving BSs with positions randomized every 50 time slots, with a pathloss of 20 dB, and there is 3 dB separation between cells. The DL transmit powers are set to $P_{\max} = 10$ dB for all BSs, the UL transmit powers $P_{\max}^{\mathrm{UE}} = P_{\max}/|\mathcal{U}_b|$ for all users, and the noise level is set such that SNR is 15 dB. We use Clarke's channel model, with 300 paths and stationary UEs.

We model the arriving data in DL as $\lambda_u(t) \sim \mathrm{Pois}(\rho_u), \forall u$, where ρ_u is the average arrival rate, and in UL as $\underline{\lambda}_u(t) \sim \mathrm{Pois}(\underline{\rho}_u), \forall u$. We consider asymmetric arrival, where a set of five BSs (denoted by $\mathcal{B}^{\mathrm{DL}}$) are under heavy DL traffic with $\{\rho_u, \underline{\rho}_u\} = \{3, 0.1\}$ bits/slot, $u\in\mathcal{U}_b$, $b\in\mathcal{B}^{\mathrm{DL}}$, and the set of the other two ($\mathcal{B}^{\mathrm{UL}}$) are under heavy UL traffic with $\{\rho_u, \underline{\rho}_u\} = \{0.1, 3\}$ bits/slot, $u\in\mathcal{U}_b$, $b\in\mathcal{B}^{\mathrm{UL}}$. The user priority weights
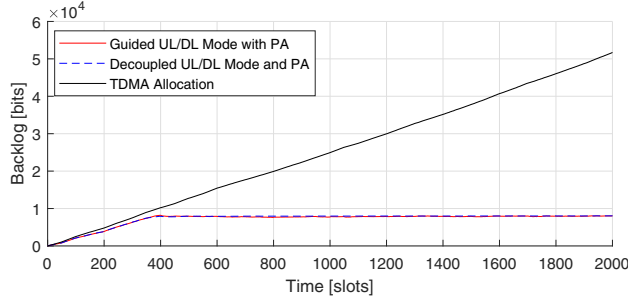
Fig. 1. Network-layer queue evolution as a function of time.

Fig. 2. Top: Objective value as a function of C.
Bottom: Total network-layer backlog as a function of C.

are set as $w_u = \rho_u$ and $\underline{w}_u = \underline{\rho}_u$, $\forall u$. For flow control, we set $A^{\max} = \underline{A}^{\max} = \log_2(1 + P_{\max}\overline{M})$, and the trade-off parameter is set to $\delta = 10$. The regularization parameter enforcing the binary solution for $\{x_b, \underline{x}_b\}$ is set to the average of $|\Theta(t)|$.

To illustrate the performance of Algorithm 1, we first plot the queue evolution of the system for $C = 500$ in Fig. 1. The guided method (red line), is obtained by solving problem (18) for the resource allocation, while the decoupled method (blue line) follows the heuristics of Section IV-A. TDMA mode allocation is plotted as reference, where we alternatingly set all BSs to DL or UL mode on consecutive time slots. From the figure we can see that the proposed algorithm achieves stable steady state after 400 time slots, while TDMA is unstable. Furthermore, in terms of queue stability, there is virtually no difference between using the guided and heuristic methods. The queue evolution behaviour is similar for other values of C, but the results are not provided due to space limitation.

Next, we investigate the performance of Algorithm 1 for different values of C, shown in Fig. 2. The results are averaged over 2000 time slots, and in case of TDMA, due to its instability, we use the achieved average rates ($\overline{r}_u, u \in \mathcal{U}$) instead of flow rates ($\overline{O_b a_u}, u \in \mathcal{U}_b, b \in \mathcal{B}$) to calculate the objective value. From the top graph of Fig. 2 we can see that for high enough value of C, the proposed algorithm performs significantly better than the TDMA reference case. This is due to high value of C emphasizing the objective, which is a measure of interference and data flow between the transport and network layers, and for which TDMA has difficulties handling the asymmetric traffic. The interference portion of the objective was noticed to be very small due to the multi-antenna users, which can already mitigate much of the inter-cell interference via beamforming. Here we can also see the performance drop due to the heuristic decoupling of Section IV-A, as the resource allocation can no longer optimize the power allocation variable along the beamformers/combiners. Finally, from the bottom graph of Fig. 2 we can see that the queue backlog increases with C, as the drift portion of (15) is less emphasized, but there are only minor differences in the steady states between the guided and heuristic mode selection methods. We've omitted the TDMA backlog due to its instability.

VI. CONCLUSIONS

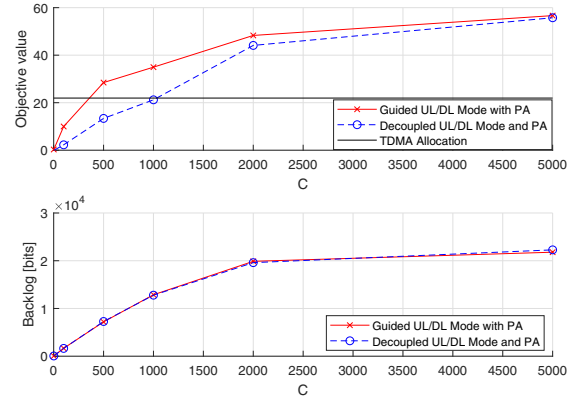We have proposed a dynamic UL/DL mode selection and resource allocation algorithm, that can dynamically account for the varying traffic demands of small cell systems. The proposed algorithm aims to maximize the long term network utility. The algorithm derivation is based on Lyapunov optimization, and provides a sequence of subproblems to solve the original problem. The numerical results show clear advantages of the proposed algorithm over the TDMA reference case under asymmetric traffic scenarios.

REFERENCES

[1] 3GPP TSG RAN WG1, "Study on new radio access technology: Physical layer aspects (3GPP TR 38.802 version 14.2.0 release 14)," *Tech. Rep., 3rd Generation Partnership Project 3GPP, www.3gpp.org*, 2017.

[2] ——, "Revised WID on cross link interference (CLI) handling and remote interference management (RIM) for NR (3GPP RP-182864)," *Tech. Rep., 3rd Generation Partnership Project 3GPP, www.3gpp.org*, 2018.

[3] W. Jeong and M. Kavehrad, "Cochannel interference reduction in dynamic-TDD fixed wireless applications, using time slot allocation algorithms," *IEEE Trans. Commun.*, vol. 50, no. 10, pp. 1627–1636, 2002.

[4] P. Omiyi, H. Haas, and G. Auer, "Analysis of TDD cellular interference mitigation using busy-bursts," *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2721–2731, 2007.

[5] I. Sohn, K. B. Lee, and Y. S. Choi, "Comparison of decentralized time slot allocation strategies for asymmetric traffic in TDD systems," *IEEE Trans. Wireless Commun.*, vol. 8, no. 6, pp. 2990–3003, 2009.

[6] H. ElSawy, E. Hossain, and M. Alouini, "Analytical modeling of mode selection and power control for underlay D2D communication in cellular networks," *IEEE Trans. Commun.*, vol. 62, no. 11, pp. 4147–4161, Nov. 2014.

[7] A. Tölli, J. Kaleva, and P. Komulainen, "Mode selection and transceiver design for rate maximization in underlay D2D MIMO systems," in *Proc. IEEE Int. Conf. Commun.*, Jun. 2015, pp. 4822–4827.

[8] P. Jayasinghe, A. Tölli, J. Kaleva, and M. Latva-aho, "Bi-directional beamformer training for dynamic TDD networks," *IEEE Trans. Signal Processing*, vol. 66, no. 23, pp. 6252–6267, Dec. 2018.

[9] P. Komulainen, A. Tölli, and M. Juntti, "Effective CSI signaling and decentralized beam coordination in TDD multi-cell MIMO systems," *IEEE Trans. Signal Processing*, vol. 61, no. 9, pp. 2204–2218, May 2013.

[10] A. Tölli, J. Kaleva, G. Venkatraman, and D. Gesbert, "Joint UL/DL mode selection and transceiver design for dynamic TDD systems," 2016, pp. 630–634.

[11] M. J. Neely, "Stochastic network optimization with application to communication and queueing systems," *Synthesis Lectures on Communication Networks*, vol. 3, no. 1, pp. 1–211, 2010.

[12] D. Tse and P. Viswanath, *Fundamentals of wireless communication*. Cambridge university press, 2005.

[13] Z. Luo and S. Zhang, "Dynamic spectrum management: Complexity and duality," *IEEE J. Select. Topics Signal Processing*, vol. 2, no. 1, pp. 57–73, Feb. 2008.

[14] G. Venkatraman, A. Tölli, M. Juntti, and L. Tran, "Traffic aware resource allocation schemes for multi-cell MIMO-OFDM systems," *IEEE Trans. Signal Processing*, vol. 64, no. 11, pp. 2730–2745, Jun. 2016.