Traffic-Aware Training and Scheduling for MISO Wireless Downlink Systems

Apostolos Destounis, Mohamad Assaad, Mérouane Debbah, Bessem Sayadi

Traffic-Aware Training and Scheduling for MISO Wireless Downlink Systems

Apostolos Destounis, *Member, IEEE,* Mohamad Assaad, *Member, IEEE,* Mérouane Debbah, *Fellow, IEEE,* and Bessem Sayadi, *Member, IEEE*

Abstract

In this paper, the problem of feedback and active user selection in MISO wireless systems such that the system's stability region is as big as possible is examined. The focus is on a system in a Rayleigh fading environment where zero forcing precoding is used to serve all active users in every slot. Acquisition of the channel states is done via uplink training in Time Division Duplexing mode by the active users. Clearly, only a subset of users can perform uplink training and the selection of this subset is a challenging and interesting problem especially in MISO systems. The stability regions of a baseline centralized scheme and two novel decentralized policies are examined analytically. In the decentralized schemes, the transmitter broadcasts periodically the queue state information and the users contend for the channel in a CSMA-based manner with parameters based on the outdated queue state information and real-time channel state information. We show that, using infrequent signaling between the base station and the users, the decentralized policies outperform the centralized policy. In addition a threshold-based user selection and training scheme for discrete-time contention is proposed. The results of this work imply that, as far as stability is concerned , the users must be involved in the active user selection and feedback/training decision. This should be leveraged in future communication systems.

A. Destounis is with CentraleSupélec, 3 rue Joliot-Curie, 91192, Gif-sur-Yvette, cedex. France.(apostolos.destounis@supelec.fr).

M. Assaad is with Laboratoire des Signaux et Systèmes (L2S, UMR CNRS 8506) CentraleSupélec , 3 rue Joliot-Curie, 91192, Gif-sur-Yvette, cedex. France. (mohamad.assaad@supelec.fr)

M. Debbah is with CentraleSupélec, 3 rue Joliot-Curie, 91192, Gif-sur-Yvette, cedex. France. (merouane.debbah@supelec.fr)

B. Sayadi is with Alcatel-Lucent Bell Labs France, Villarceaux, Route de Villejust, 91620, Nozay, cedex. France (bessem.sayadi@alcatel-lucent.com)

The work presented in this paper was done while the first author was with Alcatel-Lucent Bell Labs France. Parts of this paper have been presented at the IEEE International Symposium on Information Theory (ISIT), Honolulu, HI, U.S.A., 2014 and have been accepted for presentation at the IEEE International Conference on Communications (ICC), London, U.K., 2015.

Index Terms

MISO broadcast channel, stability, bursty traffic, zero forcing

I. Introduction

The use of multiple antennas [1] has emerged as one of the enabling technologies to increase the performance of wireless systems. The ability to serve multiple users in the same time-frequency block has made the use of multiple antennas at the base stations (BSs) particularly attractive for multiuser downlink systems, and the benefits coming from this fact are well understood [2]. The decision to be taken every timeslot then is (i) which users should be scheduled and (ii) how the corresponding signals should be precoded.

From an information theoretic point of view, the capacity region of the multiple-input multiple-output (MIMO) broadcast channel (BC) is well characterized [3], assuming perfect channel state information at the transmitter. However, achieving this region requires Dirty Paper Coding, which is complex to implement, while the assumptions of perfect channel state information and use of Gaussian codebooks are strong in practical systems. Linear precoding schemes such as zero forcing (ZF) , a scheme that cancels the interference among the scheduled users, are more desirable to use in practice. There are many works on the issue of imperfect channel state information (CSI) , see for example [4] and references therein, [5], [6], [7], however the focus is mainly on quantities like sum throughput and they do not take into account the traffic processes of the users.

It is of theoretical and practical interest to study the impact of MIMO in the higher layers [8]. For the MIMO media access control (MAC) , a precoding strategy that achieves the stability region is presented in [9], under the assumptions of perfect CSI and use of Gaussian codebooks. This policy is based on Lyapunov drift minimization given the queue lengths and channels every timeslot and makes use of superposition coding and successive decoding. This is hard to implement in practice. Regarding the BC (i.e.the downlink system), authors in [10] have proposed a technique based on ZF precoding, with a heuristic user scheduling scheme that selects users whose channel states are nearly orthogonal vectors and illustrate the stability region this policy achieves via simulations. Authors in [11] notice that the policy resulting from the minimization of the drift of a quadratic Lyapunov function is to solve a weighted sum rate maximization problem (with weights being the queue lengths) each timeslot and they propose an iterative water-filling algorithm for this purpose. In addition, authors in [12] propose to use the delays of the packets in the head of each queue along with the queue lengths as weights. All these works assume accurate CSI available at the transmitter. In the case of delayed channel state information and channels having a correlation in time, authors in [13] compare the stability and delay performance of opportunistic beamforming and space time coding while in [14] they propose a user scheduling and precoding algorithm. In addition, in [15], the authors study the impact of channel state quantization in the stability of a system using ZF precoding under a centralized scheme where the transmitter selects the users to be scheduled based only on the queue lengths. However, the fact that radio resources i.e. time and/or spectrum are needed for the BS to acquire channel state information is not accounted for in these works.

In this paper we consider a multiple-input single-output (MISO) downlink system where the BS acquires CSI from the users in time-division duplexing (TDD) mode, in order to exploit the channel reciprocity. There are two ways for this: (i) users estimate their channel and then feed back the CSI in a time-division multiple access (TDMA) manner and (ii) users send (pre-assigned) training sequences in the uplink so that the BS can estimate the channels (uplink and downlink channels are the same due to reciprocity). The latter scheme is implemented using orthogonal sequences among the users, so that the BS can decode every transmission without errors. Orthogonal sequences are produced e.g. by Walsh-Hadamard on pseudonoise sequences, and their length should be proportional to the number of users that simultaneously train in the uplink. Uplink training is considered the most promising for MIMO systems, since the length of the training sequences does not depend on the number of antennas at the BS . However, due to the orthogonality requirement, their length is proportional to the number of users that perform uplink training [1]. That means that in a system with many users, not all users should be selected to train at the same time, therefore the users that should train at every slot must also be selected. The TDD system model with uplink training has been also examined in [16], however they do not take into account that last observation, that is not all users should participate in the training at every slot. In this paper we focus on the tradeoff between having many users training (so having data transmitted to many users simultaneously) and having much time of the slot devoted to data transmission (which means having few users train). In order to simplify things, we focus on ZF precoding used at the transmitter. This scheme is widely used in the literature because it is simple to implement while capturing the fundamental tradeoffs arising from using multiple antennas and performing well in some scenarios of interest (e.g. in systems with many users [17] and/or with BSs with large antenna arrays). In addition, we will assume that all

[1]In the case where CSI acquisition is done by feedback in TDMA manner, then the time needed for CSI acquisition is also proportional to the number users that feed back.

users that perform uplink training in a slot get scheduled. This is an assumption used fairly often in the literature concerning MIMO broadcast channels; in this context, the BS should select the set of ”active users” at each timeslot and then transmit to them.

One natural approach would be to let the BS alone decide which users to schedule in every slot. This is the approach used in [15] and in current standards (e.g. Long Term Evolution (LTE) [18]), where the BS explicitly requests some users for their CSI . In the setting where traffic/queueing processes are considered, user selection can be done based on the statistics of the channels of the users and the state of their queue lengths at each slot. Unfortunately, using such centralized schemes, some scheduled users may have poor current channel states and some users with good channels may not be scheduled (i.e. may not feed back), which reduces the system performance. On the other hand, each user knows its own current channel state, and therefore decentralized feedback policies where the users decide based on their current channel states may improve the system performance. This must be done properly as the decentralized policies require additional signalling information that may decrease drastically the improvement.

It is worth noting that recent works [19], [20] have shown that, in a network with simple physical layer (e.g. on-off channel, finite discrete channel states,...), decentralized algorithms like the recently proposed Fast carrier sense multiple access (CSMA) [21] can achieve good performance. In addition, it has been shown in earlier works [22], [23] that up-to-date channel state information, which is known at the receivers, is more crucial than accurate queue length information, at least as far as stability is concerned. The scenario considered in this paper is more complicated as compared to the recent work on decentralized scheduling. In fact, in scheduling problems (e.g. orthogonal frequency-division multiplexing access (OFDMA) or TDMA), a user can directly estimate its bit rate using the current channel state. In multi-user MIMO systems, the bit rate of each user depends on the channel states of all users and the user cannot simply estimate its bit rate using its current channel state, which highly complicates the analysis.

In this paper, we examine three approaches to the user selection problem. The first one is centralized, in the sense that the transmitter decides which user will be scheduled (i.e. will train) at every slot. The second approach, which we term as decentralized, is to let the users decide which of them should actually feed back via some contention/coordination scheme. The main idea behind this approach is that every user can know its channel state, therefore a user with a very bad channel state will choose not to feed back (contrary to what can happen in the centralized approach). More specifically, in this case, the transmitter specifies the number of users to be scheduled and lets the users decide in a decentralized manner who will be the ones that will actually get scheduled in the slot. Combined with some (infrequent) signalling regarding the users queue lengths from the BS , we prove that properly combining the decentralized and centralized approaches leads to a bigger achievable stability region than using the centralized approach alone.

The rest of this paper is organized as follows. The system model and the interaction between physical layer and queueing performance are presented in Section II. Detailed description of these policies is given in Section III, after a presentation of the system model in Section II. Section IV presents some calculations regarding the rate distributions and some general intermediate results regarding stability, that will be used for the proofs in subsequent Sections. In Section V we examine in detail a special case, namely the $2-$user system with independent and identically distributed (i.i.d.) channels and single rate level. This is a case where the stability regions can be expressed in closed form and plotted, and helps illustrate why combining a decentralized and a centralized approach helps in enlarging the stability region of the system. After that, Section VI contains stability analysis in the general case of K users, while extensions to the case where multiple channels are used (e.g. via OFDMA modulation) is covered in Section VII. Finally, in Section VIII we discuss an alternative implementation where extra signalling bandwidth instead of time is used for the control signals required to be broadcasted for the decentralized/mixed policies and Section IX presents a threshold-based policy in the cases where continuous time for contention is not possible. The proofs for the derivations of the stability regions are done based on the method of first proving that the stated region is achievable by a rule that does not take into account the queue lengths, prove, using the Foster-Lyapunov criterion, that the proposed policy achieves at least as big region as the first rule and then prove that there is no policy achievng bigger than the stated region. Finally, Section X concludes the paper.

Notations: In this paper, boldface uppercase letters denote matrices, boldface lowercase letters denote column vectors and non-boldfqce letters denote scalars. In particular, x_k is the $k-$ th element of vector $\mathbf{x}$. $\mathbf{I}_N$ denotes the identity matrix of size N. The notation $||\mathbf{x}||$ is used for the Euclidean norm of vector $\mathbf{x}$, while $||\mathbf{x}||_1 = \sum_{k=1}^{K} x_k$. Superscripts T and H over a matrix or vector denote its transpose and hermitian, respectively. In addition, $\mathcal{CN}(\mu, \mathbf{R})$ denotes a complex normal random vector with mean μ and covariance matrix $\mathbf{R}$, while $\mathcal{CH}$ denotes the convex hull operation. We will use the following notations for the Gamma, upper incomplete Gamma, Beta and regularized upper incomplete beta functions, respectively: $\Gamma(N) = \int_0^\infty t^{N-1}e^{-t}dt$, $\gamma(x; N) = \int_0^x t^{N-1}e^{-t}dt$, $B(a, b) = \int_0^1 t^{a-1}(1-t)^{b-1}dt$ and $I_B(x; a, b) = \int_0^x t^{a-1}(1-t)^{b-1}dt$. Finally, $\mathbb{1}_{\{.\}}$ denotes the indicator function.

II. System Model

A. Physical Layer Model

We consider a single cell wireless system serving K users where the BS is equipped with N antennas. The users are equipped with a single antenna each. Time is slotted. Each channel is i.i.d. Rayleigh block fading, i.e. the channels stay constant in a slot of T_s channel uses and change independently in the next slot. The channel of user k can be written as an N-dimensional complex vector $\mathbf{h}_k(t) = \sqrt{\bar{g}_k}\hat{\mathbf{h}}_k(t)$ where $\hat{\mathbf{h}}_k(t) \sim \mathcal{CN}(0, \mathbf{I}_N)$ and $\bar{g}_k$ represents the channel gain due to large scale fading and is assumed to be constant (e.g. following a path loss model depending on the distance from the BS). In this case, the channel vector can be also written as $\mathbf{h}_k(t) = \sqrt{g_k(t)}\mathbf{u}_k(t)$, where $\mathbf{u}_k(t)$ is an isotropically distributed unitary vector and $g_k(t) = ||\mathbf{h}_k(t)||^2$ is the channel magnitude. We have the following:

Lemma 1. *The channel magnitude of user k is distributed with cumulative distribution function (CDF) :*

$$\mathbb{P}\left\{g_k(t) < x\right\} = \frac{\gamma\left(\frac{x}{\bar{g}_k}; N\right)}{\Gamma(N)}.$$

Proof. We can write $g_k(t) = \bar{g}_k \left|\left|\hat{\mathbf{h}}_k(t)\right|\right|^2$. Since $\hat{\mathbf{h}}_k(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$, the random variable $Z = 2\left|\left|\hat{\mathbf{h}}_k(t)\right|\right|^2$ has the chi-squared distribution with $2N$ degrees of freedom. We thus have

$$\mathbb{P}\left\{g_k(t) < x\right\} = \mathbb{P}\left\{\left|\left|\hat{\mathbf{h}}_k(t)\right|\right|^2 < \frac{x}{\bar{g}_k}\right\} = \mathbb{P}\left\{2\left|\left|\hat{\mathbf{h}}_k(t)\right|\right|^2 < 2\frac{x}{\bar{g}_k}\right\} = \frac{\gamma\left(\frac{x}{\bar{g}_k}; N\right)}{\Gamma(N)}.$$

□

Since the BS is equipped with multiple antennas, multiple users can be served simultaneously by precoding the corresponding transmit signals. In this paper, we consider ZF precoding, i.e. a linear precoder such that the intracell interference caused to any user that is served is zero. The choice is motivated by the relatively low complexity of linear precoders (and ZF in particular) with respect to the non-linear ones and by the fact that Zero Forcing is a simple scheme to analyze, that however captures the fundamental tradeoffs in multiple antenna transmission and has very good performance in many scenarios of interest (as, for example, in the case where large antenna arrays are used).

As for any linear precoder $\mathbf{W}(t) = [\mathbf{w}_1(t), ..., \mathbf{w}_K(t)]$, the signal received by user k at slot t is

$$y_k(t) = \mathbf{h}_k^H(t)\mathbf{w}_k(t)s_k(t) + \sum_{j \neq k} \mathbf{h}_k^H(t)\mathbf{w}_j(t)s_j(t) + n_k(t) \tag{1}$$

where $s_j(t)$ is the data symbol intended to user j, assumed complex Gaussian with zero mean and unit power, $n_k(t) \sim \mathcal{CN}\left(0, \sigma^2\right)$ is the white noise at the receiver or user k. The BS has available transmission power P, that is tr$(\mathbf{W}\mathbf{W}^H) \leq P$. The achievable rate of user k at slot t is $r_k(t)$ (bits/channel use), which depends on the corresponding signal-to-noise ratio (SNR) at time t. We will assume that the achievable rates can take values from the set $\mathcal{R} = \{R_1, ..., R_l, .., R_L\}$, with $R_1 = 0$, $R_{l-1} < R_l$ and $R_L < \infty$; this is the case in practice as a finite number of modulation and coding schemes are used for transmission. In addition, we assume that the possible rate levels are known to the BS and mobiles, which is rather reasonable since they are specified by the communications protocol used. Also we assume that rate R_l can be supported if the signal-to-interference-plus-noise ratio (SINR) at the receiver is above some appropriately defined threshold S_l.

Under this model, let $\mathcal{F}(t)$ be the set of users that are scheduled at slot t, $F(t) = |\mathcal{F}(t)|$ and $k(1), .., k(i), ..., k(F(t))$ the corresponding permutation of user indices. Also, define

$$\mathbf{H}_{k(i)}(t, \mathcal{F}) = [\mathbf{h}_{k(1)}(t), ..., \mathbf{h}_{k(i-1)}(t), \mathbf{h}_{k(i+1)}(t), ..., \mathbf{h}_{i(F(t))}(t)].$$

To further reduce the complexity of the transmission scheme, the total transmit power is split equally to each of the users scheduled. Then, the precoding vector for user $k(i), \forall i \in \{1, .., F(t)\}$ is given as the projection of the channel of this user on the nullspace generated by the channels of the other users:

$$\mathbf{w}_{k(i)}(t) = \sqrt{\frac{P}{F}} \frac{\left(\mathbf{I}_N - \mathbf{H}_{k(i)}(t, \mathcal{F})(\mathbf{H}^H_{k(i)}(t, \mathcal{F})\mathbf{H}_{k(i)}(t, \mathcal{F}))^{-1}\mathbf{H}^H_{k(i)}(t, \mathcal{F})\right)}{\left|\left|\left(\mathbf{I}_N - \mathbf{H}_{k(i)}(t, \mathcal{F})(\mathbf{H}^H_{k(i)}(t, \mathcal{F})\mathbf{H}_{k(i)}(t, \mathcal{F}))^{-1}\mathbf{H}^H_{k(i)}(t, \mathcal{F})\right)\mathbf{h}_{k(i)}(t)\right|\right|} \mathbf{h}_{k(i)}(t). \tag{2}$$

The corresponding SNR (since interference is suppressed) for this user will then be

$$\begin{aligned}\mathrm{SNR}_{k(i)}(t) = \frac{P\left|\left|\mathbf{h}_{k(i)}(t)\right|\right|^2}{\sigma^2 F(t)} \mathbf{u}^H_{k(i)}(t) \Big(\mathbf{I}_N \\ - \mathbf{H}_{k(i)}(t, \mathcal{F}(t))(\mathbf{H}^H_{k(i)}(t, \mathcal{F}(t))\mathbf{H}_{k(i)}(t, \mathcal{F}(t)))^{-1}\mathbf{H}^H_{k(i)}(t, \mathcal{F}(t))\Big)\mathbf{u}_{k(i)}(t).\end{aligned} \tag{3}$$

From the above, it can be seen that in order to transmit using ZF , accurate channel state information of the channels of the users that are scheduled is needed. This information is not available to the BS and must be acquired by using feedback or training from the receivers. For consistency, we will consider the case where CSI is acquired by uplink training from the users. This means that channel estimation is done in TDD mode, exploiting reciprocity; this is a promising approach, especially for large antenna arrays at the BS , since the feedback overhead does not scale with the number of antennas. It does scale with the number of users that train however, meaning that when CSI is acquired by too many users there will be little time left to transmit in the timeslot before the channels change again: This problem is exactly the focus of the paper. On the other hand, even if CSI is acquired by feedback in frequency-division duplexing

(FDD) mode, the BS must wait for the feedback from the users to be received before precoding [5]. Also under our model of i.i.d. block fading channels, outdated feedback is not useful. The above imply that the main ideas and results of the analysis presented can be useful even in systems with feedback in FDD mode, assuming accurate channel estimation from the users' side, enough bit rate in the reverse link for perfect CSI in the BS after the feedback procedure and that the bandwidth in the uplink is not enough for all users to feed back simultaneously in parallel channels. In addition, we will assume throughout this paper that there are no errors in the channel estimation, in order to focus on the impact of time needed for training in our system.

As far as power considerations are concerned, if the CSI estimation is done with an MMSE estimator via uplink pilots of length τ, (transmit) power P and error variance normalized to 1, the error variance for the channel of user k is given as (result adapted from [24]):

$$\tilde{\sigma}_k^2 = \frac{\bar{g}_k}{\bar{g}_k P \tau + 1}.$$

In the above, $\bar{g}_k$ is the channel gain due to large scale fading. In the paper, we implicitly assume that the pilot power of user k is tuned according to the large scale fading and the pilot length in a way that the above error variance is small enough to assume that the impact of the estimation error is negligible. If this was not the case, one could write the channel estimation as

$$\hat{\mathbf{h}}_k = \mathbf{h}_k + \mathbf{e}_k,$$

where the first term is the real channel realization and the second termis the estimation error, and additionally average over the error statistics to obtain the expected values of the rates. This task however would complicatea lot the analysis (as some interference from non perfect zero forcing should be accounted for as well). The same holds if one wants to take into account errors in the channel estimation at the user terminals from the training signal sent by the base station (in this case, the length of the training sequence must be bigger than the number of antennas at the base station). In addition, similar discussion holds if the base station acquires CSI from feedback (done either in the same band or in FDD mode) rather than training. In this case, the poweer and number of symbols used by each user terminal for feedback should be high enough in order totransmit a very fine quantization of the CSI.

B. Queuing model and impact of training

Each of the K users in the cell has an incoming traffic process $a_k(t)$, which is an integer-valued process, measured in bits, i.i.d. in time and independent across users with $a_k(t) < A_{max}$ almost surely for some finite constant A_{max}. This quantity is assumed to be known to the scheduler and users. The mean rate of this process is $\mathbb{E}\{a_k(t)\} = \lambda_k$. Data for user k is stored in a respective buffer until transmission and let $q_k(t)$ denote its size in bits at the beginning of slot t.

Denote now $z_k(t)$ as the schedule in timeslot t, that is $z_k(t) = 1$ if user k is scheduled for this timeslot (i.e. if user k has actually reported its channel to the BS). We are under the constraints that (i) $F(t)$ users are scheduled at each timeslot, with $F(t) \leq N$ (ii) for every channel the BS schedules a user whose channel state is known at the maximum possible rate it can support (that is ensuring transmission without errors). In addition, we will denote here $\tau(t)$ the number of channel uses used for training and signalling in the slot t. This means that data is transmitted for a scheduled user for $(T_s - \tau(t))$ channel uses, therefore, if the rate supported to user k at timeslot t is $r_k(\mathbf{W}(t), \mathbf{H}(t)) \in \mathcal{R}$ bits per channel use, the corresponding service process will be $(T_s - \tau(t))r_k(\mathbf{W}(t), \mathbf{H}(t))$ bits [2]. The queues then evolve as follows, $\forall k \in \{1, .., K\}$:

$$\begin{aligned} q_k(t+1) &= [q_k(t) - \lfloor (T_s - \tau(t)) r_k(\mathbf{W}(t), \mathbf{H}(t)) \rfloor z_k(t)]^+ + a_k(t), t \geq 0 \\ q_k(0) &= a_k(-1). \end{aligned} \tag{4}$$

In the above, $a_k(-1)$ is defined as a random variable drawn from the distribution of $a_k(t), t \geq 0$. This constraint actually means that we start measuring time after the first arrivals in the queues so that the queues do not start empty (and the broadcast of the queue lengths at time $t = 0$ not to be the zero queue). This is done for more convenience in analysing the proposed algorithm, however, since we are interested in the case when the system is left running for long enough for the corresponding Markov chains to reach their invariant distribution (i.e. $t \to \infty$) and the arrival processes are bounded, the choice of the initial condition does not really affect our results.

In this work we are interested in the stability of the system. Formally, its definition is as follows:

Definition 1 (Strong Stability)**.** *A system is said to be strongly stable if*

$$\limsup_{T \to \infty} \frac{1}{T} \sum_{t=0}^{T-1} \mathbb{E}\{q_k(t)\} < \infty, \forall k \in \{1, .., K\}$$

Intuitively stability implies that the mean queue length of every queue in the system is finite, further implying finite delays in single hop systems. The figure of interest in this work is strong stability, therefore in the remainder of the manuscript ”stable” will imply ”strongly stable” unless stated otherwise.

[2]the arguments in the rate stress that it depends on the channel state and precoder at timeslot t

The above definition holds for a fixed mean arrival rates and resource allocation policy, and leads to the concept of a stability region.

Definition 2 (Stability Region)**.** *The stability region* Λ *of a resource allocation policy is defined as the set of vectors of mean arrival rates for which the system is stable under this policy. Furthermore, an algorithm that achieves the maximum possible stability region is called throughput optimal.*

For the rest of the paper, when describing stability regions we will mean that the system is stable in the *interior* of the calculated region. The behaviour on the boundary is not examined - usually for the boundary points the system is stable in at least a weaker sense, i.e. mean rate stable [25]. If the arrivals and service rate processes are such that the Markov chain is irreducible and aperiodic with a single communicating class and the control actions are taken as functions of the queue state only, a necessary condition for strong stability is that the Markov chain is positive recurrent and the mean service rates (under the invariant distribution) are bigger than the mean arrival rates. Since we are considering the interiors of the stability regions, i.e. not examining the cases where the mean arrival rates qre equalto the mean service rates, the above condition is sufficient for our analysis. In a stable system the users experience finite delays, therefore stability is a relevant aspect for data services and users with large delay tolerance. Uplink training affects essentially the service rate, and thus the stability region, in two ways: First, more time devoted to training leads to lower service rate for the users actually scheduled in the timeslot. On the other hand, if more users participate in the training, more users can get scheduled in a timeslot, thus overall a user can get higher mean service rate. The focus of this paper is, then, this tradeoff and how to efficiently design user selection strategies to achieve large stability regions.

For simplicity, we only consider schemes where all users whose CSI is acquired get scheduled. Also, as mentioned before, transmit power is allocated equally to scheduled users. The first assumption is common in the literature concerning MIMO broadcast systems, where all ”active users” are scheduled, see e.g. [5]. In addition, even in the case of full CSI without any cost, the highest stability region would be given from solving a weighted sum maximization problem in each timeslot (in accordance to the MaxWeight rule [26], [27]). Joint scheduling and power control even in this ideal case is a hard problem, especially in our model with finite rate set (some algorithms like iterative waterfilling [11] have been proposed but using the information theoretic capacity for the service rates). Adding the cost of feedback on top of it would make the problem more complicated and result in a solution of high computational complexity (see e.g. [28] for the problem in the single antenna case, where only approximations of the optimal solution are implementable in practice). The problem then reduces to finding strategies to choose the "active" users at each time slot.

III. Proposed Policies for User Selection

In this section we present in detail the scheduling and training policies to be analyzed in the present work. Before proceeding in the descriptions, we define R_0 (bits per channel use) the rate at which the control information from the BS to the users can be broadcasted. Further, we assume that when F users perform the uplink training, they use orthogonal pilot of length βF channel uses. β is an integer system parameter, not smaller than 1 (for the pilots to be indeed orthogonal). Greater length of training sequences implies that the training symbols should have less power (for the same quality of estimation) so this parameter can be tuned according to the power capabilities of the user terminals. Finally, downlink pilots are assumed of a length of β_p channel uses. Since at least a downlink pilot and the uplink pilots must be used, the maximum number of users to be scheduled is the maximum number of users that can participate in training within the duration of the timeslot, that is

$$F_{max} = \min\left\{N, \left\lfloor \frac{T_s - \beta_p}{\beta} \right\rfloor\right\}. \tag{5}$$

In practice this number may be actually lower, for example it is usually desirable that at least half of the timeslot is used for data transmission [29].

In the following subsections, we present in detail the scheduling and feedback schemes examined. For the reader's convenience, we have put the list of variables included in the paper in Table I.

A. *Centralized policy*

Since the channel statistics are known, one approach is to let the BS decide which users to acquire CSI from and schedule based on the expectation of the rate they receive. Each expectation is found over the joint probability distribution of the channel realizations of *all* users and its expression is given in Section IV-A. For this scheme, the BS sends a downlink pilot to allow the users to estimate their channel and decode the control messages. After the pilot, there is a control phase, where the BS broadcasts the identification numbers (IDs) of the users that will get scheduled at this timeslot. For each user selected

$$\beta_c = \frac{\log_2 K}{R_0}$$

channel uses are needed. In addition, there must be a way for the users to know that the control part is over and that those among them that are scheduled should start training. Here we propose that the signal for that is the BS staying silent for one channel use. An alternative, and perhaps more robust, way would

K	number of users in the system
N	number of antennas at the base station
F	number of users feeding back
$\mathcal{R} = \{R_1, .., R_L\}$	set of possible data rates: $0 = R_1 < .. < R_L$
σ^2	variance of the Gaussian noise
$\bar{g}_k$	average channel gain for user k
T_s	number of channel uses per time slot, a time slot is one fading block
$(\lambda_k)_{k \in \{1,..,K\}}$	vector of average arrival rates in bits per time slot
A_{max}	maximum number of arrivals per time slot
R_0	bits per channel use for control information from the base station
β	positive integer, length of uplink pilot in channel uses per user for training
β_p	channel uses per slot for downlink pilot
β_c	slots per users to send IDs of the contention winners to the base station
T	number of time slots per period for the decentralized and mixed schemes
τ_c	channel uses for contention period in decentralized scheme

TABLE I

LIST OF THE VARIABLES USED IN THE MODEL.

be to assign a corresponding sequence. However we chose the one channel use of staying silent scheme for simplicity and because it will give an upper bound on the performance of the centralized scheme [3], thus the worst case for the improvement achieved with the decentralized and mixed schemes that follow. Notice that the signalling overhead β_c is a function of the total number of users admitted in the cell and the rate for the control signalling. It can thus pose some limitation in the number of users admitted in the cell (the time used for signalling in each slot should not be too big). The users that are selected then perform uplink training and then the BS serves them using ZF precoding with equal power among users, as explained in the previous Section. This procedure is illustrated in Fig. 1.

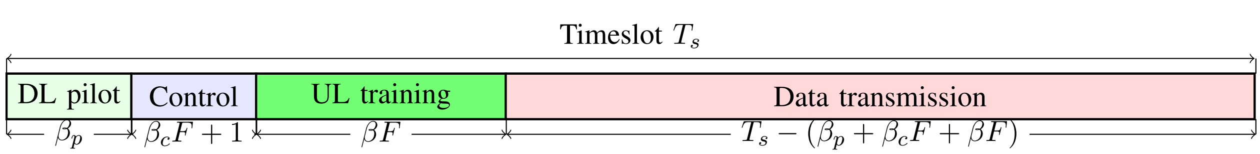

Fig. 1. Operation of the centralized scheme in a timeslot where F users have been scheduled

[3]Since the control channel must be decoded successfully at all times, a lower rate of one bit per channel use may be needed.

Alternatively, if the control phase is to remain constant irrespective of the number of scheduled users, the control phase will last for $\frac{K}{R_0}$ channel uses instead of $\beta_c F$, because a codeword of K bits, one for every user indicating if he is scheduled or not, should be used. This poses more severe restrictions to how many users the cell can support but having a control region of fixed duration may be desirable in practice e.g. for synchronization purposes.

The BS selects the set of users to be scheduled at every slot as the solution to the following problem:

$$\mathcal{F}(t) = \arg\max_{\mathcal{F}\in 2^{\mathcal{K}}} \left\{ (T_s - (1 + \beta_p + \beta_c F + \beta F)) \sum_{k\in\mathcal{F}} q_k(t)\mathbb{E}\left\{r_k(t)|F\right\} \right\}, \tag{6}$$

where the expectation is taken with respect to the joint probability distribution of the channels, as presented in detail in Section IV.

The advantage of this scheme is its (relative) simplicity. Indeed, the expectations of the rate for every user k given that $F-1$ other users are scheduled can be computed in advance and used at every slot. Furthermore, if the channels are i.i.d. among the users, it can be implemented by having F run from 1 to F_{max}, sort the users according to the values of $q_k(t)\mathbb{E}\left\{r_k(t)|F\right\}$ and select the F biggest every time. In the end select the configuration that gave the biggest expected weight. The overall computational complexity here is $O(F_{max}K\log_2(K))$.

The downside is that the actual realizations of the channels are ignored; for instance, a user chosen to be scheduled may actually have a very bad channel (i.e. channel with such a bad magnitude that cannot even support the smallest rate). This is a bigger problem when OFDMA is employed (as is actually done in modern systems e.g. LTE and Worldwide Interoperability for Microwave Access (WiMax)) because according to this scheme the same users will be scheduled for several carriers, so the frequency diversity in the fading is not exploited.

B. Decentralized policy with periodic signalling

To overcome the shortcomings of the centralized scheme, we first note that each user *can* know its actual channel realization, namely via downlink training. In this case, if each user knew its queue length (or the ranking of users based on the queue length) as well, we could exploit this knowledge and use, for example, techniques inspired by queue-based CSMA [30] or Fast CSMA [21] to find a schedule. Indeed, it has been recently shown that performing a CSMA with the backoff timer being a function of the product of the queue length times the actual rate supported by the channel realization can achieve throughput optimality in uplink systems with single carrier and single antenna fading channels [20] (under the assumption of continuous time for backoff). However, in our case the system model is more complicated because the users do not know their queue lengths and because of multiple antennas in the BS, more than one user can be scheduled simultaneously. Our proposed schemes, detailed in the next paragraph, are based on two ideas: (i) the BS periodically broadcasts the (suitably quantized) values of the queue lengths and (ii) the BS decides on the *number* of users to be scheduled and based on that lets the users contend using the queue length information they have and an estimate of their achievable rate based on channel state realization.

1) Algorithm description: To begin with, every T timeslots, that is at time $0, T, 2T, ..., mT, ...$ the BS broadcasts quantized versions of the queue lengths of the users at the beginning of this slot, i.e. broadcasts a quantization of the vector $\mathbf{q}(mT), m = 0, 1, ...,$ the restriction being that T is a finite number. The quantization of the queue lengths is discussed in detail in Section III-B2. In addition, the BS broadcasts the number $F(mT)$ of users to report the channel each timeslot for the next $T-1$ consecutive timeslots. No data transmission at this timeslot takes place in order to make broadcasting this information possible (with the BS adopting e.g. uniform precoding for transmission).

Denote now by $\hat{\mathbf{q}}(t) := \tilde{\mathbf{q}}(T\lfloor \frac{t}{T} \rfloor)$ to be the most recent information about their queue state that the users have. At each timeslot the BS sends a downlink pilot with duration β_p channel uses so that the users can estimate their channels and lets a period of τ_c channel uses for the users to contend for channel access. Assuming that contention can be done in continuous time and with signals of negligible duration (this assumption has been implicitly used in recent works dealing with Fast CSMA over fading channels [21], [19]), user k waits until time

$$\tau_k' = \frac{\tau_c'}{\hat{q}_k(t)\mathbb{E}\{r_k(t)|g_k(t), F(t)\}}. \tag{7}$$

In the above equations, the times are expressed in time units (i.e. ms or μs). The denominator is the latest broadcasted value of the queue length of this user times the expectation of the rate the user will get if it is scheduled, given its own channel realization (we have defined here $g_k(t) = ||\mathbf{h}_k(t)||^2$). This computation is detailed in the Section IV-A, and for an environment with Rayleigh fading (the case we examine here) can be done in a totally decentralized manner. Note that under this scheme, the F users with the biggest values of $\hat{q}_k(t)\mathbb{E}\{r_k(t)|g_k(t), F(t)\}$ are the ones that get actually scheduled.

Once the contention period is over, the F first users to have a signal broadcasted feed back their IDs in a TDMA manner, e.g. in the sequence in which they sent the contention signals (under the assumptions on continuous time contention and very short signals the users can know their place in the sequence). These users then perform uplink training and the BS transmits to them using Zero Forcing. The total time for transmission is then $(T_s - \beta_p - \tau_c - (\beta + \beta_c)F(m))$ channel uses. Illustration of a timeslot under this policy is given in Fig. 2.

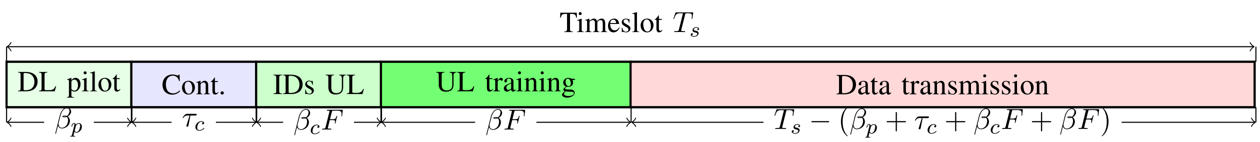

Fig. 2. Operation of the decentralized scheme in a data timeslot where the Base Station has signalled that F users are to be scheduled

What remains to be specified is the number of users to feed back in every period $\{mT+1, ..., (m-1)T\}$. As stated before, this decision is taken in the beginning of timeslot mT, based on the corresponding queue length information and channel statistics. Here we consider that the number of users to get scheduled is given as the solution of the following problem.

$$F^*(m) = \arg\max_{F=1,..,F_{\max}} \left\{ (T_s - \beta_p - \tau_c - (\beta + \beta_c)F)\, \mathbb{E}_{\mathbf{g}(t)} \left\{ \max_{\mathcal{F}:|\mathcal{F}|=F} \sum_{k\in\mathcal{F}} \hat{q}_k(mT)\mathbb{E}\left\{r_k(mT)|g_k(t), F\right\} \right\} \right\}. \tag{8}$$

In the above, the outer expectation is with respect to the joint distribution of the channel magnitudes while the inner expectation is with respect to the joint distributions of the channel directions (for the channels of all users in the system). A way to do these calculations is to use the recently proposed framework in [31], [32] for partial sums of order statistics of non-identically distributed random variables.

2) Queue length quantization scheme: As noted in the description of the algorithm, the BS broadcasts quantized versions of the queue lengths of the users. This quantization is essential because, as noted in the beginning of the Section the rate at which the BS can broadcast signalling information is R_0 bits per channel use; this means that if a slot is used for signalling, at most $T_s R_0$ bits can be sent to the users. Given that the BS should broadcast K queue lengths and how many users are to feed back, the number of bits b_q used for quantization of each queue should satisfy the following:

$$K b_q + \log_2 F_{max} \leq T_s R_0. \tag{9}$$

The above inequality poses limitations as to how many users can be supported by the system if it operates under this scheduling policy and a limitation to the accuracy of the queue length feedback for a given number of users in the cell. However, if multiple antennas and carriers are used, this limitation is not very severe.

We now detail the way the queue lengths are actually quantized for a given number of bits per user, b_q. To this end, we define

$$Q = \max\{TR_L, TA_{max}\} \tag{10}$$

and the intervals $[0, Q], [Q, 2Q], ..., [(p-1)Q, pQ], ...$ Note that Q is the biggest change that can possibly happen to a queue length after T slots (the queue will decrease by at most TR_L-if at every slot is served at the maximum rate with no further arrivals- and increase by at most by TA_{max}-if it s not served at all and has the maximum possible arrivals at each slot). Therefore, every T slots each queue length will be in one of the aforementioned intervals, and furthermore, given that at mT a queue length was at the p-th interval, at $(m+1)T$ it will be at intervals $p-1$, p or $p+1$. The idea is then to set the quantization interval to $[0, Q]$ at the beginning and inform each user every T slots if it stays in the same interval of one of the neighbouring intervals (this can be done at a signalling as low as 2 bits per user or even $1.5K$ bits in total). Then, the quantized queue length is sent, assuming uniform quantization within each interval using the rest of the signalling bits available. More concretely, if b_q bits per user are used for quantization 2 bits are used to denote the quantization interval and the rest are used to point to one of the 2^{b_q-2} levels within the interval. If $b_q = 2$, then the middle value of the interval is used as an estimation of the queue length. Note that this way, the difference between each of the broadcasted queue lengths, denoted by $\tilde{\mathbf{q}}(mT)$ and the corresponding real queue length from $\mathbf{q}(mT)$ is at most Q .

Finally, some remarks are in order. First, we have assumed that the control information broadcasted by the BS are always decodable at the user terminals. This is a rather frequent and reasonable assumption in the literature (i.e. that signalling and control data are transmitted without errors). In practice, control data are transmitted using low rate modulations and strong coding schemes. We can also always assume that we do downlink power control for the signalling information. In addition, if the number of carriers and antennas is high, the diversity of the system is so big that control data can always be received successfully even if some channels of the users are in deep fade (also, users with very bad average channel conditions should not be admitted into the cell). Accurate estimation of the users channel can be similarly argued by using high power pilots.

C. Mixed Policy: Combining Centralized and Decentralized Schemes

The decentralized approach to the user selection problem should lead to users with better channel conditions being selected in general, however it requires some extra time overhead for the contention period. In some cases, some queues may be empty or a few queues may be much bigger than others. If this happens it may be better for the users with the much bigger queues to be scheduled for the $T-1$ timeslots without any additional signalling overhead. Note that since the same users get scheduled every $T-1$ consecutive slots, no overhead for their IDs must be used either. The operation of this scheme in a data timeslot is illustrated in Fig. 3 [4]. We will refer to this scheme as ”periodic centralized policy” for the rest of the paper.

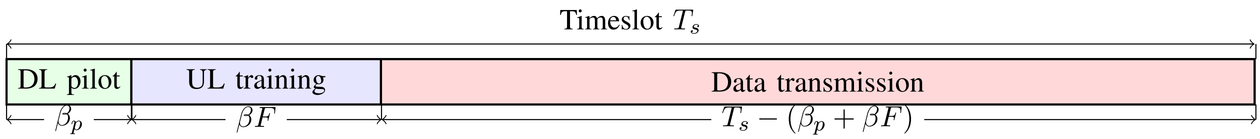

Fig. 3. Operation of centralized scheme with periodic user selection in a data timeslot where the Base Station has signalled a set of F users to be scheduled

The set of users to schedule according to this periodic centralized policy is set as

$$\mathcal{F}^*(m) = \arg\max_{\mathcal{F}\in 2^{\mathcal{K}}} \left\{ (T_s - (\beta_p + \beta F)) \sum_{k\in\mathcal{F}} \hat{q}_k(mT)\mathbb{E}\left\{r_k(t)|F\right\} \right\}. \tag{11}$$

The mixed policy here is, therefore, that the BS at every slot $t = mT, m = 0, 1, ...$ decides that, for the next $T-1$ slots, either the decentralized policy will be used, with the optimal number of users to get scheduled as given by eq. (8) or select a set $\mathcal{F}(m)$ of users to schedule according to the maximization in eq. (11). This decision is made based on which of the two policies will maximize the quantity $\mathbb{E}\left\{\sum_{k=1}^{K} \hat{q}_k(mT)(T_s - \tau(t))r_k(t) \middle| \mathbf{q}(mT)\right\}$, with the expectation taken over the channel distributions (which affect the outcomes of the policies). More concretely, let $F^*(m)$ be given by (8) and $\mathcal{F}^*(m)$ be given by (11). Then, the BS selects the decentralized scheme with $F^*(m)$ users to get scheduled if

$$\begin{aligned}
&(T_s - \beta_p - \tau_c - (\beta + \beta_c))\mathbb{E}_{\mathbf{g}(t)}\left\{ \max_{\mathcal{F}:|\mathcal{F}|=F^*(m)} \sum_{k\in\mathcal{F}} \hat{q}_k(mT)\mathbb{E}\left\{r_k(mT)|g_k(t), F\right\} \right\} \\
&\leq (T_s - (\beta_p + \beta F)) \sum_{k\in\mathcal{F}^*(m)} \hat{q}_k(mT)\mathbb{E}\left\{r_k(t)||\mathcal{F}^*(m)|\right\}
\end{aligned}$$

and the centralized scheme with the users in $F^*(m)$ otherwise.

This policy should have a bigger stability region than either of the two policies mentioned before in this Section, since it essentially combines the ideas behind both. However, we would like to point out that the decentralized policy alone can achieve points that can not be achieved by the centralized (or the

[4]The DL pilot is still needed for the users to set the power of their training sequences such that the SNR in the reverse link suffices for perfect estimation

version of the centralized policy used in the mixed scheme) policy. Normally while applying the mixed policy, the decentralized scheme should be the one most frequently selected, with the centralized mode selected in special cases (for example, when few queues are not empty or there is a queue that is much bigger than all the others). It needs 1 additional bit of signalling compared to the other policies, in order to inform the users if the centralized or decentralized scheme will be employed. In addition, the BS needs to broadcast the number F of users to be scheduled in the decentralized policy if used ($\log_2(F_{max})$ bits) or the users to get scheduled in case the centralized policy is used ($\min\{K, F_{max}\log_2 K\}$ bits). Since $F_{max} \leq K$, there must hold

$$Kb_q + \min\{K, F_{max}\log_2 K\} \leq T_s R_0, \tag{12}$$

which gives a bound on the number of users to be admitted to the cell in order to operate under the mixed policy.

IV. Calculation of Parameters and Stability Results

In this Section we give the expressions for the SNR (and subsequently rate) distributions. Also, we give some useful lemmas about stability of systems where control decisions are done periodically and not in a slot-per-slot basis.

A. Calculation of average rates

The decentralized scheme requires that every user should calculate their average rate given their current channel state realization, as seen by eq. (7). Indeed, since the system operates under isotropic fading directions, we can calculate the probability distribution over the other users' channels and zero forcing precoding of a user's SNR given its channel. We have:

Proposition 2. *The probability that the received SNR at user k exceeds s given its channel strength and that this user and other $F-1$ users are scheduled is given by*

$$\begin{aligned}\mathbb{P}\{SNR_k(t) > S | g_k(t), F\} &= \mathbb{P}\{SNR_k(t) > S | \mathbf{h}_k(t), F\} \\ &= 1 - I_B\left(\frac{F\sigma^2}{Pg_k(t)}S; N-F+1, F-1\right).\end{aligned} \tag{13}$$

This distribution is with respect to the direction of the channel of user k and the channels of the other users that get scheduled.

Proof. Please refer to Appendix A for the proof. □

From the above result and the proof we can see that only the magnitude and not the direction of the channel realization comes into the equation. In addition, the long-term statistics of the other users do not play any part in the computation either. Intuitively these remarks are due to the isotropic direction of the channel vectors. Indeed, since we are considering ZF precoding, the loss of SNR comes due to the fact that the channels are not orthogonal, therefore the demand of causing zero interference can constrain a lot the precoder selection. Since the directions are isotropic, knowledge of one channel direction does not imply anything about how nullspace of the other users should behave. A consequence of these remarks is that a user can actually calculate this distribution (and hence the average rate he will get given its channel) with only the knowledge that the whole system operates under Rayleigh fading.

In the centralized scheme, where the BS does not have knowledge of the magnitude realization of the channels, the probability that the SNR of user k exceeds S when $F-1$ other users are scheduled is the following [15], [33][5]

Proposition 3. *Given a number of users to be scheduled, F, the probability that the SNR of user k exceeds S is given as*

$$\mathbb{P}\left\{SNR_k(t) > S|F\right\} = 1 - \frac{\gamma\left(\frac{F\sigma^2}{\bar{g}_k P}S; N-F+1\right)}{\Gamma(N-F+1)}.$$

From the results of Propositions 2 and 3 we can find the average rates conditioned on the channel realization of user k and the expected rate of this user without knowing the channel realization, respectively. More concretely, if we define

$$L_{0,k}(t,F) = \max\left\{l \in 1, ..., L : S_l \leq \frac{g_k(t)P}{F\sigma^2}\right\}, \tag{14}$$

i.e. the index of the highest rate that could be supported by user k if he is scheduled and his channel is orthogonal to the channels of the other $F-1$ scheduled users, we have:

$$\begin{aligned}\mathbb{E}\left\{r_k(t)|F, g_k(t)\right\} &= \sum_{l=1}^{L_{0,k}(t,F)} R_l \mathbb{P}\left\{S_l \leq SNR_k(t) < S_{l+1}|F, g_k(t)\right\} \\ &= \sum_{l=1}^{L_{0,k}(t,F)} R_l \left(\mathbb{P}\left\{SNR_k(t) \geq S_l|F, g_k(t)\right\} - \mathbb{P}\left\{SNR_k(t) \geq S_{l+1}|F, g_k(t)\right\}\right).\end{aligned} \tag{15}$$

[5]It can also be calculated by integrating (13) for $g_k(t)$ from zero to infinity. Although we could not obtain the exact form given in Proposition 3, numerical results indicate that the numerical values are the same with either method. We will use the latter expression since it is in a closed form

and

$$\mathbb{E}\left\{r_k(t)|F\right\} = \sum_{l=1}^{L} R_l \left(\mathbb{P}\left\{SNR_k(t) \geq S_l|F\right\} - \mathbb{P}\left\{SNR_k(t) \geq S_{l+1}|F\right\}\right). \tag{16}$$

Notice that, since the statistics are assumed known, the rates in (16) can be calculated only once and used by the BS for the centralized scheme.

B. Stability results

In this subsection we are interested in deriving some stability results for the system under the policies where slots $\{0, T, T+1, .., mT, ..\}$ are used for signalling and/or broacasting of the queue lengths. First, we define the queueing system that results when examining the original system at time instances $0, T, ...,$ i.e. at the beginning of the slot in which the broadcasting takes place. Formally:

$$\tilde{\mathbf{q}}(m) := \mathbf{q}(mT), m = 0, 1, 2, ... \tag{17}$$

The equations regarding the evolution of this system are, thus $\forall k \in \{1, ..., K\}$:

$$\tilde{q}_k(m+1) = \tilde{q}_k(m) + \sum_{t=0}^{T-1} a_k(mT+t) - \sum_{t=1}^{T-1} z_k(mT+t)\mu_k(mT+t) + \sum_{t=1}^{T-1} y_k(mT+t) \tag{18}$$

where $z_k(t)$ is the indicator function, set to 1 if user k is scheduled in timeslot t and zero otherwise, $\mu_k(t)$ is the total number of bits assigned for transmission to user k at timeslot t (that is $\mu_k(t) = r_k(t)(T_s - \tau(t))$ (recall that $\tau(t)$ is the total time of the slot used for pilot transmission, coordination and training), and $y_k(mT+t) = [z_k(mT+t)\mu_k(mT+t) - q_k(mT+t)]^+$ the number of "wasted" bits if the offered rate at one timeslot is bigger than the available bits in the buffer. Note that the process $\tilde{\mathbf{q}}(m)$ is a discrete time Markov chain evolving on a countable state space. The following result holds:

Lemma 4. *The system $\mathbf{q}(t)$ is strongly stable if and only if the system $\tilde{\mathbf{q}}(m)$ is strongly stable.*

Proof. Assume first that $\mathbf{q}(t)$ is strongly stable. We have $\frac{1}{M}\sum_{m=0}^{M-1}\mathbb{E}\{\tilde{q}_k(m)\} \leq T\frac{1}{MT}\sum_{\tau=0}^{T(M-1)}\mathbb{E}\{q_k(\tau)\}$ therefore

$$\lim_{M\to\infty}\sup\frac{1}{M}\sum_{m=0}^{M-1}\mathbb{E}\{\tilde{q}_k(m)\} \leq T\lim_{M\to\infty}\sup\frac{1}{MT}\sum_{\tau=0}^{T(M-1)}\mathbb{E}\{q_k(\tau)\} < +\infty,$$

where the second inequality follows from the assumption. Therefore $\tilde{\mathbf{q}}(m)$ is indeed strongly stable. Assume now that $\tilde{\mathbf{q}}(m)$ is strongly stable. We can write

$$\begin{aligned}
&\lim_{t\to+\infty}\sup\frac{1}{t}\sum_{\tau=0}^{t-1}\mathbb{E}\{q_k(\tau)\} = \lim_{M\to+\infty}\sup\frac{1}{MT}\sum_{\tau=0}^{MT-1}\mathbb{E}\{q_k(\tau)\} \\
&= \lim_{M\to\infty}\sup\left(\frac{1}{MT}\sum_{m=0}^{M-1}\mathbb{E}\{q_k(mT)\} + \frac{1}{MT}\sum_{m=0}^{M-1}\sum_{\tau'=1}^{T-1}\mathbb{E}\{q_k(mT+\tau')\}\right).
\end{aligned} \tag{19}$$

Since $\tilde{\mathbf{q}}(m)$ is strongly stable and $\tilde{\mathbf{q}}(m) = \mathbf{q}(mT)$, there exists some $0 < C_0 < \infty$ such that $\forall k \in \{1,..,K\}$:

$$\lim_{M\to\infty} \sup \frac{1}{MT} \sum_{m=0}^{M-1} \mathbb{E}\{q_k(mT)\} \leq \frac{C_0}{T}. \tag{20}$$

Also, note that $\forall \tau' \in \{1,..,T-1\}$, $\forall m = 0,1,....$ it holds $q_k(mT+\tau') \leq q_k(mT) + \tau' A_{max}$. This implies that, $\forall m = 0,1,2,...$ we have $\sum_{\tau'=1}^{T-1} \mathbb{E}\{q_k(mT+\tau')\} \leq \mathbb{E}\{q_k(mT)\} + \frac{(T-1)T}{2} A_{max}$. Replacing we get

$$\begin{aligned}\lim_{t\to+\infty} \sup \frac{1}{t} \sum_{\tau=0}^{t-1} \mathbb{E}\{q_k(\tau)\} &\leq \lim_{M\to\infty} \sup \left(\frac{2}{MT} \sum_{m=0}^{M-1} \mathbb{E}\{q_k(mT)\} + \frac{T(T-1)}{2} A_{max} \right) \\ &\leq \frac{2C_0}{T} + \frac{T(T-1)}{2} A_{max} < \infty,\end{aligned}$$

which implies that $\mathbf{q}(t)$ is stable. □

The above Lemma implies that a throughput optimal policy for the process $\tilde{\mathbf{q}}(m)$ should be also throughput optimal for the original system.

Define now $V(\mathbf{x}) = \frac{1}{2}\sum_{k=1}^{K} x_k^2$ a Lyapunov function and $\Delta V(\mathbf{x})$ its drift, i.e.

$$\Delta V(\mathbf{x}) = \mathbb{E}\left\{ V(\tilde{\mathbf{q}}(m+1)) - V(\tilde{\mathbf{q}}(m)) | \tilde{\mathbf{q}}(m) = \mathbf{x} \right\}. \tag{21}$$

The expectation is over the arrival and channel processes as well as the possibly randomized feedback policy (i.e. the set of F users that feed back). Then the following holds for the drift of the sampled system:

Lemma 5. *The drift of the quadratic Lyapunov function for the system $\tilde{\mathbf{q}}(m)$ under a scheduling policy π is upper bounded as follows (note that the number of users to feed back, F is included in the policy):*

$$\Delta V_\pi(\tilde{\mathbf{q}}(m)) \leq \tilde{B} + T \sum_{k=1}^{K} \tilde{q}_k(m)\lambda_k - (T-1)\sum_{k=1}^{K} \tilde{q}_k(m) \mathbb{E}\left\{ \tilde{\mu}_k^\pi(m) | \tilde{\mathbf{q}}(m) \right\} \tag{22}$$

where $\tilde{\mu}_k^\pi(m)$ is the rate of user k in any of the slots $mT+1,...,mT+(T-1)$ for a given channel state realization and outcome of the policy π (i.e. set of users actually fed back), $\tilde{B}$ is a constant depending only on the system parameters. The expectation is taken over the joint distribution of the channels and possible randomization of the policy.

Proof. Please refer to Section B in the Appendix of this paper for the proof. □

As a final remark, we note that the same stability results with Lemma 4 hold for the system operating under the centralized policy as well. That is, the system operating under the centralized policy is stable if and only if the system that results from sampling the queue lengths of the original at timeslots mT is strongly stable;the proof is essentially the same as the proof of Lemma 4.

V. A Special Case: The 2-User MISO BC With Single Rate

In this Section we will consider a simple case, namely a system with $K = 2$ users with identical channel statistics (i.i.d. Rayleigh with mean power gain $\bar{g}$) where a user gets rate R bits per channel use if SNR exceeds the threshold $\hat{S}$ and zero otherwise. This setting is of interest as the stability regions admit easy mathematical expressions and can be plotted, thus giving some insight on the outcomes of the policies.

To begin with, we define some parameters to be used frequently in the sequel. Define the probabilities that a user's SNR exceeds the threshold if only one or both users are scheduled as $\bar{p}(1)$ and $\bar{p}(2)$, respectively. Since the channels are statistically identical for both users, these probabilities are the same for any of them. The numerical values of these probabilities are given by

$$\begin{aligned} \bar{p}(1) &= 1 - \frac{\gamma\left(\frac{\sigma^2 \hat{S}}{P\bar{g}}; N\right)}{\Gamma\left(N\right)} \\ \bar{p}(2) &= 1 - \frac{\gamma\left(\frac{2\sigma^2}{\bar{g}P}\hat{S}; N-1\right)}{\Gamma(N-1)}. \end{aligned} \tag{23}$$

These expressions are derived by specializing the results in Sections II and IV for $F = 2$ and $\bar{g}_k = \bar{g}$. The system parameters are the same as in the original description: Downlink training requires β_p channel uses and uplink training requires pilots of length β channel uses for each user.

We now turn to characterizing the form and stability region of each policy. The general shapes of the policies are sketched in Fig. 4 that follows, and Fig. 5 in the next Section depicts a numerical example with specific values of the parameters.

A. *Centralized policy*

In this policy, in every slot t the transmitter selects either one or both the receivers to be scheduled. In the latter case, there is an overhead of $2\beta_c$ channel uses to broadcast the IDs of the two users and in the former, of $\beta_c + 1$ to broadcast the ID of the scheduled user and a signal that the control period is over.

The expected rate that a user gets if both users are scheduled or if this user only is scheduled at timeslot t is given by

$$\bar{\mu}_c(2) = (T_s - (\beta_p + 2\beta_c + 2\beta))\bar{p}(2)R \tag{24}$$

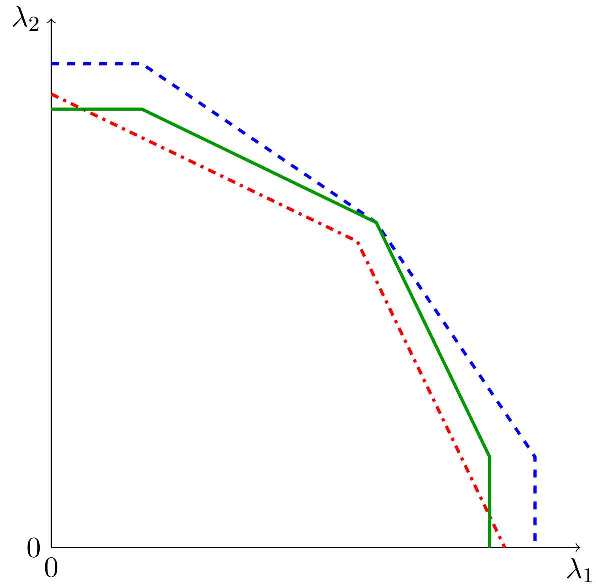

Fig. 4. An illustration of the stability regions of the centralized (red dashed-dotted line), decentralized (green solid line) and mixed (blue dashed line) policies in the 2 user system with single rate level

and by

$$\bar{\mu}_c(1) = (T_s - (1 + \beta_p + \beta_c + \beta))\bar{p}(1)R, \tag{25}$$

respectively. The set to be scheduled at slot t is then chosen at the beginning of this slot by the rule that follows:

$$\mathcal{F}_c(t) = \begin{cases} \{1,2\}, \text{ if } (q_1(t) + q_2(t))\bar{\mu}_c(2) \geq \max\{q_1(t), q_2(t)\}\bar{\mu}_c(1) \\ \mathcal{F}_c(t) = \{\arg\max\{q_1(t), q_2(t)\}\}, \text{otherwise} \end{cases}$$

The stability region of the system under this policy is characterized as follows:

Theorem 6. *The stability region of the centralized policy in the 2 i.i.d. user case with one rate level is*

$$\Lambda_c^{(2)} = \mathcal{CH}\Big\{(0, \bar{\mu}_c(1)), (\bar{\mu}_c(2), \bar{\mu}_c(2)), (\bar{\mu}_c(1), 0)\Big\}. \tag{26}$$

Proof. First we will prove that the region in the statement of the Theorem is indeed achievable by the centralized policy. To this end, consider a randomized policy that at the beginning of timeslot t selects the set $\mathcal{F}$. of users to serve with probability $\pi_{\mathcal{F}}$. In our setting the set $\mathcal{F}$ can be one of $\{1\}, \{2\}, \{1, 2\}$.

The achievable stability region by this policy is

$$\begin{array}{c}\lambda_1 < \pi_{\{1\}}\bar{\mu}_c(1) + \pi_{\{1,2\}}\bar{\mu}_c(2) := \hat{\mu}_1 \\ \lambda_1 < \pi_{\{2\}}\bar{\mu}_c(1) + \pi_{\{1,2\}}\bar{\mu}_c(2) := \hat{\mu}_2\end{array}, \tag{27}$$

with $\pi_{\{1\}} + \pi_{\{2\}} + \pi_{\{1,2\}} \leq 1$ and $0 \leq \pi_{\mathcal{F}} \leq 1$. This is exactly the algebraic characterization of the set (26).

Define now the quadratic Lyapunov function $V(\mathbf{x}) = x_1^2 + x_2^2$. Its drift

$$\Delta V(\mathbf{q}) = \mathbb{E}\left\{V(\mathbf{q}(t+1)) - V(\mathbf{q}(t))|\mathbf{q}(t) = \mathbf{q}\right\}$$

can be shown to be bounded as (with some positive constant B)

$$\Delta V(\mathbf{q}) \leq B - \sum_{k=1}^{2}(q_k(t)\mathbb{E}\{\mu_k(t)\} - q_k(t)\lambda_k) \leq B - \sum_{k=1}^{2} q_k(t)(\hat{\mu}_k - \lambda_k).$$

The second inequality follows by the definition of our centralized policy. From (27) it follows that $\forall \lambda \in \Lambda_c^{(2)}, \exists \epsilon > 0$ such that $\Delta V(\mathbf{q}) \leq B - \epsilon \sum_{k=1}^{2} q_k(t)$, hence the system under the centralized policy is indeed stable for all mean arrival rates in the asserted region.

We then need to prove the converse, that is, if a centralized policy achieves stability, then the mean arrival rate lies in (the interior of) the region given by (26). Indeed, assume that the system is stable for a mean arrival rate vector λ. The centralized policy depends only on the queue lengths at the beginning of slot t, which we denote by $\mathcal{F}(\mathbf{q})$. The assumptions, thus, on the channel and arrival processes make the system a discrete time Markov chain with a single communicating class. In this case, stability implies the existence of an invariant distribution $\pi(\mathbf{q})$. The mean service rate user k gets is then equal to

$$\lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mu_k(\tau) = \bar{\mu}_c(1) \sum_{\mathbf{q}\in\mathbb{Z}_+^2:\mathcal{F}(\mathbf{q})=\{k\}} \pi(\mathbf{q}) + \bar{\mu}_c(2) \sum_{\mathbf{q}\in\mathbb{Z}_+^2:\mathcal{F}(\mathbf{q})=\{1,2\}} \pi(\mathbf{q}).$$

Since the sums are probabilities themselves, we can see that the service rate have the same form as in (27). Also, since the system is assumed stable, there should be $\lambda_k < \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mu_k(\tau)$. From these we conclude that $\lambda \in \Lambda_c^{(2)}$, completing the proof. □

The stability region for the centralized scheduling algorithm looks like a trapeze with corner points $(0,0), (0,\bar{\mu}_c(1)), (\bar{\mu}_c(2), \bar{\mu}_c(2)), (\bar{\mu}_c(1), 0)$ if it holds that $\bar{\mu}_c(1) < 2\bar{\mu}_c(2)$ and like a triangle with corners at $(0,0), (0,\bar{\mu}_c(1)), (\bar{\mu}_c(1), 0)$ otherwise. This follows from the fact that in the latter case only one user will get scheduled.

B. Decentralized Policy

Let the time for contention be τ_c channel uses. For consistency, the contention period will be present even if $F = 2$, i.e. when both users are to be scheduled (this will be improved by the mixed policy). From the results of Section IV-B it suffices to look at the stability of the system examined in the beginning of each signalling slot $mT, m = 0, 1, ...$

As described in Section III-B, the BS broadcasts the quantized queue lengths $\hat{q}_1(mT), \hat{q}_2(mT)$ at time mT. In the particular case with 2 users, the decision taken by the BS is to either select both users ($F(m) = 2$) or signal that one user will be selected ($F(m) = 1$) and the user who will be scheduled is the one with the lowest timer from eq. (7).

1) Contention procedure: If $F(m) = 1$, in each of these slots the receivers are given a contention period of τ_c channel uses to decide which one is to be scheduled based on the (quantized and outdated) queue length information they have and the realization of their channels. This can be done using a contention scheme, assuming contention in continuous time e.g. like [21], where each user waits until time $\frac{\tau_c}{\hat{q}_k(m)r_k(t)}$: if both have the same timer, e.g. the user with the smallest ID is scheduled. Another alternative, that can be used thanks to our model, is to divide the contention period into minislots (TDMA manner) where each receiver sends a signal in its corresponding minislot if its SNR is above the threshold $\hat{S}$. If both receivers send a signal, in their corresponding minislots, then the receiver with the largest broadcasted queue length gets scheduled for training (this analysis/comparison can be done independently by each receiver since the queue lengths of all receivers are broadcasted). Otherwise, if only one user sends a signal in a minislot, then this user will be scheduled for training. Then, the user to be scheduled sends its ID to the BS , taking β_c channel uses, and trains. Using the above ”decentralized” procedure, the user that will eventually get served in the slot will be the one with the maximum product of quantized queue length at mT times achievable rate. Due to our model here, denoting $SNR_k^{(1)}(t) = \frac{Pg_k(t)}{\sigma^2}$, the user to be scheduled will be

- If $\forall k = 1, 2$ holds $SNR_k^{(1)}(t) > \hat{S}$, then $k^*(t) = \arg\max[\hat{q}_1(mT), \hat{q}_2(mT)]$
- The user for which $SNR_k^{(1)}(t) > \hat{S}$ otherwise

The scheduled receiver will always be given rate of R bits per channel use, except in the case where no one has sufficiently high SNR, in which no receiver can be scheduled anyway. Defining the permutation $k(1), k(2)$, where $\hat{q}_{k(1)}(mT) \geq \hat{q}_{k(2)}(mT)$, the average service rates of these users under $F = 1$ for the next $T-1$ slots are

$$\begin{aligned}\bar{\mu}^{d,(1)}_{k(1)}(t) &= (T_s - (\beta_p + \tau_c + \beta))\bar{p}(1)R := \bar{\mu}_d(1)\\ \bar{\mu}^{d,(1)}_{k(2)}(t) &= (T_s - (\beta_p + \tau_c + \beta))\bar{p}(1)(1-\bar{p}(1))R.\end{aligned} \tag{28}$$

2) F(m)=2: Both users train just after the coordination period. The average rate per slot for each user in this case will be

$$\bar{\mu}_d(2) = (T_s - (\beta_p + \tau_c + 2\beta))\bar{p}(2)R \tag{29}$$

Based on the above, the transmitter decides at $t = mT$ the number of users to get scheduled for the next $T-1$ slots by:

$$F(m) = \begin{cases} 2, \text{if } \hat{q}_{k(1)}(m) + \hat{q}_{k(2)}(m))\bar{\mu}^{d,(1)}_{k(1)}(t) \geq \\ (\hat{q}_{k(1)}(m) + \hat{q}_{k(2)}(t)(1-\bar{p}(1)))\bar{\mu}^{d,(1)}_{k(1)}(t) \\ 1, \text{otherwise} \end{cases}$$

In the case of $F(m) = 1$, the contention procedure is followed.

The stability region of this policy is described as follows :

Theorem 7. *The stability region of the decentralized scheme for the* 2 *user MISO broadcast system with a single rate level is*

$$\begin{aligned}\Lambda^{(2)}_d = \left(1 - \frac{1}{T}\right)\mathcal{CH}\Big\{&(0, \bar{\mu}_d(1)), (\bar{\mu}_d(1)(1-\bar{p}(1)), \bar{\mu}_d(1)),\\ &(\bar{\mu}_d(2), \bar{\mu}_d(2)), (\bar{\mu}_d(1), \bar{\mu}_d(1)(1-\bar{p}(1))), (\bar{\mu}_d(1), 0)\Big\}\end{aligned}$$

Proof. The proof consists in four parts. For the first two parts we compute the stability region for policies that select all the time $F = 2$ and $F = 1$. Then we prove the convex combination of the two is achievable by the decentralized policy and we finish by proving the converse. In the proof we examine the system $\tilde{\mathbf{q}}(m) = \mathbf{q}(mT)$, since from Lemma 4 stability of this system is sufficient for stability of the original queueing system.

Step 1: We first find the stability region if $F = 2$ for every signalling slot mT. In this case, the mean rate a user gets for each data slot is $\bar{\mu}_d(2)$. Thus, for the system $\tilde{\mathbf{q}}(m)$, the mean arrival rate for user k is $T\lambda_k$ and the mean service rate is $(T-1)\bar{\mu}_d(2)$, thus the stability region here is $\lambda_k < \frac{T-1}{T}\bar{\mu}_d(2), \forall k = 1, 2$.

Step 2: We then find the stability region if $F = 1$ in every signalling slot. We define a hypothetical policy where a the BS knows from the start of a data slot the achievable rates for both users and, based on this knowledge, chooses one of the two users to train and get scheduled, probably at random (while keeping the same time for data transmission in the slot as the corresponding in the decentralized policy). More concretely, only one user can support the rate R then this user should be scheduled, otherwise if both support the rate R then user 1 gets scheduled with some probability π_1 and user 2 with a probability π_2. In this case, taking into account the model for the system $\tilde{\mathbf{q}}(m)$ the mean arrival rates λ_1, λ_2 that can be supported by the system are the one for which there exist probabilities π_1, π_2 such that (the quantities in the right hand side are the mean rates given to each user):

$$\begin{aligned} T\lambda_1 &< (T-1)\left((1-\bar{p}(1))\bar{\mu}_d(1) + \pi_1 \bar{p}(1)\bar{\mu}_d(1)\right) := (T-1)\hat{\mu}_{d,1} \\ T\lambda_2 &< (T-1)\left((1-\bar{p}(1))\bar{\mu}_d(1) + \pi_2 \bar{p}(1)\bar{\mu}_d(1)\right) := (T-1)\hat{\mu}_{d,2} \\ & \qquad\qquad\qquad\qquad\qquad\qquad 0 \le \pi_1 + \pi_2 \le 1. \end{aligned} \tag{30}$$

This is (for λ) the algebraic representation of the convex hull of the points $(0, \frac{T-1}{T}\bar{\mu}_d(1))$, $(\frac{T-1}{T}(1-\bar{p}(1))\bar{\mu}_d(1), \frac{T-1}{T}\bar{\mu}_d(1))$, $(\frac{T-1}{T}\bar{\mu}_d(1), \frac{T-1}{T}(1-\bar{p}(1))\bar{\mu}_d(1))$, $(\frac{T-1}{T}\bar{\mu}_d(1), 0)$. Now assume a vector λ inside this region denoting $\hat{\mu}_k$ the mean rate of user k under a hypothetical policy such that the system is stable. From Lemma 5 we have that the drift of the quadratic Lyapunov function here is

$$\Delta V_\pi(\tilde{\mathbf{q}}(m)) \le \tilde{B} + T\sum_{k=1}^{K} \tilde{q}_k(m)\lambda_k - (T-1)\sum_{k=1}^{K} \tilde{q}_k(m)\mathbb{E}\left\{\tilde{\mu}_k^{\pi}(m)|\tilde{\mathbf{q}}(m)\right\}$$

Recall that $\hat{\mathbf{q}}(m)$ is vector containing the quantized versions of the queue lengths at the beginning of the signalling slot, therefore $\tilde{q}_k(m) - Q \le \hat{q}_k(m) \le \tilde{q}_k(m) + Q$. Also that the decentralized policy here selects the user with the maximum product of rate times quantized queue length, thus we get

$$\begin{aligned} \Delta V_\pi(\tilde{\mathbf{q}}(m)) &\le \tilde{B} + T\sum_{k=1}^{K} \tilde{q}_k(m)\lambda_k - (T-1)\sum_{k=1}^{K} \tilde{q}_k(m)\mathbb{E}\left\{\tilde{\mu}_k^{d}(m)|\tilde{\mathbf{q}}(m)\right\} \\ &\le \tilde{B} + TQ\sum_{k=1}^{2}\lambda_k + (T-1)KR_{max}Q + T\sum_{k=1}^{K}\hat{q}_k(m)\lambda_k \\ &- (T-1)\sum_{k=1}^{K}\hat{q}_k(m)\mathbb{E}\left\{\tilde{\mu}_k^{d}(m)|\tilde{\mathbf{q}}(m)\right\} \\ &\le \tilde{C} + \sum_{k=1}^{2}\hat{q}_k(m)\left(\lambda_k - \hat{\mu}_k\right) \\ &\le \tilde{C} - \epsilon\sum_{k=1}^{2}\hat{q}_{d,k}(m). \end{aligned}$$

The drift is negative for $\sum_{k=1}^{2}\hat{q}_k(m) > \tilde{C}/\epsilon \implies \sum_{k=1}^{2}\hat{q}_k(m) > 2Q + \tilde{C}/\epsilon$, thus the system under the decentralized policy achieves indeed the stability region given by (30).

Step 3: Here we prove that $\Lambda_d^{(2)}$ is achievable by the decentralized policy. Consider a randomized policy between $F=1$ and $F=2$ with probabilities $\pi(F=1)$ and $\pi(F=2)$ (independent on anything), respectively and the randomized hypothetical policy for the case of $F=1$ given in the above paragraph. The mean arrival rates supported under this policy should then be such that there exist these probabilities while satisfying the conditions

$$\begin{aligned} & T\lambda_k < (T-1)\left(\pi(F=1)\left((1-\bar{p}(1))\bar{\mu}_d(1)+\pi_k\bar{p}(1)\bar{\mu}_d(1)\right)+\pi(F=2)\bar{\mu}_d(2)\right) \\ & := (T-1)\hat{\mu}_{d,k}, k=1,2 \\ & 0 \leq \pi(F=1)+\pi(F=2) \leq 1 \\ & 0 \leq \pi_1+\pi_2 \leq 1. \end{aligned} \tag{31}$$

The region defined by the above equations is the convex hull of the two regions defined by (30) and (28), thus the set in the statement of the theorem. Under the proposed policy, using the same calculations as above, the drift of the quadratic Lyapunov function becomes

$$\begin{aligned} \Delta V_\pi(\tilde{\mathbf{q}}(m)) &\leq \tilde{C} + T\sum_{k=1}^{2}\hat{q}_k(m)\lambda_k - (T-1)\sum_{k=1}^{2}\hat{q}_k(m)\mathbb{E}\{\mu_k^\pi(m)\} \\ &\leq \tilde{C} + \sum_{k=1}^{2}\hat{q}_k(m)(T\lambda_k-(T-1)\hat{\mu}_{d,k}), \end{aligned}$$

where the second inequality follows from the fact that by definition of the policy the quantity $\sum_{k=1}^{2}\hat{q}_k(m)\mathbb{E}\{\mu_k^\pi(m)\}$ is maximized. Then, by (31) we get that for some $\epsilon>0$, $\Delta V_\pi(\tilde{\mathbf{q}}(m)) < \tilde{C} - \epsilon\sum_{k=1}^{2}\hat{q}_k(m)$, which is negative for (as above) $\sum_{k=1}^{2}\tilde{q}_k(m) \geq 2Q + \frac{\tilde{C}}{\epsilon}$, therefore the decentralized policy can support any rate of the (interior of the) set in the statement of the Theorem.

Step 4: To finish, we prove the converse, that is any mean arrival rate vector λ for which the system under the decentralized policy is stable lies in the interior of the set $\Lambda_d^{(2)}$. We have that (i) the number of users scheduled by the decentralized policy depends on the quantized queue lengths and that the user scheduled for $F=1$ depends on the quantized queue length and the channel state realizations and (ii) the quantized queue lengths are functions of the actual queue lengths at the start of slot mT, the system is an aperiodic markov chain with countable state space ($\mathbb{Z}_+^2$) and a single communicating class, thus strong stability implies ergodicity of the chain, therefore existence of an invariant distribution $\pi(\mathbf{q})$. The mean service rate a user 1 gets is therefore

$$\lim_{M\to\infty}\frac{1}{M}\sum_{m=0}^{M-1}\sum_{t=mT+1}^{mT+T-1}\mu_1(t)=(T-1)\Bigg(\bar{\mu}_d(1)\sum_{\mathbf{q}\in\mathbb{Z}_+^2:F(\mathbf{q})=1,\hat{q}_1\geq\hat{q}_2}\pi(\mathbf{q})$$
$$+(1-\bar{p}(1))\bar{\mu}_d(1)\sum_{\mathbf{q}\in\mathbb{Z}_+^2:F(\mathbf{q})=1,\hat{q}_1<\hat{q}_2}\pi(\mathbf{q})+\bar{\mu}_d(2)\sum_{\mathbf{q}\in\mathbb{Z}_+^2:F(\mathbf{q})=2}\pi(\mathbf{q})\Bigg)$$

and similar for user 2. By assumption the system is stable therefore $T\lambda_k < \lim_{M\to\infty}\frac{1}{M}\sum_{m=0}^{M-1}\sum_{t=mT+1}^{mT+T-1}\mu_k(t)$ for both users. Combining the above, and since the summations in the right hand side of the mean rate expression are probabilities, we get that $\lambda\in\Lambda_d^{(2)}$. □

In the special case of $K=2$ we consider here, if the overhead for the contention τ_c is 2 channel uses then the scheme can be implemented even dropping the assumption of continuous time for contention. Indeed, the first of the two channel uses can be dedicated for the user with the biggest quantized queue length and the second for the other user (since the queue lengths are broadcasted, each users knows the queue length of the other), with the ranking be based on the user ID in case of a tie. Then, for $F=1$, the first user in the ranking sends a signal if its SNR exceeds the threshold and remains silent otherwise, same for the second user. Note that based on the same idea, we can have even $\tau_c=1$ channel use: the first user in the ranking only signals if its channel can support the rate, if yes the user is scheduled, if not the other user is scheduled (though this would consume extra power from the BS if the channel of other user is also bad).

C. Mixed Policy

The mixed policy is a combination of both the ideas behind the centralized and decentralized policies. As in the decentralized policy, slot mT is used to broadcast signalling regarding the quantized queue lengths and the action that specifies how scheduling will be done in the next $T-1$ slots.

In the signalling slot, the BS there can choose one of the following actions: $\mathcal{F}=\{1\}$, $\mathcal{F}=\{2\}$, $\mathcal{F}=\{1,2\}$ and $F=1$. In the first three actions the user(s) specified train directly in the uplink for the $T-1$ slots after the signalling slot, without any control or contention/uplink of the IDs phase. In the case of $F=1$ one user is scheduled according to the contention procedure explained in Section V-B. In detail, for the rates at a slot t corresponding to each of the BS actions and assuming $\hat{q}_{k(1)}(m)\geq\hat{q}_{k(2)}(m)$

we have for $t \in \{mT+1, ..., mT+T-1\}$:

$$\begin{aligned}
&\mathbb{E}\left\{\mu_1(t)\right\} = (T_s - (\beta_p + \beta))\bar{p}(1)R, \mu_2(t) = 0, \mathcal{F} = 1 \\
&\mathbb{E}\left\{\mu_1(t)\right\} = 0, \mu_2(t) = (T_s - (\beta_p + \beta))\bar{p}(1)R, \mathcal{F} = 2 \\
&\mathbb{E}\left\{\mu_1(t)\right\} = \mathbb{E}\left\{\mu_2(t)\right\} = (T_s - (\beta_p + 2\beta))\bar{p}(1)R, \mathcal{F} = \{1,2\} \\
&\mathbb{E}\left\{\mu_{k(1)}(t)\right\} = \bar{\mu}_d(1) \\
&\mathbb{E}\left\{\mu_{k(2)}(t)\right\} = (1 - \bar{p}(1))\bar{\mu}_d(1), F = 1.
\end{aligned} \tag{32}$$

We define further $\bar{\mu}_m(\{k\}) = (T_s - (\beta_p + \beta))\bar{p}(1)R$ and $\bar{\mu}_m(\{1,2\}) = (T_s - (\beta_p + 2\beta))\bar{p}(2)R$. The mixed policy selects, at every slot mT, the following action:

- $\mathcal{F} = \{k(1)\}$, if
  $\hat{q}_k(1)(mT)\bar{\mu}_m(\{k\}) >$
  $\max\left\{(\hat{q}_1(mT) + \hat{q}_2(mT))\bar{\mu}_m(\{1,2\}), (\hat{q}_1(mT) + (1-\bar{p}(1))\hat{q}_2(mT))\bar{\mu}_d(1)\right\}$
- $\mathcal{F} = \{1,2\}$, if
  $(\hat{q}_1(mT) + \hat{q}_2(mT))\bar{\mu}_m(\{1,2\}) \leq$
  $\max\left\{\hat{q}_k(1)(mT)\bar{\mu}_m(\{k\}), (\hat{q}_1(mT) + (1-\bar{p}(1))\hat{q}_2(mT))\bar{\mu}_d(1)\right\}$
- $F = 1$ if
  $(\hat{q}_1(mT) + (1-\bar{p}(1))\hat{q}_2(mT))\bar{\mu}_d(1) >$
  $\max\left\{\hat{q}_k(1)(mT)\bar{\mu}_m(\{k\}), (\hat{q}_1(mT) + \hat{q}_2(mT))\bar{\mu}_m(\{1,2\})\right\}$

The main result here is summarized in the following

Theorem 8. *The stability region of the mixed scheme in the* 2 *user case with i.i.d. channels and one rate level is*

$$\begin{aligned}
\Lambda_m^{(2)} = \left(1 - \frac{1}{T}\right)\mathcal{CH}\Big\{&(0, \bar{\mu}_m(\{k\})), (\bar{\mu}_d(1), (1-\bar{p}(1))\bar{\mu}_d(1)), (\bar{\mu}_m(\{1,2\}), \bar{\mu}_m(\{1,2\})), \\
&((1-\bar{p}(1))\bar{\mu}_d(1), \bar{\mu}_d(1)), (\bar{\mu}_m(\{k\}), 0)\Big\}.
\end{aligned} \tag{33}$$

Proof. The proof is in the same spirit as the proof of Theorem 7, that is deriving the stability region of every action first. Due to the high similarity for the proofs of Theorems 6 and 7, only the outline is given to avoid repetition.

The stability region for the action $F = 1$ for every signalling slot has already been derived in the proof of Theorem 7. In addition, if the action $\mathcal{F} = \{1,2\}$ is chosen all the time, the mean arrival rates that can be supported must satisfy

$$T\lambda_k < (T-1)\bar{\mu}_m(\{1,2\}), \forall k \in \{1,2\}.$$

Finally, if only the user with the biggest (quantized) queue length at the beginning of slot mT is scheduled in the slots $t \in \{mT+1,..,mT+T-1\}$, the region with mean arrival rates such that there exist probabilities $\pi_{\{1\}}, \pi_{\{2\}}$ so that

$$\begin{aligned} T\lambda_1 &< (T-1)\pi_{\{1\}}\bar{\mu}_m(\{k\}) \\ T\lambda_2 &< (T-1)\pi_{\{2\}}\bar{\mu}_m(\{k\}) \\ 0 \leq \pi_{\{1\}} &+ \pi_{\{2\}} \leq 1 \end{aligned} \tag{34}$$

is satisfied. The proof uses the same ideas with Theorem 6 (i.e. the randomized policy) and the bound of the Lyapunov drift including the quantized queue lengths seen in theorem Theorem 7. Same arguments as the ones of Theorem 6 give that the mixed policy achieves the stability region of this Theorem and the converse, i.e. that every mean arrival rate vector for which the mixed policy stabilizes the system is in the stability region given in the Theorem. □

In short, as with the centralized and decentralized policies, the mixed policy aims to maximize the quantity $\mathbb{E}\left\{\sum_{k=1]}^{2} \hat{q}_k(mT)\mu_k(t)|\hat{\mathbf{q}}(mT)\right\}$ over all allowed actions.

VI. General case

In this Section we consider the general case with K users and L possible transmission rates, as described in Section II.

A. Centralized Policy

We begin by considering the centralized policy and characterizing its stability region. We denote by $\mu_k^{(c)}(t)$ the service in bits given to user k at slot t under the centralized policy.

Theorem 9. *The stability region Λ_c of the centralized policy consists in all rate vectors $\lambda = [\lambda_1,..,\lambda_K]$ for which there exist $0 \leq p(\mathcal{F}) \leq 1, \mathcal{F} \in 2^{\mathcal{K}}$ with $\sum_{\mathcal{F}\in 2^{\mathcal{K}}} p(\mathcal{F}) = 1$ such that*

$$\lambda_k < \sum_{\mathcal{F}\in 2^{\mathcal{K}}} p(\mathcal{F}) I_{\{k\in\mathcal{F}\}}(T_s - (1+\beta_p+(\beta+\beta_c)|\mathcal{F}|))\mathbb{E}\left\{r_k||\mathcal{F}|\right\}, \forall k \in \mathcal{K}. \tag{35}$$

Proof. First we prove that the centralized policy achieves the region characterized above. Note that this stability region is achieved by a randomized policy that every time slot schedules users in the set $\mathcal{F}$ randomly with probability $p(\mathcal{F})$ such that for every user (35) is satisfied. Denoting μ_k^* the mean rate user k gets under this policy, satisfaction of the aforementioned condition implies that for some $\epsilon > 0$

there is $\mu_k^* - \lambda_k \geq \epsilon$. Defining the quadratic Lyapunov function $V(\mathbf{x}) = \frac{1}{2}\sum_{k=1}^{K} x_k^2$, we have under the centralized policy (the expectation is over the arrival, channel and scheduling policies):

$$\begin{aligned}\Delta V(\mathbf{q}) &= \mathbb{E}\left\{V(\mathbf{q}(t+1)) - V(\mathbf{q}(t))|\mathbf{q}(t) = \mathbf{q}\right\}\\ &\leq B + \sum_{k=1}^{K} q_k(t)\lambda_k - \sum_{k=1}^{K} q_k(t)\mathbb{E}\left\{\mu_k^{(c)}(t)|\mathbf{q}(t) = \mathbf{q}\right\}\\ &\leq B + \sum_{k=1}^{K} q_k(t)(\lambda_k - \mu_k^*) \leq B - \epsilon \sum_{k=1}^{K} q_k(t),\end{aligned}$$

which implies that the system under the centralized policy is stable. The second inequality holds because the centralized policy chooses $\mathcal{F}$ to maximize the second sum; under the randomized policy the expectation of the rate of user k in every slot is μ_k^*, since this schedule is chosen independent of everything.

Conversely, assume that the centralized policy renders the system stable for a mean arrival rate vector λ in the interior of Λ_c. Then, the system is a Markov chain and since it is strongly stable it has a unique invariant distribution $\pi(\mathbf{q})$. In addition, since it is stable, there must hold $\lambda_k < \lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mu_k^{(c)}(\tau)$. On the other hand, the scheduling decision depends on the queue length only, this dependency denoted below by $\mathcal{F}(\mathbf{q})$, therefore the mean service rate for user k under the centralized policy is given as

$$\begin{aligned}\lim_{t\to\infty} \frac{1}{t}\sum_{\tau=0}^{t-1} \mu_k^{(c)}(\tau) &= \sum_{\mathbf{q}\in\mathbb{Z}_+^K} \pi(\mathbf{q})\mathbb{E}\left\{\mu_k^{(c)}(t)|\mathcal{F}(\mathbf{q})\right\}\\ &= \sum_{\mathcal{F}} \mathbb{E}\left\{\mu_k^{(c)}(t)|\mathcal{F}\right\} \sum_{\mathbf{q}\in\mathbb{Z}_+^K:\mathcal{F}(\mathbf{q})=\mathcal{F}} \pi(\mathbf{q}) > \lambda_k.\end{aligned}$$

Replacing for the service, the above can be written indeed as (35) by setting $p(\mathcal{F}) = \sum_{\mathbf{q}\in\mathbb{Z}_+^K:\mathcal{F}(\mathbf{q})=\mathcal{F}} \pi(\mathbf{q})$.

□

Geometrically the stability region of the centralized policy is the convex hull generated by the points $\left(\mathbb{E}\left\{\mu_1^{(c)}(t)|\mathcal{F}\right\}, ..., \mathbb{E}\left\{\mu_K^{(c)}(t)|\mathcal{F}\right\}\right)$ for every $\mathcal{F}$ subset of $\{1, ..., K\}$. The expectation is over the channel state distributions.

B. Decentralized Policy

Denoting as $\Lambda_d(F)$ the stability region of the decentralized policy for a fixed number of users to feed back every timeslot, i.e. setting $F(m) = F, \forall m \geq 0$. In this case, from the description of the contention scheme, the users with the F maximum values of $\hat{q}_k(t)\mathbb{E}\left\{\mu_k(t)|g_k(t), F\right\}$ will eventually get scheduled in every data slot t, where the channel state realizations are such that the magnitudes are $g_k(t)$. This observation leads to the following:

Lemma 10. *The region* $\Lambda_d(F)$ *consists in all mean arrival rate vectors* $\lambda \in \mathbb{R}_+^K$ *for which there exist* $\phi_{\mathcal{F}}(\mathbf{g}) \geq 0$ *such that*

$$\begin{aligned} T\lambda_k &< (T-1)\int_0^\infty p_1(g_1)dg_1 ... \int_0^\infty p_K(g_K)dg_K \sum_{\mathcal{F}:|\mathcal{F}|=F} \phi_{\mathcal{F}}(\mathbf{g})\mathbb{E}\left\{\mu_k(t)|g_k,\mathcal{F}\right\} \\ &:= (T-1)\mu_k^*(F) \\ &\sum_{\mathcal{F}:|\mathcal{F}|=F} \phi_{\mathcal{F}}(\mathbf{g}) \leq 1, \forall \mathbf{g} \in \mathbb{R}_+^K \end{aligned} \quad . \tag{36}$$

The expectation is with respect to the directions of the channel vectors and $\mathbf{g}$ *is the vector containing a realization of the channel magnitudes.*

Proof. The region above is achieved by a hypothetical policy where the BS at each slot t knows the realizations of all the channel magnitudes and, based on this knowledge, selects each set $\mathcal{F}$ with a probability $\phi_{\mathcal{F}}(\mathbf{g})$, while not transmitting any data on slot $mT, m = 0, 1, 2, ...$ We will show that the decentralized policy stabilized the system when it is can be stabilized by the aforementioned hypothetical policy. By Lemma 4, it suffices to prove that the decentralized policy stabilizes the system defined by $\tilde{\mathbf{q}}(m) = \mathbf{q}(mT)$ if the hypothetical policy renders it stable. Assume a mean arrival rate vector λ such that the system is stable under the hypothetical policy. Every timeslot $t \in \{mT+1, ..., mT+T-1\}$, the outcome of the decentralized policy is the F users with the greatest values of $\hat{q}_k(m)\mathbb{E}\left\{r_k(t)|g_k(t), F\right\}$. The drift of the quadratic Lyapunov function for the decentralized policy for the system $\tilde{\mathbf{q}}(m)$ is then (the inner expectations are with respect to the channel directions and the outer with respect to the channel magnitudes):

$$\begin{aligned} \Delta V_{(d)}(\mathbf{q}) &\leq \tilde{B} + T\sum_{k=1}^K \tilde{q}_k(m)\lambda_k - (T-1)\mathbb{E}\left\{\sum_{k=1}^K \tilde{q}_k(m)\mathbb{E}\left\{\tilde{\mu}_k^{(d)}(m)|g_k, F\right\}|\hat{\mathbf{q}}(m)\right\} \\ &\leq (\tilde{B} + TKQ\lambda_k + (T-1)KR_LQ) + T\sum_{k=1}^K \hat{q}_k(m)\lambda_k \\ &- (T-1)\mathbb{E}\left\{\sum_{k=1}^K \hat{q}_k(m)\mathbb{E}\left\{\tilde{\mu}_k^{(d)}(m)|g_k, F\right\}|\hat{\mathbf{q}}(m)\right\} \\ &\leq \tilde{C} + T\sum_{k=1}^K \hat{q}_k(m)\lambda_k \\ &- (T-1)\mathbb{E}\left\{\sum_{k=1}^K \hat{q}_k(m)\sum_{\mathcal{F}:|\mathcal{F}=F|} \phi_{\mathcal{F}}(g_1, ..g_K)I_{\{k\in\mathcal{F}\}}\mathbb{E}\left\{\tilde{\mu}_k^{(d)}(m)|g_k, F\right\}\hat{\mathbf{q}}(m)\right\} \\ &\leq \tilde{C} + \sum_{k=1}^K (\hat{q}_k(m)(T\lambda_k - (T-1)\mu_k^*(F)) \leq \tilde{C} - \epsilon\sum_{k=1}^K \hat{q}_k(m), \end{aligned}$$

for some $\epsilon > 0$, with $\tilde{C} = \tilde{B} + TKQ\lambda_k + (T-1)KR_LQ$ and the last inequality stems from the fact that λ is inside the stability region of the hypothetical policy. The drift gets negative for $||\hat{\mathbf{q}}||_1 > \tilde{C}/\epsilon$. Since $||\tilde{\mathbf{q}}||_1 > KQ + ||\hat{\mathbf{q}}||_1$, the drift is negative for $||\tilde{\mathbf{q}}||_1 \geq KQ + \frac{\tilde{C}}{\epsilon}$. From the Lyapunov-Foster criterion this implies that the system $\hat{\mathbf{q}}(m)$, therefore $\mathbf{q}(t)$, is stable under the decentralized policy if it is stable under a hypothetical policy that can achieve the stability region in the statement of this Lemma; therefore, this stability region is achievable by the decentralized policy.

We now proceed to show the converse, that is, if the system is stable under the decentralized policy for a mean arrival rate vector λ then this vector belongs in the set $\Lambda_d(F)$. Indeed, the set of users that are served at each timeslot $t = mT+1, ..., mT+T-1$ is a function on the realizations of the channel magnitude at slot t and the queue lengths $\mathbf{q}(mT)$. Denote this as $\mathcal{F}(\mathbf{q}, \mathbf{g})$. Since the channels are i.i.d. in time, the system $\tilde{\mathbf{q}}(m)$ is Markovian on a countable state space with a single communicating class. Stability then implies positive recurrence of the chain $\tilde{\mathbf{q}}(m)$, which further implies that there exist a (unique) stationary distribution, $\pi(\mathbf{q})$. The mean service rate (in bits per slot) user k gets is:

$$\lim_{M\to\infty} \frac{1}{M} \sum_{m=0}^{M-1} \sum_{t=mT+1}^{(m+1)T-1} \mu_k(t) = (T-1) \sum_{\mathbf{q}\in\mathbb{Z}_+^K} \pi(\mathbf{q}) \mathbb{E}\left\{\mu_k|\mathbf{q}\right\}$$
$$= (T-1) \sum_{\mathbf{q}\in\mathbb{Z}_+^K} \pi(\mathbf{q}) \int_0^\infty p_1(g_1)dg_1 ... \int_0^\infty p_K(g_K)dg_K \mathbb{E}\left\{\mu_k|\mathcal{F}(\mathbf{q},\mathbf{g}),\mathbf{g}\right\}$$
$$= (T-1) \int_0^\infty p_1(g_1)dg_1 ... \int_0^\infty p_K(g_K)dg_K \sum_{\mathbf{q}\in\mathbb{Z}_+^K} \pi(\mathbf{q}) \mathbb{E}\left\{\mu_k|\mathcal{F}(\mathbf{q},\mathbf{g}),\mathbf{g}\right\}$$
$$= (T-1) \int_0^\infty p_1(g_1)dg_1 ... \int_0^\infty p_K(g_K)dg_K \sum_{\mathcal{F}:|\mathcal{F}|=F} \sum_{\mathbf{q}\in\mathbb{Z}_+^K} \pi(\mathbf{q}) I_{\{\mathcal{F}=\mathcal{F}(\mathbf{q},\mathbf{g})\}} \mathbb{E}\left\{\mu_k|\mathcal{F},\mathbf{g}\right\}.$$

Denote $\phi_{\mathcal{F}}(\mathbf{g})' = \sum_{\mathbf{q}\in\mathbb{Z}_+^K} \pi(\mathbf{q}) I_{\{\mathcal{F}=\mathcal{F}(\mathbf{q},\mathbf{g})\}}$. For every $\mathbf{g}$ this is a probability distribution over the sets $\mathcal{F}$. In addition, since the system is stable, the mean arrival rate should be less that the mean service rate for each user [25], therefore

$$T\lambda_k < (T-1) \int_0^\infty p_1(g_1)dg_1 ... \int_0^\infty p_K(g_K)dg_K \phi_{\mathcal{F}}(\mathbf{g})' \mathbb{E}\left\{\mu_k|\mathbf{g}\right\},$$

which implies that the vector of mean arrival rates λ indeed belongs to the set $\Lambda_d(F)$. □

Based on the above Lemma and the fact that according to the decentralized policy the BS selects every signalling slot $mT$the number of users to be scheduled in the next $T-1$ slots based on the state of the queue lengths at the beginning of slot mT and the channel statistics, the main result of this subsection follows:

Theorem 11. *The stability region of the decentralized policy is the convex combination of the regions given by Lemma 10 for every F, i.e.*

$$\Lambda_d = \mathcal{CH}\left\{\Lambda_d(1), ..., \Lambda_d(F), ..., \Lambda_d(F_{max})\right\}.$$

Proof. To begin with, note that this region is achievable by a randomized policy that selects a number of users F to get scheduled according to a properly defined probability distribution $\pi(F(m) = F)$ [6] and the exact set of users to be scheduled is determined by the contention procedure of the decentralized scheme. Assume any vector $\lambda \in \Lambda_d$. In this case, there exist a probability distribution $\pi(F(m) = F)$ and $\epsilon > 0$ such that, for the mean service rate each user gets in each slot it holds

$$(T-1)\bar{\mu}_k - T\lambda_k \geq \epsilon. \tag{37}$$

We will show that the system under the decentralized policy is stable for this vector of mean arrival rates, meaning that Λ_d is achievable under the decentralized policy. Noticing that according to the decentralized policy the number $F(m)$ is selected in order to maximize the sum $\sum_{k=1}^{K} \hat{q}(m)\mathbb{E}\left\{\mu_k|F(m)\right\}$, we get for the Lyapunov drift of the system $\tilde{\mathbf{q}}(m)$:

$$\begin{aligned}\Delta V_{(d)}(\mathbf{q}) &\leq \tilde{B} - T\sum_{k=1}^{K}\tilde{q}_k(m)\lambda_k + (T-1)\sum_{k=1}^{K}\tilde{q}(m)\mathbb{E}\left\{\mu_k^{(d)}|F(m)\right\} \\ &\leq \tilde{C} - \sum_{k=1}^{K}\hat{q}_k(m)(T\lambda_k - (T-1)\mathbb{E}\left\{\mu_k^{(d)}|F(m)\right\}) \leq \tilde{C} - \epsilon\left|\left|\hat{\mathbf{q}}(m)\right|\right|_1.\end{aligned}$$

Following the same reasoning as in the proof of Lemma 10, this implies that the system is indeed stable under the decentralized policy for $\lambda \in \Lambda_d$.

To prove the converse, i.e. that every mean arrival rate vector λ for which the decentralized policy renders the system stable belongs to the set λ_d, we proceed in the same way as in the proof of Lemma 10. Since the system is stable for this vector, there exists a unique stationary distribution $\pi(\mathbf{q})$ of the markov chain that describes $\tilde{\mathbf{q}}(m)$. We have that $F(m)$ is in fact function of the queue lengths $\tilde{\mathbf{q}}(m)$ only, and we denote it by $F(\mathbf{q})$. Also, the mean service every user gets in a timeslot should be greater than the mean arrival rate. Based on that, and denoting $\bar{\mu}_k^{(d)}(F)$ the mean service (in bits per slot) user k gets under the decentralized policy when $F(m) = F$ we have

$$T\lambda_k < (T-1)\sum_{\mathbf{q}\in\mathbb{Z}_+^K}\pi(\mathbf{q})\bar{\mu}_k^{(d)}(F(\mathbf{q})),$$

[6] Knowledge of the statistics of the arrival processes is needed along with the statistics of the channels to select this probability distribution

which is indeed the convex hull of the sets $\Lambda_d(F)$. $\square$

The stability region of the system under the decentralized policy is thus the same as the biggest one achieved in the hypothetical case where all channel magnitude realizations were available to the scheduler, keeping the same timing overheads as in the decentralized policy. This shows why in the decentralized policy the users scheduled get high rates in *bits per channel use*. This is an advantage compared to the centralized policy, since in the latter a user with a bad channel realization may be scheduled. On the other hand, the decentralized policy comes with the disadvantage of spending one every T slots for signalling rather than data transmission and needing more time for exchange of control information in the data slot in order to implement the contention phase (though if the contention period can be made small and the decentralized policy tends to schedule fewer users than the centralized the additional overhead for the contention period may not pose a problem - however this is not guaranteed to happen in practice).

C. Mixed Policy

The mixed policy uses the same signalling structure as the centralized policy, however essentially switches between using the centralized and decentralized schemes every slot mT for the $T-1$ slots that follow. Note that if the mixed scheme operates in centralized mode selecting a set $\mathcal{F}(m)$, $T-1$ slots are used for data transmission, however only $(\beta_p + \beta F(m))$ channel uses are devoted to overhead for pilots and control signalling (since the IDs of the users to participate are fixed from the information at slot mT). Denote

$$\Lambda_c' = \left(1 - \frac{1}{T}\right) \mathcal{CH} \left\{ \frac{T_s - (\beta_p + \beta F)}{T_s - (1 + \beta_p + \beta F + \beta_c F)} \Lambda_c(\mathcal{F}) \right\}, \tag{38}$$

where

$$\Lambda_c(\mathcal{F}) = \left\{ \lambda \in \mathbb{R}_K^+ : \lambda_k < \mathbb{E}\left\{(T_s - (\beta_p + \beta F + \beta_c F)) r_k(t) | F\right\} I_{\{k \in \mathcal{F}\}} \right\}, \tag{39}$$

i.e. the stability region if the set of users $\mathcal{F}$ was scheduled all the time according to the mixed policy. Then we have the following:

Theorem 12. *The stability region of the mixed policy is*

$$\Lambda_m = \mathcal{CH}\left\{\Lambda_c', \Lambda_d\right\}.$$

The proof can be done in the same way as the ones in the previous subsections and is thus omitted to avoid repetition.

D. Comparison and Discussion

By selecting a high enough value of T, we can guarantee that the mixed scheme increases the stability region of the system. Denote

$$\hat{m} = \min_{1 \le F \le F_{max}} \left\{ \frac{T_s - (\beta_p + \beta F)}{T_s - (\beta_p + \beta F + \beta_c F + 1)} \right\}. \tag{40}$$

Then the following result holds:

Proposition 13. *A sufficient condition for the mixed scheme to have greater stability region than the centralized scheme is*

$$T > \frac{1}{1 - \hat{m}^{-1}}. \tag{41}$$

In this case, $\Lambda_m \supseteq \rho(T)\Lambda_c$ *with* $\rho(T) = \left(1 - \frac{1}{T}\right)\hat{m}$*.*

Proof. An immediate corollary of Theorem VI-C is that $\Lambda_m \supseteq \Lambda'_c$. Note that, by the definition of this region, the signalling overhead required in a data slot is smaller than the one of the centralized policy. A sufficient condition, thus for $\Lambda'_c \supset \Lambda_c$ is

$$\left(1 - \frac{1}{T}\right) \min_{1 \le F \le F_{max}} \left\{ \frac{T_s - (\beta_p + \beta F)}{T_s - (\beta_p + \beta F + \beta_c F + 1)} \right\} > 1. \tag{42}$$

Replacing from (40), we get the stated result. □

For proving the expansion of the stability region by using the mixed policy, only the centralized mode of the mixed policy was used: The proof uses the fact that if the set of users to get scheduled is broadcasted periodically then, in each of the data slots, the signalling overhead is reduced (no control section is needed). Having the mixed policy select the same users as a centralized policy for most of the time can happen, for example, when the traffic patterns are such that a few users request very high rate or (in the extreme case) only one user requests nonzero rate. In these cases, it may be better to directly serve the users with the heavy traffic demands most of the time, in order to avoid the extra overhead for implementing the contention of the decentralized policy. In fact, in these cases the centralized policy might even perform better than the decentralized due to reduced overhead. On the other hand, the decentralized policy can get used in cases with more uniformly distributed traffic load. As seen in the system with the two users, using the decentralized policy along with the centralized one for the mixed policy can satisfy rate demands that the centralized policy alone could not. The result of Proposition 13 is thus a sufficient condition and expansion of the region in some directions of mean arrival rate vectors can be much higher than $\rho(T)$.

As far as the exact difference between the stability regions of the centralized policy and the mixed or decentralized policies is concerned, this is a challenging task as it depends on the system parameters (as does the exact shape of the regions) and it is not the same for different directions ofthe vectors of mean arrival rates, λ. For some insights, Fig. 5 depicts the stability region for the two-user system under specific values of the parameters. We can observe that, for most of the traffic demands, the decentralized mode is used in the mixed policy for most of the traffic demand vectors. In fact, the centralized policy achieves higher stability region than the decentralized one only near the points where the demands of one user is much smaller than the other: in these cases indeed the user with the high demand should be scheduled for much greater fraction of time. Also note that the rightmost points of the regions do not appear: this means that the system parameters (power, number of antennas, SNR threshold for correct decoding) are such that scheduling both users is actually not beneficial. Intuitively, for the two user system we can expect that the decentralized policy is much better than the centralized when channel conditions are on average bad, and on regimes with relatively low transmit power: there the knowledge of the channel realization is crucial. As the nulber of users grows, we can also expect that the decentralized policy would be getting betterthan the centralized in most cases (again, except the cases that few users have much largertraffic demands than the rest): As small subsets compared to the total number of users have to be scheduled (due to limited nulber of BS antennas and overhead issues), selecting the ones with good instantaneous channels becomes more and more important.

A last point to note is that increasing the time T between two signalling slots enlarges the stability region, however this expansion is eventually bounded by system parameters. In addition, selecting a very big value of T can have a negative impact in terms of delays experienced by the users and can lead to slow convergence of the queueing system to its stationary distribution.

VII. The case of OFDMA systems

In this Section we extend the analysis to systems using multiple orthogonal channels in frequency, i.e. OFDMA , on top of multiple antennas. We consider a system with N_c channels in frequency. Transmit power P is used for each channel, that is the total transmit power of the BS here is N_cP. The other parameters are the same as in the general description of Section II (i.e. N antennas, L rate levels). Channels are assumed to be independent in frequency and time, the channel state of user k on channel ν given as $\mathbf{h}_{k\nu}(t) = \sqrt{g_k}\hat{\mathbf{h}}_{k\nu}(t)$, where $\hat{\mathbf{h}}_{k\nu}(t) \sim \mathcal{CN}(0, \mathbf{I}_N)$. On each channel, Zero-Forcing precoding is employed to serve (potenitially) multiple users per time-frequency slot. Let $\mathcal{F}_\nu(t)$ the set of users that are scheduled at channel ν at timeslot t.

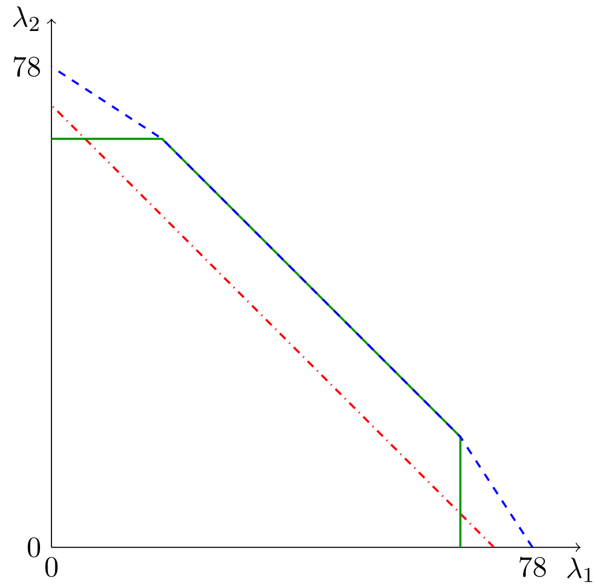

Fig. 5. Stability regions (in bits per slot) of the centralized (red dashed-dotted line), decentralized (green solid line) and mixed (blue dashed line) policies in the 2 user system with $N = 2$ antennas and a single rate level $R = 1$. The parameters of the system are here $P = \hat{S} = 1, \bar{g}_1 = \bar{g}_2 = 1, T = 100, \sigma^2 = 1, \tau_c = 7, \beta_c = 7, \beta = \beta_p = 4$ and $T_s = 100$ channel uses.

In this case, the policies of Section III can be carried over, running in parallel for each channel ν. Since the channels are assumed i.i.d. among frequencies for the same user, the parameters set by each policy (e.g. the sets $\mathcal{F}$ in the centralized policy and centralized mode of the mixed policy and the number of users F to get scheduled) are the same for every carrier. For the stability region of any policy π in the multiple channel case given its stability region in each carrier ν (i.e. if only this carrier was used to serve the users in the system) we have the following general result:

Theorem 14. *Let Λ^π_ν the stability region of the system using only carrier ν and Λ^π the region of the system using all carriers. Then the following holds:*

$$\Lambda^\pi = \oplus_{\nu=1}^{N_c} \Lambda^\pi_\nu. \tag{43}$$

Moreover, if channels in different frequencies are i.i.d. for the same user, the above reduces to

$$\Lambda^\pi = N_c \Lambda^\pi_\nu, \tag{44}$$

where Λ_ν is the stability region of one carrier.

Proof. The second part from the theorem follows directly from the first, so we will prove the first part only. We begin by showing that Λ^π is indeed achievable by policy π modified for the multicarrier case.

Assume $\lambda \in \Lambda^{\pi}$. Then, we can write

$$\lambda = \sum_{\nu=1}^{N_c} \lambda_{\nu}, \text{with } \lambda_{\nu} \in \Lambda_{\nu}^{\pi}, \forall \nu = 1, ..., N_c, \tag{45}$$

for some proper λ_{ν}. In other words, there exist $0 \leq \phi_{\nu k} \leq 1$ such that $\sum_{\nu=1}^{N_c} \phi_{\nu k} = 1, \forall k \in \mathcal{K}$ such that $\forall \nu = 1, .., N_c, \;\; [\phi_{\nu 1}\lambda_1, ..., \phi_{\nu K}\lambda_K]^T \in \Lambda_{\nu}^{\pi}$. This implies that a policy achieving this capacity region is a policy that creates a queue for every user at each channel ν, $q_{\nu k}(t)$ and routes the traffic arriving at time t for user k at queue ν with probability $\phi_{\nu k}$. These queues otherwise operate independently under the policy π applied to each channel ν. In this case, the system $q_{\nu k}(t)$ is stable and therefore every queue $q_k'(t) = \sum_{\nu=1}^{N_c} q_{\nu k}(t)$ is stable. Letting the average rate of user k at channel ν in this case be $\bar{\mu}_{\nu k}$, we have then

$$\bar{\mu}_{\nu k} > \lambda_{\nu k} = \phi_{\nu k}\lambda_k, \forall \nu, k. \tag{46}$$

On the other hand, every policy π from the ones presented in Section III tries to minimize some kind of Lyapunov drift based on the available information. For the mixed and decentralized policies, for the system $\tilde{\mathbf{q}}(m) = \mathbf{q}(mT)$, the above relationship becomes

$$(T-1)\bar{\mu}_{\nu k} > T\lambda_{\nu k} = T\phi_{\nu k}\lambda_k, \forall \nu, k, \tag{47}$$

therefore

$$\begin{aligned}
\Delta V_{\pi}(\mathbf{q}(m)) &\leq \tilde{C} + T\sum_{k=1}^{K} \lambda_k \hat{q}_k(m) - (T-1)\sum_{k=1}^{K} \hat{q}_k(m) \sum_{\nu=1}^{N_c} \mathbb{E}\left\{\tilde{\mu}_{\nu k}^{\pi}(m) | \hat{\mathbf{q}}(m)\right\} \\
&\leq \tilde{C} + T\sum_{k=1}^{K} \lambda_k \hat{q}_k(m) - (T-1)\sum_{k=1}^{K} \hat{q}_k(m) \sum_{\nu=1}^{N_c} \bar{\mu}_{\nu k} \\
&\leq \tilde{C} - \sum_{k=1}^{K} \hat{q}_k(m) \left(\epsilon + T\sum_{\nu=1}^{N_c} \phi_{\nu k}\lambda_k - T\lambda_k\right) = \tilde{C} - \epsilon \sum_{k=1}^{K} \hat{q}_k(m),
\end{aligned}$$

for some $\epsilon > 0$, due to (47). This implies that the system is stable under the policy π, thus π can achieve Λ^{π}.

For the centralized policy the achievability proof is essentially the same, just taking the system every slot t instead every slot mT and taking the actual queue lengths in the drift expressions.

To finish we need to prove the converse, that is that if the system is stable under π for a mean arrival rate vector λ then $\lambda \in \Lambda^{\pi}$. Then the system is an ergodic Markov chain, and since the service rates actually allocated to each user in each timeslot at each subcarrier depend on the queue lengths and the channel states, the limit $\lim_{T\to\infty} \frac{1}{T} \sum_{t=0}^{T-1} \mu_{\nu k}^{\pi}(t)$ exists, for every user and carrier and it is equal to $\bar{\mu}_{\nu k}^{\pi}$

such that (since the system is assumed stable)

$$\lambda_k < \sum_{\nu=1}^{N_c} \bar{\mu}^{\pi}_{\nu k}, \forall k \in \mathcal{K}.$$

This means that, for some $\epsilon_k > 0$ we have $\forall k$: $\lambda_k = \sum_{\nu=1}^{N_c}(\bar{\mu}^{\pi}_{\nu k} - \frac{\epsilon_k}{N_c})$. However, at each carrier ν there should be $\bar{\mu}^{\pi}_{\nu} \in \Lambda^{\pi}_{\nu}$ (if not, these rates are not achievable in this carrier), therefore $\bar{\mu}^{\pi}_{\nu} - \frac{1}{N_c}\epsilon \in \Lambda^{\pi}_{\nu}$. Setting $\lambda_{\nu k} = \bar{\mu}^{\pi}_{\nu k} - \frac{\epsilon_k}{N_c}$, the mean arrival rate vector can be written as an element of Λ^{π}, therefore the converse is proved. □

An application of the theorem above gives that, in the case of N_c carriers used, the stability regions of the centralized, decentralized and mixed policies are $N_c\Lambda_c$, $N_c\Lambda_d$, $N_c\Lambda_m$, respectively.

VIII. Trading signalling time for signalling bandwidth

In the above analysis we focused on the case where the quantized queue lengths and the number of users to be scheduled once every T timeslots, occupying the whole slot with this information. An alternative approach is to use an additional (low rate) control channel on which to broadcast this information.

More in detail, we add a downlink control channel of rate, say T_sR_c bits per time slot. Denote b the quantization bits per user; then the total number of signalling bits is $Kb + \log_2(K)$, transmission of which can be done in a time of $T = \frac{Kb+\log_2(K)}{T_sR_c}$ timeslots. The scheme here is to start broadcasting the signalling information of time mT at this slot and the users should use the most recent broadcasted information, i.e. at time $mT \leq t \leq (m+1)T - 1$ the users will use the broadcasted queue lengths and F of time $(m-1)T$. The error in the queue length estimation remains bounded and in addition all slots are now used for data transmission. It follows then that the stability region of this schemes will be the same as the corresponding schemes that take one slot for broadcast without the factor of $\left(1 - \frac{1}{T}\right)$. This statement can be proven using the same steps as the preceding analysis. For stability purposes, the rate of the added control channel can be arbitrarily low, with lower rate however probably deteriorating the delay performance of the scheme.

IX. A dynamic threshold scheme for discrete time contention

The contention-based protocol analyzed in the above Sections performs better that a centralized policy, however the assumptions of continuous time for contention and very short signals, although used in the literature dealing with CSMA, can be challenging to realize in practice. In this Section we present a scheme to still leverage the fact that the receivers know their channel states (channel gain in the particular) for the more practical case where contention is done in discrete time. We will assume throughout the

Section that the channels of the users are i.i.d. with $\mathbf{h}_k(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N), \forall k \in \{1, .., K\}$. In practice (where the users are in different distances from the BS), this can be achieved by appropriate power control.

The main idea here is to have users with channel magnitude above a threshold train and get scheduled. This threshold is updated dynamically by the BS, based on the state of the queues. We would like to mention here that there is a rich literature concerning the idea of threshold-based feedback in single-input single-output (SISO) systems in order to reduce the feedback overhead while selecting a ”good” user to schedule. It was initiated in [34], [35] for sum-rate maximization without considering the queues of the users. The recent work [36] addressed the problem of stability region expansion. Therein, the channel state of the user with the largest queue is used as threshold. The above frameworks cannot be simply extended to MIMO systems. In fact, in MIMO systems using ZF precoding (or in more general beamforming), the bit rate of each user does not depend only on its channel conditions but also on the channel realization of the other users that get served simultaneously, making the problem more challenging. Finally, authors in [37] employ similar ideas for a multiuser MIMO system where the receivers employ ZF. Apart from focusing on the receivers and not the precoding at the transmitter, this work addresses sum rate maximization, so the traffic processes of the users are not taken into account.

A. Protocol Description

More in detail, the BS uses slots $0, T, 2T, .., mT, ...$ for signaling and setting the control parameters to be used for the next $T-1$ slots. T is a (positive and possibly high) constant whose role in the performance of the algorithm will be examined in the next subsection. There are therefore two distinct phases:

1) Broadcast phase (t=0,T,2T,...): At slot mT the BS broadcasts IDs of the $\bar{K}(m)$ users with the biggest queue lengths (the set $\bar{\mathcal{K}}(m)$) ordered by decreasing queue length, the preferred number of users to be served $F(m)$ and the threshold $G(m)$, given as the solution of (51) that follows.

2) Uplink Training phase in each time slot: In all $T-1$ slots (i.e. between slots mT and $(m+1)T$), the user selection for uplink training is done as follows. A portion of the time slot is reserved to receive information from the users in $\bar{\mathcal{K}}(m)$ about their channels. This portion of time is divided into minislots of total number equal to the number of users in $\bar{K}(m)$. In each minislot (which lasts for one channel use), a user sends a signal (e.g. 1 bit signal) if its channel gain is above $G(m)$ and does nothing otherwise. The minislots are shared among the users in $\bar{\mathcal{K}}(m)$ in a TDMA manner according to the ordering broadcasted by the BS at slot mT. The procedure stops in the following cases:

i. When $F(m)$ users with channel above the threshold are found. In this case these users get scheduled for uplink training and the user selection period lasts for $L(G(m), F(m))$ channel uses. These $F(m)$ will send then their uplink training sequences that will last $\beta F(m)$ channel uses. The remaining time in the current time slot is then used for data transmission (we will see that even with this additional signaling the performance of this algorithm is better than the baseline user selection algorithm).

ii. When all users in $\bar{\mathcal{K}}$ have sent the 1 bit signal (as above) and $U \leq F(m)$ users have signaled that their channel gain is above the threshold. These U users will then get scheduled for training and transmission and the user selection period lasts for $\bar{K}$ channel uses. These U users will then send their training sequences (i.e. βU channel uses are used for uplink training). The remaining of the time slot is then used for data transmission.

The parameters to be optimized here are (i) the ordering of the users, (ii) the candidate users in the set $\bar{\mathcal{K}}(m)$,(iii) the number of users, $F(m)$, to get scheduled and (iv) the threshold for the channel magnitudes, G, which should be in general a function of the queue lengths, $F(m)$ and $\bar{K}(m)$.

Under the setting of i.i.d. channels for the users, they will be ordered according to their queue lengths at (the beginning of) timeslot mT, with the user with the largest queue first. This implies that the set of candidate users consists in the $\bar{K}(m)$ users with the biggest queue lengths.

We now turn to the impact of the threshold G. The probability that the channel magnitude of a user is above this threshold is given as

$$p(G) = \mathbb{P}\{g_k(t) > G\} = 1 - \frac{\gamma(G;N)}{\Gamma(N)} = 1 - \int_G^{+\infty} \frac{x^{N-1}e^{-x}}{\Gamma(N)}dx. \tag{48}$$

The above follows from the fact that since $\mathbf{h}_k(t) \sim \mathcal{CN}(\mathbf{0}, \mathbf{I}_N)$, the quantity $2\,||\mathbf{h}_k(t)||^2$ follows a chi-squared distribution with $2N$ degrees of freedom. In addition, let

$$\bar{r}_d(F,G) = \mathbb{E}\left\{ r_k(t) \middle| F, g_k(t) > G \right\}$$

be the mean rate in bits per channel use that user k gets if his channel magnitude is above the threshold and F users are scheduled in total. These can be calculated using the results of Section IV-A.

Intuitively, smaller threshold will cause the signalling phase to end faster, however users with worse channel condition may be selected. Assume a user ordering $\{k(1), k(2)..., k(K)\}$ such that $q_{k(1)}(mT) \geq q_{k(2)}(mT) \geq ... \geq q_{k(K)}(mT)$. Then, defining

$$M(i,m,U) = \sum_{j=\max\{0,U-1-m+i\}}^{\min\{U-2,i-1\}} \binom{i-1}{j}\binom{m-i-1}{U-(j-2)},$$

we have the following

Lemma 15. *The mean rate, $\bar{\mu}_{k(i)}(F,\bar{K},G)$ the i-th user in the ordering gets if the threshold is G and the BS signalled for F users out of the $\bar{K}$ users with the bigguest queue lengths to be scheduled is, for $i \le \bar{K}$:*

$$\bar{\mu}_i(1,\bar{K},G) = [T_s - (\beta_p + \beta_c i + \beta)]^+ p(G)(1-p(g^*))^{i-1}\bar{r}_d(F,G) \tag{49}$$

for $F(m) = 1$ and

$$\begin{aligned}\bar{\mu}_i(F,\bar{K},G) &= \bar{r}_d(F,G)\,[T_s - \beta_p - \beta_c i - \beta F]^+\, p^F(G)(1-p(G))^{i-F}\mathbb{1}_{\{F\le i\}} \\ &+ \bar{r}_d(F,G)p^F(G)\sum_{m'=\max[F,i+1]}^{\bar{K}}\left[T_s - \beta_p - \beta_c m' - \beta F\right]^+ M(i,m',F)(1-p(G))^{m'-F} \\ &+ \sum_{F'=1}^{F-1}\binom{\bar{K}-1}{F'-1}[T_s - \beta_p - \bar{K} - \beta F']^+ p^{F'}(G)(1-p(G))^{\bar{K}-F'}\bar{r}_d(F',G),\end{aligned} \tag{50}$$

for $2 \le F(m) \le F_{max}$. In addition, $\bar{\mu}_{k(i)}(F,\bar{K},G) = 0$ for $i > \bar{K}$.

Proof. Please refer to Appendix C for the proof. □

The expected weight for $F(m) = F$ and threshold for the channel gain set to g^* then is given as

$$\mathbb{E}\left\{\sum_{k=1}^{K} q_k(mT)\mu_k(t)z_k(t)\middle| F,\bar{K},G\right\} = \sum_{i=1}^{K} q_{k(i)}(mT)\bar{\mu}_i(F,\bar{K},G).$$

From the above, the BS sets at slot mT the parameters $F(m), g^*(m)$ to be used in slots $mT+1,...,mT+T-1$ as a solution to the following optimization problem:

$$\{F(m),\bar{K}(m),G(m)\} = \arg\max_{1\le F\le F_{max},1\le\bar{K}\le K,G\in\mathcal{G}}\left\{\sum_{i=1}^{K} q_{k(i)}(mT)\bar{\mu}_i(F,\bar{K},G)\right\} \tag{51}$$

where $\mathcal{G}$ denotes the set of possible thresholds for the channel magnitude. The stability region achieved by this protocol is then the following:

Theorem 16. *The stability region of the threshold-based scheme when signaling is done every T slots is*

$$\Lambda_{thr}(T) = \left(1-\frac{1}{T}\right)\mathcal{CH}_{F\in\{1,..,F_{max}\},\bar{K}=\{1,..,K\},G\in\mathcal{G}}\left\{\Lambda'(F,\bar{K},G)\right\},$$

where $\Lambda'(F,\bar{K},G)$ is the convex hull of all points in $\mathbb{R}^K$ resulting from all possible permutations of the numbers $\bar{\mu}_i(F,\bar{K},G)$.

The proof of the theorem is along the lines of the other proofs in this paper and thus is omitted.

B. Simulation Results

In order to illustrate the impact of the threshold-based in the stability region of the system, we simulated a single cell system where the BS uses $N = 4$ antennas to serve $K = 10$ users. A single modulation and coding scheme with a rate of R bits per channel use if the SNR of a scheduled user exceeds S is assumed. Specifically, we set $R = 1$ bit per channel use. Each user has Poisson traffic with different average arrival rates as $\lambda_K := \lambda$ and $\lambda_{k-1} = 1.05\lambda_k$: we parametrize the mean arrival rates by the smallest one and see how the system behaves as this smallest rate varies. That means that the mean arrival rate vector is

$$\lambda_k = 1.05^{10-k}\lambda, \forall k \in \{1, .., 10\}. \tag{52}$$

For simplicity, a threshold rule, where the threshold for the channel gain depends only on the preferred number F of users to be served and threshold $G = \frac{FS}{P}$ is used. The reason for choosing this is that if the channel gain of a user is smaller and $F - 1$ other users are served then the rate of this user will be zero. We set the slot duration to be $T_s = 1000$ channel uses, and overheads for contention, downlink pilot and uplink pilot length per user $\beta_c = 1$, $\beta_p = 50$ and $\beta = 100$ channel uses, respectively. The rate for signalling is set to $R_0 = 0.2$ bits per channel use. We assume that $P = S = 1$, that is the total transmit power and the SNR threshold for correct decoding are at the same level of the noise floor.

We simulated the system under the centralized policy, as well as under the threshold-based policy for different values of signalling period T and λ, each simulation lasting 10^6 timeslots. Notice that the decentralized policies, described in the previous sections, achieve obviously better performance than the centralized and threshold based policies. In order to illustrate the stability performance, we present in Figure 6 the average total queue lengths, i.e. the quantity $\frac{1}{10^6}\sum_{t=0}^{10^6-1}\sum_{k=1}^{K} q_k(t)$. The point where the system becomes unstable is the point where the total average queue length begins to increase very steeply.

We can see in the figure that, indeed, by selecting the signalling period T high enough, the threshold-based algorithm outperforms the baseline centralized policy. Notice here that under the setting regarding the physical layer, the total average service rate at each slot given that users with $Pg_k(t)/F > S$ are scheduled can be at most $805, 1287, 1262, 954$ if $F = 1, 2, 3, 4$ users are scheduled, respectively. The average total number of incoming bits is 12.57λ, therefore the proposed schemes achieve a good fraction of the biggest traffic achievable under our signalling constraints. As compared to the centralized policy, the gain is around 20%. In addition, the bigger the signalling period T is, the bigger stability region is achievable. Notice, however, that when T takes very high values, the increase in the stability region from increasing T further becomes negligible. This can be interpreted theoretically, using the result in Theorem 16 since the convex hull is multiplied by the $\left(1 - \frac{1}{T}\right)$ term.

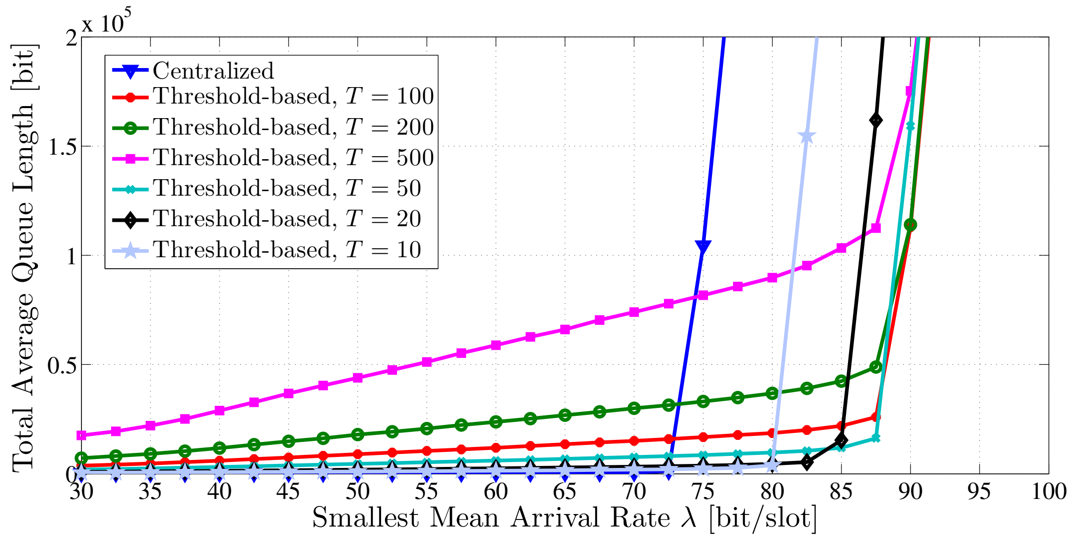

Fig. 6. Average Total Queue Length for Different Mean Arrival Rates with $S = 1$ and $P = 1$. The arrival rates are given by (52) for different values of the smallest rate, λ.

X. Conclusions

In this paper we addressed the problem of user selection in a system were CSI is acquired by the BS via uplink training from the intended receivers. We have demonstrated that a feedback/training policy that combines decentralized schemes for user selection along with a centralized one can achieve greater stability region than using the standard centralized policy alone, in the case of a MISO broadcast system. These results suggest that, in future systems, decentralized methods should be considered for feedback and user scheduling along with the traditional centralized ones, at least for data-oriented services. We have also proposed a scheme based on a threshold for the channel magnitude in the case where contention is implemented in a slotted manner. These results imply that the fact that the users know their channel states should be leveraged in future communication systems, especially when delivering data services. Important extensions can be addressing the problem when the channels of the users are correlated in time and/or taking into account errors in the channel estimation. For the setting where the channels are correlated in time, the decision of which users should feed back in each slot affects now the information that the base station will have in subsequent slots. This coupling between the control decision and information availabe makes the problm very challenging, but interesting both from a theoretical and practical perspective. In addition, another future direction can be characterizing the tradeoff between stability region and delay, which depends on the signalling period of the decentralized/mixed schemes.

ACKNOWLEDGEMENTS

The work of M. Debbah has been funded by the ERC Grant 305123 MORE (Advanced Mathematical Tools for Complex Network Engineering). The research of M. Assaad is funded by the european Celtic project SHARING.

APPENDIX A
PROOF OF PROPOSITION 2

To begin with, we denote $k = k(i)$ for some $1 \leq i \leq F$ and define $\{k(j)\}_{j=1,..,F}$ the set of users that reported their CSI . For notational convenience we will drop the notation for dependence on t and F. For all other users except i, define the matrix of the vector of the respective channel corresponding to fast fading as $\hat{\mathbf{H}}_{k(i)} = \left[\hat{\mathbf{h}}_{k(1)}, ..., \hat{\mathbf{h}}_{k(i-1)}, \hat{\mathbf{h}}_{k(i+1)}, .., \hat{\mathbf{h}}_{k(F)}\right]$, the matrix of the respective directions, $\mathbf{U}_{k(i)} = \left[\mathbf{u}_{k(1)}, ..., \mathbf{u}_{k(i-1)}, \mathbf{u}_{k(i+1)}, .., \mathbf{u}_{k(F)}\right]$, the matrix of large scale channel gains $\bar{\mathbf{G}}_{k(i)} = \text{diag}\left\{\bar{g}_{k(1)}, .., \bar{g}_{k(i-1)}, \bar{g}_{k(i+1)}, ..., \bar{g}_{k(F)}\right\}$ and finally the diagonal matrix containing the instantaneous channel gains, $\mathbf{G}_{k(i)} = \text{diag}\left\{g_{k(1)}, .., g_{k(i-1)}, g_{k(i+1)}, ..., g_{k(F)}\right\}$. Using these notations we can write:

$$\mathbf{H}_{k(i)} = \hat{\mathbf{H}}_{k(i)}\bar{\mathbf{G}}_{k(i)} = \mathbf{U}_{k(i)}\mathbf{G}_{k(i)}. \tag{53}$$

Replacing the above in (3), we get that the SNR at user $k(i)$ then can be written as

$$SNR_{k(i)} = \frac{g_{k(i)}P}{\sigma^2 F}\mathbf{u}_{k(i)}^H\left(\mathbf{I}_N - \mathbf{U}_{k(i)}\left(\mathbf{U}_{k(i)}^H\mathbf{U}_{k(i)}\right)^{-1}\mathbf{U}_{k(i)}^H\right)\mathbf{u}_{k(i)}. \tag{54}$$

However, we have from (53) that

$$\mathbf{U}_{k(i)} = \hat{\mathbf{H}}_{k(i)}\bar{\mathbf{G}}_{k(i)}\mathbf{G}_{k(i)}^{-1} = \hat{\mathbf{H}}_{k(i)}\mathbf{D}_{k(i)}, \tag{55}$$

where we have defined $\mathbf{D}_{k(i)} = \bar{\mathbf{G}}_{k(i)}\mathbf{G}_{k(i)}^{-1} = \text{diag}\left\{\frac{\bar{g}_{k(1)}}{g_{k(1)}}, .., \frac{\bar{g}_{k(i-1)}}{g_{k(i-1)}}, \frac{\bar{g}_{k(i+1)}}{g_{k(i+1)}}, ..., \frac{\bar{g}_{k(F)}}{g_{k(F)}}\right\}$. In addition, since $\hat{\mathbf{H}}_{k(i)}$ is a matrix with $F-1$ i.i.d. $\mathcal{CN}(0, \mathbf{I}_N)$ random vectors as columns, we can write its Singular Value Decomposition as [38]

$$\hat{\mathbf{H}}_{k(i)} = \mathbf{V}\mathbf{\Sigma}\mathbf{\Delta}, \tag{56}$$

where $\mathbf{V} \in \mathbb{C}^{N\times N}$ and $\mathbf{\Delta} \in \mathbb{C}^{(F-1)\times(F-1)}$ are isotropically distributed unitary matrices (i.e. Haar matrices) and $\mathbf{\Sigma} \in \mathbb{C}^{N\times(F-1)}$ is the matrix containing the singular values of $\hat{\mathbf{H}}_{k(i)}$. Replacing in the SNR expression (54) and noting that $\mathbf{V}\mathbf{V}^H = \mathbf{I}_N$ we eventually get

$$\begin{aligned} SNR_{k(i)} &= \frac{g_{k(i)}P}{\sigma^2 F}\mathbf{u}_{k(i)}^H\left(\mathbf{V}\mathbf{V}^H - \mathbf{V}\mathbf{\Sigma}\left(\mathbf{\Sigma}^H\mathbf{\Sigma}\right)^{-1}\mathbf{\Sigma}^H\mathbf{V}^H\right)\mathbf{u}_{k(i)} \\ &= \frac{g_{k(i)}P}{\sigma^2 F}\mathbf{u}_{k(i)}^H\mathbf{V}\left(\mathbf{I}_N - \mathbf{\Sigma}\left(\mathbf{\Sigma}^H\mathbf{\Sigma}\right)^{-1}\mathbf{\Sigma}^H\right)\mathbf{V}^H\mathbf{u}_{k(i)}. \end{aligned} \tag{57}$$

Defining now

$$\mathbf{v} = \mathbf{V}^H \mathbf{u}_{k(i)},$$

$$\mathbf{A} = \begin{bmatrix} \mathbf{0}_{F-1,F-1} & \mathbf{0}_{F-1,N-F+1} \\ \mathbf{0}_{N-F+1,F-1} & \mathbf{I}_{N-F+1} \end{bmatrix}$$

we have eventually

$$SNR_{k(i)} = \frac{g_{k(i)}P}{\sigma^2 F}\mathbf{v}^H\mathbf{A}\mathbf{v} = \frac{g_{k(i)}P}{\sigma^2 F}\sum_{j=F+1}^{N}|u_j|^2. \tag{58}$$

Note that the summation is over the squared magnitudes of the $N-F+1$ last components of the vector $\mathbf{v}$. Since $\mathbf{V}$ is a Haar matrix and $\mathbf{u}_{k(i)}$ is unitary, $\mathbf{v} \in \mathbb{C}^{N\times 1}$ is a random unitary isotropic vector with distribution [39]

$$p(\mathbf{v}) = \frac{\Gamma(N)}{\pi^N}\delta\left(||\mathbf{v}||^2 - 1\right), \tag{59}$$

where $\delta(.)$ denotes the Dirac function. We now define the vectors $\psi_1 \in \mathbb{C}^{(F-1)\times 1}$ and $\psi_2 \in \mathbb{C}^{(N-F+1)\times 1}$ such that $\psi_2 = [v_{F+1}, ..., v_N]^T$ and $\mathbf{v} = [\psi_1^T, \psi_2^T]^T$. In this case, we can write the SNR as

$$SNR_{k(i)} = \frac{g_{k(i)}P}{\sigma^2 F}\psi_2^H\psi_2. \tag{60}$$

In addition, note that $||\mathbf{v}||^2 = ||\psi_1||^2 + ||\psi_2||^2$ and the channel coefficient corresponding to each antenna is independent of the coefficients corresponding to the other antennas.

We will deal with the probability distribution function of the SNR (58) first. From the aforementioned expression, if follows that

$$p_{SNR_{k(i)}}(s|g_{k(i)}, F) = \frac{F\sigma^2}{g_{k(i)}P}p_{\psi_2^H\psi_2}\left(\frac{F\sigma^2}{g_{k(i)}P}s\right) \tag{61}$$

To proceed further, we note that the p.d.f. of $\psi_2^H\psi_2$ is in fact the probability that a vector is drawn from the distribution of unitary isotropic vectors with the sums of the squared magnitudes of its $N-F+1$ last elements equal to $\psi_2^H\psi_2$, thus:

$$\begin{aligned} p_{\psi_2^H\psi_2}(x) &= \int_{\mathbf{v}\in\mathbb{C}^N:||\mathbf{v}||=1,\sum_{k=F}^N|v_k|^2=x} p(\mathbf{v})d\mathbf{v} = \int_{\mathbf{v}\in\mathcal{C}^N:\sum_{k=F}^N|v_k|^2=x}\frac{\Gamma(N)}{\pi^N}\delta\left(\mathbf{v}^H\mathbf{v}-1\right)d\mathbf{v} \\ &= \frac{\Gamma(N)}{\pi^N}\int_{\psi_2\in\mathbb{C}^{N-F+1}} d\psi_2\delta\left(\psi_2^H\psi_2 - x\right)\int_{\psi_1\in\mathbb{C}^{F-1}} d\psi_1\delta\left(\psi_1^H\psi_1 + \psi_2^H\psi_2 - 1\right) \\ &= \frac{\Gamma(N)}{\pi^N}\left(\int_{\psi_2\in\mathbb{C}^{N-F+1}}\delta\left(\psi_2^H\psi_2 - x\right)d\psi_2\right)\left(\int_{\psi_1\in\mathbb{C}^{F-1}}\delta\left(\psi_1^H\psi_1 - (1-x)\right)d\psi_1\right) \end{aligned} \tag{62}$$

In order to calculate the integrals we make use the Fourier transform of the Dirac function, a method used also in [40]. We have the following general result (it is similar to a result in [40] but we present its proof here for completeness) :

Lemma 17. *For $M \in \mathbb{Z}_+$ and $r > 0$ it holds*

$$\int_{\mathbf{u}\in\mathbb{C}^M} \delta\left(\mathbf{u}^H\mathbf{u} - r\right) d\mathbf{u} = \frac{\pi^M}{\Gamma(M)} r^{M-1}.$$

Proof: For this proof, i will denote the imaginary unit (i.e. $i^2 = -1$). We begin by replacing the Dirac function with its Fourier representation, that is $\delta(x) = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega e^{i\omega x}$ and we get, for (any) $\alpha > 0$,

$$\begin{aligned}
\int_{\mathbf{u}\in\mathbb{C}^M} \delta\left(\mathbf{u}^H\mathbf{u} - r\right) d\mathbf{u} &= \int_{\mathbf{u}\in\mathbb{C}^M} d\mathbf{u}\frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega e^{i\omega(\mathbf{u}^H\mathbf{u}-r)} \\
&= \frac{1}{2\pi}\int_{\mathbf{u}\in\mathbb{C}^M} d\mathbf{u} e^{-\alpha\mathbf{u}^H\mathbf{u}} e^{+\alpha\mathbf{u}^H\mathbf{u}} \int_{-\infty}^{+\infty} d\omega e^{i\omega(\mathbf{u}^H\mathbf{u}-r)} \\
&= \frac{e^{\alpha r}}{2\pi}\int_{\mathbf{u}\in\mathbb{C}^M} d\mathbf{u} e^{-\alpha\mathbf{u}^H\mathbf{u}} \int_{-\infty}^{+\infty} d\omega e^{i\omega(\mathbf{u}^H\mathbf{u}-r)} \qquad (63) \\
&= \frac{e^{\alpha r}}{2\pi}\int_{-\infty}^{+\infty} d\omega e^{-i\omega r}\int_{\mathbf{u}\in\mathbb{C}^M} d\mathbf{u} e^{-\mathbf{u}^H\mathbf{u}(\alpha - i\omega)} \\
&= \frac{e^{\alpha r}}{2\pi}\int_{-\infty}^{+\infty} d\omega e^{-i\omega r}\frac{\pi^M}{(\alpha - i\omega)^M} \\
&= e^{\alpha r}\pi^M\left(\frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega e^{i\omega r}\frac{1}{(\alpha + i\omega)^M}\right). \qquad (64)
\end{aligned}$$

Line (63) is obtained by noting that the integrand with respect to $\mathbf{u}$ is nonzero only for $\mathbf{u}^H\mathbf{u} = r$. The term in the parenthesis in (64) is the inverse Fourier transform of the function $\frac{1}{(\alpha+i\omega)^M}$. Noting that the Fourier transform of $e^{-\alpha r}\mathbb{1}_{\{r\geq 0\}}$ is $\frac{1}{\alpha + i\omega}$ and using the properties of Fourier transforms (mainly the properties about their derivatives and linearity) we get, for $r > 0$

$$\int_{\mathbf{u}\in\mathbb{C}^M} \delta\left(\mathbf{u}^H\mathbf{u} - r\right) d\mathbf{u} = e^{\alpha r}\pi^M\frac{r^{M-1}e^{-\alpha r}}{(M-1)!} = \frac{\pi^{M-1}}{(M-1)!}r^{M-1},$$

which, since for any positive integer M there is $(M-1)! = \Gamma(M)$, completes the proof. ■

Applying Lemma 17 in equation (62) we get eventually

$$p_{\psi_2^H\psi_2}(x) = \frac{\Gamma(N)}{\Gamma(F-1)\Gamma(N-F+1)}(1-x)^{N-F}x^{F-2} = \frac{(1-x)^{N-F}x^{F-2}}{B(N-F+1, F-1)}. \qquad (65)$$

Using (61) and replacing with (65) we have

$$\begin{aligned}\mathbb{P}\left\{SNR_{k(i)} > S|g_{k(i)}, F\right\} &= \int_{S}^{\frac{Pg_{k(i)}}{\sigma^2 F}} p_{SNR_{k(i)}}(s|g_{k(i)}, F)ds \\ &= \int_{S}^{\frac{Pg_{k(i)}}{\sigma^2 F}} \frac{F\sigma^2}{g_{k(i)}P} p_{\psi_2^H\psi_2}\left(\frac{F\sigma^2}{g_{k(i)}P}s\right) ds \\ &= \int_{\frac{F\sigma^2}{Pg_{k(i)}}S}^{1} p_{\psi_2^H\psi_2}(x)dx \\ &= \frac{1}{B(N-F+1, F-1)} \int_{\frac{F\sigma^2}{Pg_{k(i)}}S}^{1} (1-x)^{N-F} x^{F-2} dx,\end{aligned}$$

which is the stated result.

Appendix B

Proof of Lemma 5

From the evolution equation for the sampled queue lengths we have

$$\begin{aligned}\tilde{q}_k^2(m+1) = &\Bigg(\tilde{q}_k(m) + \sum_{t=0}^{T-1} a_k(mT+t) - \sum_{t=1}^{T-1} z_k(mT+t)\mu_k(mT+t) + \\ &\sum_{t=1}^{T-1} y_k(mT+t)\Bigg)^2 \\ &\leq \left(\tilde{q}_k(m) + \sum_{t=0}^{T-1} a_k(mT+t) - \sum_{t=1}^{T-1} z_k(mT+t)\mu_k(mT+t)\right)^2 \\ &= \tilde{q}_k^2(m) + \left(\sum_{t=0}^{T-1} a_k(mT+t)\right)^2 + \left(\sum_{t=1}^{T-1} z_k(mT+t)\mu_k(mT+t)\right)^2 \\ &+ 2\tilde{q}_k(m)\sum_{t=0}^{T-1} a_k(mT+t) - 2\tilde{q}_k(m)\sum_{t=1}^{T-1} z_k(mT+t)\mu_k(mT+t) \\ &\leq \tilde{q}_k^2(m) + T^2 A_{max}^2 + (T-1)^2 R_L^2 \\ &- 2\tilde{q}_k(m)\sum_{t=0}^{T-1} a_k(mT+t) + 2\tilde{q}_k(m)\sum_{t=1}^{T-1} z_k(mT+t)\mu_k(mT+t).\end{aligned}$$

From the above, setting $\tilde{B} = K(T^2 A_{max}^2 + (T-1)^2 R_L^2)$ and taking expectations it follows that

$$\begin{aligned}
\Delta V(\tilde{\mathbf{q}}(m)) &\leq \tilde{B} + \mathbb{E}\left\{\sum_{k=1}^{K} 2\tilde{q}_k(m) \sum_{t=0}^{T-1} a_k(mT+t)|\tilde{q}_k(m)\right\} \\
&- \mathbb{E}\left\{\sum_{k=1}^{K} 2\tilde{q}_k(m) \sum_{t=1}^{T-1} z_k(mT+t)\mu_k(mT+t)|\tilde{q}_k(m)\right\} \\
&= \tilde{B} + 2\sum_{k=1}^{K} \tilde{q}_k(m) \sum_{t=0}^{T-1} \mathbb{E}\{a_k(mT+t)\} \\
&- 2\sum_{k=1}^{K} \tilde{q}_k(m) \sum_{t=1}^{T-1} \mathbb{E}\{z_k(mT+t)\mu_k(mT+t)|\tilde{\mathbf{q}}(m)\} \\
&= \tilde{B} + 2T\sum_{k=1}^{K} \tilde{q}_k(m)\lambda_k \\
&- 2\sum_{k=1}^{K} \tilde{q}_k(m) \sum_{t=1}^{T-1} \mathbb{E}\{z_k(mT+t)\mu_k(mT+t)|\tilde{\mathbf{q}}(m)\}
\end{aligned}$$

Where the last equality follows from the fact that the arrival processes are i.i.d. in time and independent of anything else. Moreover, we have that the schedule $z_k(mT+t)$ and the rate of the user in the slot $\mu_k(mT+t)$ depend only on the queue length vector $\tilde{\mathbf{q}}(m)$ and the realizations of the channels and not on the time index. This implies that the product $z_k(mT+t)\mu_k(mT+t)$ has the same distribution for every slot between $mT+1 \leq \tau \leq m(T+1)$. Defining $\tilde{\mu}_k(m)$ as this final rate (including the scheduling decision) user k gets, for a given scheme it is as if we have $(T-1)$ independent copies of this random variable (with probability distribution over the probability distribution of the channels), hence the statement follows.

APPENDIX C
PROOF OF LEMMA 15

Proof of equation (49): Notice that here only one user is to be scheduled (since $F(m) = 1$). This means that the i-th user in the ordering gets scheduled if (i) his channel magnitude is above the threshold *and* (ii) the channel magnitudes of the $i-1$ users before him are below the threshold. In addition, if this user is scheduled the contention period stops right after, i.e. lasts for i minislots.

Proof of equation (50): We now deal with the case where $2 \leq F(m) = F \leq F_{max}$. The mean rate of the i-th user in the ordering can be written as follows:

$$\begin{aligned}\bar{\mu}_{k(i)}(F,\bar{K},G) =& \mathbb{P}\left\{F \text{ users above } G, k(i)\in\mathcal{F}\right\}\mathbb{E}\Big\{\bar{\mu}_{k(i)}|F \text{ above } G, k(i)\in\mathcal{F}\Big\}\\ &+\sum_{F'=1}^{F-1}\mathbb{P}\left\{F' \text{ users above } G, k(i)\in\mathcal{F}\right\}\mathbb{E}\Big\{\bar{\mu}_{k(i)}|F' \text{ above} G, k(i)\in\mathcal{F}\Big\}\\ &:= \hat{\mu}_i(F)+\sum_{F'=1}^{F-1}\hat{\mu}_i(F').\end{aligned} \tag{66}$$

For the rest of the proof, we will denote the event that user $k(i)$ is scheduled as $k(i)\in\mathcal{F}$. For the first term of the above equation, which corresponds to the event that $F(m)=F$ users get scheduled in the end, we have

$$\hat{\mu}_i(F)=\sum_{m=F}^{\bar{K}}\mathbb{P}\left\{M=m,k(i)\in\mathcal{F},|\mathcal{F}|=F\right\}[T_s-(\beta_p+\beta F+\beta_c m)]^{+}\mathbb{E}\left\{r_{k(i)}(t)\middle|F,g_{k(i)}(t)>G\right\}, \tag{67}$$

where M denotes the duration of the contention period (i.e. how many minislots are used till F users are found with the channel magnitude above the threshold). Since $k(i)$ should be in the set of users that are scheduled, the contention period should not stop before his minislot, that is

$$\mathbb{P}\left\{M=m,k(i)\in\mathcal{F},|\mathcal{F}|=F\right\}=0,\forall 0\le m<i.$$

In addition, we should also have $m\ge F$. Combined with the above, it implies that $m\ge\max\{F,i\}$. Tanking also into account that, by definition $\mathbb{E}\left\{r_{k(i)}(t)\middle|F,g_{k(i)}(t)>G\right\}=\bar{r}_d(F,G)$, the sum in (67) can be then rewritten as follows:

$$\begin{aligned}\hat{\mu}_i(F)=&\mathbb{1}_{\{i\ge F\}}[T_s-(\beta_p+\beta F+i]^{+}\mathbb{P}\left\{M=i,k(i)\in\mathcal{F},|\mathcal{F}|=F\right\}\bar{r}_d(F,G)\\ &+\sum_{m=\max\{F,i+1\}}^{\bar{K}}\mathbb{P}\left\{M=m,k(i)\in\mathcal{F},|\mathcal{F}|=F\right\}[T_s-(\beta_p+\beta F+\beta_c m)]^{+}\bar{r}_d(F,G),\end{aligned} \tag{68}$$

The first term of (68) corresponds to the case where $k(i)$ is the last user to send the signal that its channel gain is greater than G. In this case, exactly $F-1$ out of the $i-1$ previou users must have channel gain above G, therefore

$$\mathbb{P}\Big\{M=i,k(i)\in\mathcal{F},|\mathcal{F}|=F\Big\}=\binom{i-1}{F-1}p^{F}(G)(1-p(G))^{i-F} \tag{69}$$

For the second term of (68), we note that the event in the probability is equivalent to the union of events where (i) the channel magnitude of user $k(i)$ is above G (ii)the channel magnitude of user $k(m)$

is above G and (iii) f users with lower order than i, that is between and including $k(1)$ and $k(i-1)$ and $F-f-2$ between and including $k(i+1)$ and $k(m-1)$ have channel magnitudes above the threshold, for all values of f. The values f is allowed to take should satisfy the following properties:

$$\begin{aligned} &0 \leq f \leq i-1 \\ &f \leq F-2 \\ &F-(f+2) \leq m-1-i \implies f \geq F-1-(m-i). \end{aligned} \tag{70}$$

The first condition in (70) comes from the fact that the number of users before the i-th user is $i-1$, the second from the fact that since $k(i)$ and $k(m)$ should both be scheduled, not more than $F-2$ users higher in the ranking of $k(i)$ should have channel magnitudes above the threshold and the third because there are left $m-i$ users in the ranking between, and not including users $k(i)$ and $k(m)$. We thus have

$$\begin{aligned} \mathbb{P}\left\{M=m, k(i) \in \mathcal{F}, |\mathcal{F}|=F\right\} = p^2(G) \sum_{f=\max\{0,F-1-m+i\}}^{\min\{F-2,i-1\}} \mathbb{P}\Big\{& f \text{ in } \{k(1),...,k(i-1) \text{ above } G, \\ & F-(f+2) \text{ in } \{k(i+1),...,k(m-1) \text{ above } G\Big\}, \end{aligned}$$

which, since the channels are i.i.d. among users reduces to

$$\begin{aligned} &\mathbb{P}\left\{M=m, k(i) \in \mathcal{F}, |\mathcal{F}|=F\right\} = \\ &p^2(G) \sum_{f=\max\{0,F-1-m+i\}}^{\min\{F-2,i-1\}} \mathbb{P}\Big\{f \text{ out of } i-1 \text{ above } G\Big\} \mathbb{P}\Big\{F-(f+2) \text{ out of } m-1-i \text{ above } G\Big\}. \end{aligned}$$

The event in the first term inside the sum happens in $\binom{i-1}{f}$ possible ways with probability $p^f(G)(1-p(G))^{i-1-f}$ each, while the event in the second term happens in $\binom{m-1-i}{F-(f+2)}$ ways, with probability $p^{F-(f+2)}(G)(1-p(G))^{m-i-(F-(f+1))}$ each. Replacing and taking into account that $F \leq m$ (all F users should get scheduled) and we get eventually

$$\begin{aligned} &\mathbb{P}\left\{M=m, k(i) \in \mathcal{F}, |\mathcal{F}|=F\right\} = p(G)^F(1-p(G))^{m-F} \sum_{f=\max\{0,F-1-m+i\}}^{\min\{F,i\}-1} \binom{i-1}{f}\binom{m-1-i}{F-(f+2)}, \\ &\text{for } \max\{F, i+1\} \leq m \leq \bar{K} \end{aligned} \tag{71}$$

We now turn to the second term of (66), i.e. the sum $\sum_{F'=1}^{F-1} \hat{\mu}_i(F')$. This is in fact due to the probability that less than F users can have channel magnitude above the threshold. The event that exactly $F' < F$ users have channel magnitude above the threshold and user $k(i)$ is scheduled is the same as user $k(i)$ having channel magnitude over the threshold and exactly $F'-1$ out of the remaining $\bar{K}-1$ users do. This event can happen in $\binom{\bar{K}-1}{F'-1}$ combinations, each having a probability of $p(g^*)p^{F'-1}(g^*)(1-p(g^*))^{\bar{K}-F'}$. On the other hand, if less than F users are above the threshold then all $\bar{K}$ minislots are used, therefore the contention period lasts for $\bar{K}$ channel uses. Finally, F' users participate in the uplink training. We then have

$$\hat{\mu}_i(F') = \binom{\bar{K}-1}{F'-1} p^{F'}(G)(1-p(G))^{\bar{K}-F'}[T_s - (\beta_p + \beta_c\bar{K} + \beta F')]^+ \mathbb{E}\left\{ r_{k(i)}(t) \middle| F', g_{k(i)}(t) > G \right\}. \tag{72}$$

The result follows combining (66) with (67), (68), (69), (71) and (72).

References

[1] G. Foschini and M. Gans, "On limits of wireless communications in a fading environment when using multiple antennas," *Wireless Personal Communications*, vol. 6, no. 3, pp. 311–335, 1998.

[2] D. Gesbert, M. Kountouris, R. Heath, C.-B. Chae, and T. Salzer, "Shifting the MIMO paradigm," *IEEE Signal Process. Mag.*, vol. 24, no. 5, pp. 36–46, 2007.

[3] H. Weingarten, Y. Steinberg, and S. Shamai, "The Capacity Region of the Gaussian Multiple-Input Multiple-Output Broadcast Channel," *IIEEE Trans. Info. Theory*, vol. 52, no. 9, pp. 3936–3964, Sept. 2006.

[4] D. Love, R. Heath, V. Lau, D. Gesbert, B. Rao, and M. Andrews, "An overview of limited feedback in wireless communication systems," *IEEE J. Sel. Areas Commun.*, vol. 26, no. 8, pp. 1341 –1365, Oct. 2008.

[5] G. Caire, N. Jindal, M. Kobayashi, and N. Ravindran, "Multiuser mimo achievable rates with downlink training and channel state feedback," *IEEE Transactions on Information Theory*, vol. 56, no. 6, Jun. 2010.

[6] M. Kobayashi, N. Jindal, and G. Caire, "Training and Feedback Optimization for Multiuser MIMO Downlink," *IEEE Trans. Commun.*, vol. 59, no. 8, pp. 2228–2240, Aug.t 2011.

[7] L. Song, Z. Han, Z. Zhang, and B. Jiao, "Non-cooperative feedback-rate control game for channel state information in wireless networks," *IEEE Journal on Selected Areas in Communications*, vol. 30, no. 1, pp. 188 –197, Jan. 2012.

[8] H. Boche and M. Wiczanowski, "The interplay of link layer and physical layer under MIMO enhancement: benefits and challenges," *IEEE Wireless Commun. Mag.*, vol. 13, no. 4, pp. 48–55, Aug. 2006.

[9] ——, "Optimization-Theoretic Analysis of Stability-Optimal Transmission Policy for Multiple-Antenna Multiple-Access Channel," *IEEE Trans. Signal Process.*, vol. 55, no. 6, pp. 2688–2702, Jun. 2007.

[10] C. Swannack, E. Uysal-Biyikoglu, and G. Wornell, "Low Complexity Multiuser Scheduling for Maximizing Throughput in the MIMO Broadcast Channel," in *Allerton Conference on Communications, Control, and Computing*, Monticello, IL, USA, 2004.

[11] M. Kobayashi and G. Caire, "An iterative water-filling algorithm for maximum weighted sum-rate of Gaussian MIMO-BC," *IEEE J. Sel. Areas Commun.*, vol. 24, no. 8, pp. 1640–1646, Aug. 2006.

[12] C. Wang and R. Murch, "Optimal downlink multi-user MIMO cross-layer scheduling using HOL packet waiting time," *IEEE Transactions on Wireless Communications*, vol. 5, no. 10, pp. 2856–2862, Oct. 2006.

[13] M. Kobayashi, G. Caire, and D. Gesbert, "Transmit diversity versus opportunistic beamforming in data packet mobile downlink transmission," *IEEE Transactions on Communications,*, vol. 55, no. 1, pp. 151–157, Jan. 2007.

[14] M. Kobayashi and G. Caire, "Joint Beamforming and Scheduling for a Multi-Antenna Downlink with Imperfect Transmitter Channel Knowledge," *IEEE J. Sel. Areas Commun.*, vol. 25, no. 7, pp. 1468–1477, Sep. 2007.

[15] K. Huang and V. Lau, "Stability and delay of zero-forcing SDMA with limited feedback," *IEEE Trans. Inf. Theory*, vol. 58, no. 10, pp. 6499–6514, Oct. 2012.

[16] J. Jose, A. Ashikhmin, P. Whiting, and S. Vishwanath, "Channel Estimation and Linear Precoding in Multiuser Multiple-Antenna TDD Systems," *IEEE Transactions on Vehicular Technology*, vol. 60, no. 5, pp. 2102–2116, Jun. 2011.

[17] T. Yoo and A. Goldsmith, "On the optimality of multiantenna broadcast scheduling using zero-forcing beamforming," *IEEE Journal on Selected Areas in Communications*, vol. 24, no. 3, pp. 528–541, Mar. 2006.

[18] S. Sesia, I. Toufik, and M. Baker, *LTE, The UMTS Long Term Evolution: From Theory to Practice*. Wiley, 2009.

[19] S. Lakshminarayana, B. Li, M. Assaad, A. Eryilmaz, and M. Debbah, "A fast-CSMA based distributed scheduling algorithm under SINR model," in *IEEE nternational Symposeum on Information Theory (ISIT)*, Princeton, NJ, USA, 2012, pp. 1598–1602.

[20] M. Karaca, E. Ekici, and Ö. Erçetin, "Throughput-optimal distributed algorithm for wlans over fading channels," *ArXiv e-prints*, vol. abs/1301.2750, 2013.

[21] B. Li and A. Eryilmaz, "Distributed channel probing for efficient transmission scheduling over wireless fading channels," in *IEEE International Conference on Computer Communications (INFOCOM)*, Mar. 2012, pp. 3026 –3030.

[22] A. Eryilmaz, R. Srikant, and J. R. Perkins, "Stable scheduling policies for fading wireless channels," *IEEE/ACM Transactions on Networking*, vol. 13, no. 2, pp. 411–424, Apr. 2005.

[23] K. Kar, X. Luo, and S. Sarkar, "Throughput-optimal scheduling in multichannel access point networks under infrequent channel measurements," *IEEE Transactions on Wireless Communications*, vol. 7, no. 7, pp. 2619–2629, Jul. 2008.

[24] J. Hoydis, M. Kobayashi, and M. Debbah, "Optimal Channel Training in Uplink Network MIMO Systems," *IEEE Transactions on Signal Processing*, vol. 59, no. 6, pp. 2824–2833, Jun. 2011.

[25] 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.

[26] L. Tassiulas and A. Ephremides, "Stability properties of constrained queueing systems and scheduling policies for maximum throughput in multihop radio networks," *IEEE Trans. Autom. Control*, vol. 37, no. 12, pp. 1936 –1948, Dec. 1992.

[27] A. L. Stolyar, "Maxweight Scheduling in a Generalized Switch: State Space Collapse and Workload Minimization in Heavy Traffic," *The Annals of Applied Probability*, vol. 14, no. 1, pp. 1–53, 2004.

[28] P. Chaporkar, A. Proutiere, H. Asnani, and A. Karandikar, "Scheduling with limited information in wireless systems," in *ACM International Symposium on Mobile Ad Hoc Networking and Computing (MobiHoc)*, 2009, pp. 75–84.

[29] T. Marzetta, "Noncooperative cellular wireless with unlimited numbers of base station antennas," *IEEE Transactions on Wireless Communications*, vol. 9, no. 11, pp. 3590–3600, Sep. 2010.

[30] D. Shah and J. Shin, "Randomized scheduling algorithm for queueing networks." *Annals of Applied Probability*, vol. 22, no. 1, pp. 128–171, 2012.

[31] S. S. Nam, H.-C. Yang, M.-S. Alouini, and D. I. Kim, "Exact capture probability analysis of GSC receivers over i.n.d. rayleigh fading channels," in *IEEE International Symposium on Information Theory (ISIT), 2013*, July 2013, pp. 56–60.

[32] ——, "Technical report: An MGF-based unified framework to determine the joint statistics of partial sums of ordered i.n.d. random variables," *ArXiV e-prints*, vol. abs/1307.8199, 2013.

[33] N. Jindal, J. Andrews, and S. Weber, "Multi-antenna communication in ad hoc networks: Achieving mimo gains with simo transmission," *IEEE Transactions on Communications*, vol. 59, no. 2, pp. 529–540, Feb. 2011.

[34] D. Gesbert and M.-S. Alouini, "How much feedback is multi-user diversity really worth?" in *IEEE Int. Conf. Commun. (ICC)*, vol. 1, Paris, France, 2004, pp. 234–238.

[35] V. Hassel, D. Gesbert, M.-S. Alouini, and G. Oien, “A Threshold-Based Channel State Feedback Algorithm for Modern Cellular Systems,” *IEEE Trans. Wireless Commun.*, vol. 6, no. 7, pp. 2422 –2426, Jul. 2007.

[36] M. Karaca, Y. Sarikaya, O. Ercetin, T. Alpcan, and H. Boche, “Joint Opportunistic Scheduling and Selective Channel Feedback,” *IEEE Trans. Wireless. Commun.*, vol. 12, no. 6, pp. 3024–3034, Jun. 2013.

[37] T. Tang, R. W. Heath, S. Cho, and S. Yun, “Opportunistic feedback for multiuser MIMO systems with linear receivers,” *IEEE Trans. Commun.*, vol. 55, no. 5, pp. 1020–1032, 2007.

[38] R. Couillet and M. Debbah, *Random Matrix Methods for Wireless Communications*. New York, NY, USA: Cambridge University Press, 2011.

[39] T. Marzetta and B. Hochwald, “Capacity of a mobile multiple-antenna communication link in Rayleigh flat fading,” *IEEE Transactions on Information Theory*, vol. 45, no. 1, pp. 139–157, Jan. 1999.

[40] B. Hassibi and T. Marzetta, “Multiple-antennas and isotropically random unitary inputs: the received signal density in closed form,” *IEEE Transactions on Information Theory*, vol. 48, no. 6, pp. 1473–1484, Jun. 2002.