

Satisfying Demands in a Multicellular Network: An Universal Power Allocation Algorithm

Sreenath Ramanath¹, Veeraruna Kavitha^{1,2}, Merouane Debbah³

¹INRIA, Sophia-Antipolis, France; ²Universite d'Avignon, Avignon, France;

³SUPELEC, Paris, France;

Abstract

Power allocation to satisfy user demands in the presence of large number of interferers in a multicellular network is a challenging task. Further, the power to be allocated depends upon the system architecture, for example upon components like coding, modulation, transmit precoder, rate allocation algorithms, available knowledge of the interfering channels, etc. This calls for an algorithm via which each base station in the network can simultaneously allocate power in a decentralized way to their respective users so as to meet their demands (when they are within the achievable limits), using whatever information is available of the other users. The goal of our research is to propose one such algorithm which in fact is universal: the proposed algorithm works from a fully co-operative setting to almost no co-operation (becomes completely decentralized) and or for any configuration of modulation, rate allocation, etc. schemes. The algorithm asymptotically satisfies the user demands, running simultaneously and independently within a given total power budget at each base station. Further, it requires minimal information to achieve this: every base station needs to know its own users demands, its total power constraint and the transmission rates allocated to its users in every time slot. We formulate the power allocation problem in a system specific game theoretic setting, define system specific capacity region and analyze the proposed algorithm using ordinary differential equation (ODE) framework. Simulations confirm the effectiveness of the proposed algorithm.

Index Terms

Cellular networks; MIMO; Power Allocation; Stochastic Approximation; Ordinary Differential Equations;

I. INTRODUCTION

Multi-input multi-output (MIMO) combined with network densification promise improved network coverage and capacity for mobile broadband access. But, due to an increased number of transmit antennas and or the proximity of base stations (BS), users at cell edges experience a higher degree of interference from neighboring base stations.

Network MIMO [2] or other forms of base station co-operation [3], [4] can alleviate interference (example by exchanging information regarding the channel state), but demand a high capacity backhaul network and computation power. Often, the available backhaul capacity is limited and needs to be utilized judiciously, for example, as in the case of small cell networks [5] (eg. micro, pico and femto).

Mobile broadband users run various applications. Simple chat sessions demand very low data rates, while, streaming applications demand high data rates. Base stations need to satisfy users with varying demands. Since interference plays a key role in deciding the rate achievable at the user's end, various forms of BS co-operation can be implemented at the BS to cater to user application demands. The given power budget at the base station can be effectively utilized in conjunction with various levels of BS co-operation to satisfy user demands.

If the cellular network had unlimited backhaul capacity combined with good channel estimation algorithms, every BS can have complete channel state information at transmitter (CSIT). If the channel estimation has errors it

will have partial CSIT. With limited backhaul capacity, a BS may have only statistical information of the channel states of its neighbors. With good estimation techniques, a BS may have local CSIT, i.e., the knowledge of the channel state of its users. Further, it may not even have access to the local CSIT, instead might have to work using a minimal feedback information: an ACK-NACK feedback is provided from its users indicating the QoS of the current transmissions. So, depending upon the situation, we have different scenarios and thus various degrees of co-operation amongst base stations. However the goal is the same: satisfy the demands of all the users. We may require higher power profiles to satisfy the same demands when we have to work with lesser information. More the BS knows about the underlying channel, the better it utilizes the available power. Further diverse situations can arise because of the system configuration like modulation, precoding, channel coding, resource allocation etc.

For a given vector of power constraints at various base stations, Shannon capacity gives the maximum achievable rate, i.e., the capacity region. This is an upper bound. We define "system specific capacity region" (achievable rate region of a given system) which depend on coding (space-time, channel), modulation, channel state information availability, synchronization, feedback errors and many other things. Given a system architecture with a chosen set of parameters which define its rate allocation, modulation, etc, the achievable rates are usually inferior to the theoretical rates and the system specific capacity region is defined based on these rates. The system-specific capacity region for the same power constraint varies: for example it shrinks if the number of supported discrete rates reduce. Thus, the power allocated to any user to achieve the same demand rate varies with the set of system parameters.

The main contribution of this paper is an universal algorithm which can work with many of the systems mentioned above. It satisfies asymptotically the demands of all the users irrespective of the system in which it is operating, albeit with different power profiles. *Each base station requires minimal information: its user's demands, its total power constraint and the current transmission rates to its users.* The current transmission rates are decided by the serving base stations either using complete CSIT (algorithm can also be used as a centralized scheme in this case) or has to be estimated completely blindly (a fully distributed algorithm in this case) or using some partial information. The following are the contributions of this paper:

- A system specific game theoretic problem formulation using the system specific capacity region.
- A *Stochastic Approximation* based decentralized power allocation algorithm in an interference limited multicell networks.
- Various properties (eg. convergence) of the proposed algorithm is analyzed using an ODE framework.
- Simulation results demonstrate the effectiveness of the proposed algorithm for a variety of systems.

Related Work: Power control or power allocation is a very basic and important problem in any inference limited cellular network. For an excellent survey on power control in wireless networks, the reader is referred to [8] and the references there-in. Much of the study on cellular network power control started in the early 90's with the seminal work of [9], followed by a series of results that solve the basic problem formulation, where transmit power is the only variable, constrained by fixed and feasible target SIRs, and optimized to minimize the total power [10], [11]. Next set of problems in power control address variable SIR to achieve a certain QoS constraint to support mixed traffic types (eg. [12], [13], [14]). Opportunistic power control (eg. [15], [16]) exploit channel variations to improve performance and fairness. In the recent years, tools from game theory have been used to address many variants of the power control problem (eg. [17], [18]). And finally there are problems which address power control jointly with beam-forming or Base station assignment or scheduling (eg. [20], [21], [22]). In recent years, several authors have addressed distributed power control strategies with various levels of co-operation for a given system configuration (eg. [23], [24], [25], [26], [28], [29]). Typically, the design objective in distributed processing is to *maximize* the total sum rate of all the users subject to BS power constraints or to *minimize* the total transmit power

satisfying some SINR constraints of the users.

Most of the existing algorithms aim at either optimizing the total power spent keeping the QoS above a required level and or optimize the QoS while keeping the power utilized within a given budget. In our paper, we develop an universal algorithm, that asymptotically meets the demands (in the form of average transmission rates) on average. The power profile that can meet the demands of all the users (if feasible) depends upon the system configuration like the modulation scheme, rate allocation scheme, precoding, channel state information availability and many other system related factors. The proposed algorithm, when working on a specific system converges to a system specific power profile which meets the given demands (QoS) of the users. More importantly, unlike many existing algorithms, it requires minimal information at the transmitters: rates at which the information is correctly transmitted to the user in every slot. Data is pumped out from the transmitter and hence these rates are readily known to the transmitter.

Organization: We introduce the system model in section II. In section III, we describe the system specific problem formulation. The algorithm and its analysis is presented in section IV. Section V provides simulations and the paper concludes in section VI.

Notations: Boldface lower-case symbols represent vectors, capital boldface characters denote matrices (\mathbf{I}_N is the $N \times N$ identity matrix). Hermitian transpose is denoted $(\cdot)^H$ while $\text{tr}[\mathbf{X}]$ represents the trace of matrix \mathbf{X} . All logarithms are base-2 logarithms. Small letters represent the scalars. Let a_k represent the k^{th} component of the vector \mathbf{a} . If the vector is already indexed like for example in \mathbf{p}_j , then $p_{k,j}$ represents its k^{th} component. Let $(\mathbf{p}.\mathbf{s})$ represent the component-wise product, i.e., $(\mathbf{p}.\mathbf{s})_k = p_k s_k$ for all k while $\sqrt{\mathbf{p}}$ represents component wise square root. $\mathbf{E}[\cdot]$ denotes expectation and \mathbf{E}_s is expectation w.r.t to s when conditioned (if any) on the other random variables. Other notations are introduced as and when required.

II. SYSTEM MODEL

We consider a multicell MIMO system. Each base station has M transmit antennas and is communicating with K single-antenna users (see figure 1). Every user experiences both intra-cell (transmissions from parent BS) and inter-cell (transmissions from neighboring BS) interference. Each user in a cell demands a certain rate and all these rates have to be jointly satisfied by the BS (present in the cell) while operating within a total power constraint.

Let $\mathbf{H}_{j,\bar{j}}$ represent the $K \times M$ channel matrix, when the users in cell j receive signals from the BS of cell \bar{j} and let its elements be given by zero-mean unit-variance i.i.d. complex Gaussian entries. Let \mathbf{n}_j represent the additive white Gaussian noise at the receivers of cell j , \mathbf{x}_j be the M length transmit vector in cell j and $\gamma_j \in [0, 1]$ be the interference factor, representative of the level of interference. For example, as base stations become denser, interference increases and hence $\gamma_j \rightarrow 1$. The signal vector (of length K) received by users in cell j is given by,

$$\mathbf{y}_j = \mathbf{H}_{j,j}\mathbf{x}_j + \sum_{\bar{j}=1, \bar{j} \neq j}^J \gamma_{\bar{j}} \mathbf{H}_{j,\bar{j}}\mathbf{x}_{\bar{j}} + \mathbf{n}_j \text{ for all } j \leq J.$$

In the above expression the first term represents the useful signal part as well as the intra-cell interference while the second term (summation) represents the inter-cell interference to the j^{th} cell from its neighbors.

If \bar{P}_j represents the total power constraint in cell j , then $\text{tr}(\mathbf{E}[\mathbf{x}_j\mathbf{x}_j^H]) \leq \bar{P}_j$ to satisfy the power constraint. As an example if the BS in cell j uses power levels specified by \mathbf{p}_j and a precoding matrix \mathbf{G}_j (of size $M \times K$), then the transmit vector is given by $\mathbf{x}_j = \mathbf{G}_j(\sqrt{\mathbf{p}_j}.\mathbf{s}_j)$ (\mathbf{s}_j is a K length independent symbol vector of zero mean and unit variance components). In this case the power constraint leads to,

$$\text{tr}(\mathbf{E}[\mathbf{x}_j\mathbf{x}_j^H]) \leq \text{tr}(\mathbf{E}[\mathbf{G}_j\sqrt{\mathbf{p}_j}(\mathbf{G}_j\sqrt{\mathbf{p}_j})^H]) \leq \bar{P}_j \text{ for any } j.$$

Given a precoding scheme, this constraint can equivalently be represented by (for a possibly different \bar{P}_j)

$$\sum_k p_{k,j} \leq \bar{P}_j.$$

The symbol, $y_{k,j}$, received by the user k of cell j is,

$$y_{k,j} = \mathbf{h}_{k,j,j}^H \mathbf{x}_j + \sum_{i=1, i \neq k}^K \mathbf{h}_{i,j,j}^H \mathbf{x}_j + \sum_{\bar{j}=1, \bar{j} \neq j}^J \sum_{i=1}^K \gamma_{\bar{j}} \mathbf{h}_{i,j,\bar{j}}^H \mathbf{x}_{\bar{j}} + n_{k,j} = u_{k,j} + i_{k,j,j} + \sum_{\bar{j} \neq j} i_{k,j,\bar{j}} + n_{k,j} \quad (1)$$

where $\mathbf{h}_{k,j,\bar{j}}$, the k^{th} row of matrix $\mathbf{H}_{j,\bar{j}}$, represents the M length channel vector for user k of cell j as received from the BS of cell \bar{j} . In the above, $u_{k,j}$, $i_{j,j,k}$ and $i_{k,j,\bar{j}}$ respectively represent the useful, intra-cell interference and inter-cell interference signal, respectively.

System with No Precoding:

This paper proposes an algorithm which works for any system in general. By *system*, we mean a particular multicell network with a given configuration like, precoding scheme, channel coding, resource allocation etc. We will derive the exact received signal characteristics for one such example system. The received signal characteristics of the others system can be derived in a similar way. We consider a system with no precoding (for example, systems which does not have access to channel state information). Further we consider a system with $M = K$ and with $\mathbf{x}_j = (\sqrt{p_j} \cdot \mathbf{s}_j)$. The average power in the useful, intra-cell, inter-cell interference signals of the received signal (after channel coding at the transmitter and channel decoding at the receiver) after averaging w.r.t. to the symbol statistics $\{\mathbf{s}_j\}$ for any given channel state:

$$\begin{aligned} \mathbf{E}_{\mathbf{s}_j, 1 \leq j \leq J} [|u_{k,j}|^2] &= p_{k,j} |h_{k,j,j,k}|^2, \\ \mathbf{E}_{\mathbf{s}_j, 1 \leq j \leq J} [|i_{k,j,j}|^2] &= \sum_{\bar{k} \neq k} p_{\bar{k},j} |h_{k,j,j,\bar{k}}|^2 \text{ and} \\ \mathbf{E}_{\mathbf{s}_j, 1 \leq j \leq J} [|i_{k,j,\bar{j}}|^2] &= \sum_{\bar{k}} \gamma_{\bar{j}} p_{\bar{k},\bar{j}} |h_{k,j,\bar{j},\bar{k}}|^2 \end{aligned} \quad (2)$$

where, $h_{k,j,\bar{j},\bar{k}}$ is the $(k, \bar{k})^{\text{th}}$ component of the matrix $\mathbf{H}_{j,\bar{j}}$. In the above we used $E[s_{k,j} s_{k',j'}^*] = 1_{\{k=k', j=j'\}}$.

III. SYSTEM SPECIFIC PROBLEM FORMULATION

Every BS has to meet its users demands, for example BS j has to meet the demand rates represented by $\mathbf{r}_j := \{r_{k,j}, k \leq K\}$. It has to tune its power levels \mathbf{p}_j to achieve this. But the rates achieved will also depend upon the powers used by the other base stations. Our goal is to find a simple decentralized power allocation algorithm which runs independently and simultaneously at all the base stations and tunes the power levels to achieve the demands of the users. The power levels depend upon the system configuration (for example, a given channel precoding scheme and a rate allocation scheme). We consider some interesting example systems described in the following paragraphs. The brief three part code, I-II-III, which represents a system is given in Table 1 and is explained below:

1) The first part (I) represents the **availability of channel state (CS) information at transmitter**¹: a) **A** represents an asymptotic (large number of antennas/users) system, where achievable rates for all most all CS are approximated

¹One can also consider systems which have an estimate of the CS.

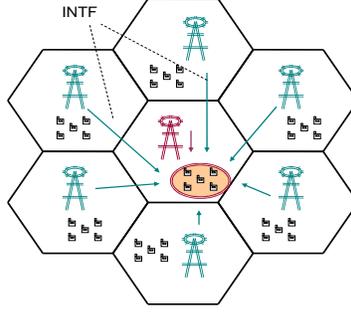


Fig. 1. 2D Wyner model

I CSIT		II TX rate		III Precoder	
A	Asymptotic	I	Ideal	ZF	Zero-forcing
C	Complete CSIT	D	Discrete	NO	No precoder
L	Local CSIT	RA	Rate adaptation		
N	No CSIT	RAE	Rate adaptation with errors		

Table 1: System specification (I-II-III)

by a constant limit (see [31] and references there in), b) **C** for systems having complete CSIT, c) **L**, systems with local CSIT, i.e., BS j knows $\mathbf{H}_{j,j}$ part of the CS, d) **N**, systems with no CS information.

2) The second part (II) represents **the transmission rates used at the system**²: a) **I** for ideal systems which can channel code to achieve any feasible rate, b) **D** for the those systems which can only operate at one of the discrete rates in the set $\mathbb{R} = \{r_1, r_2, \dots, r_{N_R}\}$ (arranged in decreasing order), c) **RA** for systems which estimate the current rate (i.e., pick the current maximum possible rate from the set \mathbb{R}) using (blind) rate adaptation schemes (eg. [33]) without CSIT, d) **RAE** when there are estimation errors in the rate adaptation algorithm.

3) The third part (III) represents the **precoder**³: a) **ZF** for zero forcing precoding, b) **NO** for no channel precoding.

We now describe some example systems. The brief summary of the example systems considered in this paper is also available in Table 2.

1) *Asymptotic Ideal Rate systems* : In a multicellular system with large number of antennas at the BS and large number of users, the rate for a given CS can be obtained using random matrix theory. For example, in [31] the asymptotic rates are derived for a zero forcing (ZF) precoder. It is shown that for almost all realizations of CS, the rate can be approximated by the expression given below in equation (3). Further, we consider a system in which, the base stations can use channel coding schemes to transmit very close to the theoretical rates. When this system (which we call as *asym-ideal-zeroforcing* or in short **A-I-ZF** according to our notations) uses power profile \mathcal{P} and when the channel state (CS) is \mathcal{H} , the BS j transmits to the user k at rate ([31]) (when $M > K$):

$$R_{k,j}^{\mathbf{A-I-ZF}}(\mathcal{P}, \mathcal{H}) \approx \log \left(1 + \frac{p_{k,j}}{\frac{1}{\beta-1} \sum_{j=1, j \neq k}^J \gamma_j \frac{\text{tr}(\mathbf{P}_j)}{K} + \sigma_{k,j}^2} \right) \quad (3)$$

where, $\beta = M/K$ is the ratio of number of transmit antennas on the BS to the number of users and $\gamma_j \in (0, 1)$

²We are illustrating the concepts using simple rate allocation schemes. One can also extend it to other rate allocations, for example schemes that incorporate fairness.

³One can also consider other types of precoders (eg. MMSE). Our analysis and proofs hold for these configurations as long as they satisfy the simple assumptions A.1 to 4.

represents the interference from cell \bar{j} . This rate is same for almost all CS \mathcal{H} as it is an asymptotic rate. The above expression is for a Zero-Force (ZF) precoder. Similar expression is available for the case with $M = K$ in [31].

2) *Ideal rates using complete CSIT* : If the number of antennae/number of users is not large enough, the asymptotic results are not accurate. If BS has access to CSIT (and if each BS could channel code to obtain rates closer to the ideal rate) then with ZF precoder it transmits at rate:

$$R_{k,j}^{\mathbf{C-I-ZF}}(\mathcal{P}, \mathcal{H}) = \log \left(1 + \frac{p_{k,j}^t}{\sum_{\bar{j}=1, \bar{j} \neq j}^J \gamma_{\bar{j}} \frac{\text{tr}(\mathbf{H}_{j,\bar{j}}^t Q_{\bar{j}} \mathbf{H}_{j,\bar{j}}^{tH})}{K} + \sigma_{k,j}^2} \right); \quad Q_{\bar{j}} := \mathbf{H}_{\bar{j}}^{tH} (\mathbf{H}_{\bar{j}}^t \mathbf{H}_{\bar{j}}^{tH})^{-1} \mathbf{P}_{\bar{j}}^t (\mathbf{H}_{\bar{j}}^t \mathbf{H}_{\bar{j}}^{tH})^{-1} \mathbf{H}_{\bar{j}}^t$$

For the same configuration, but without transmitter precoding, the the instantaneous transmission rate (as obtained using Shannon's capacity expression), from equation (1) is:

$$\begin{aligned} R_{k,j}^{\mathbf{C-I-NO}}(\mathcal{P}, \mathcal{H}) &= \log(1 + \eta_{k,j}) \text{ where} & (4) \\ \text{SINR, } \eta_{k,j} &:= \frac{\mathbf{E}[|u_{k,j}|^2]}{\rho_{k,j}} \text{ with} \\ \text{noise + interference, } \rho_{k,j} &:= \sum_{\bar{j}} \mathbf{E}[|i_{k,j,\bar{j}}|^2] + \sigma_{k,j}^2. \end{aligned}$$

3) *Finite number of Rates* : Ideal rate systems are not realistic, they can't be implemented in practice. We consider a system, in which the BS can transmit at one of the available discrete rates from the set \mathbb{R} . When transmitter has CSIT, it knows the exact theoretical rate and hence will pick the largest rate from set \mathbb{R} that is smaller than the current theoretical rate:

$$R_{k,j}^{\mathbf{C-D-ZF}}(\mathcal{P}, \mathcal{H}) = \inf_{r \in \mathbb{R}} \{r \leq R_{k,j}^{\mathbf{C-I-ZF}}(\mathcal{P}, \mathcal{H})\}, \quad (5)$$

$$R_{k,j}^{\mathbf{C-D-NO}}(\mathcal{P}, \mathcal{H}) = \inf_{r \in \mathbb{R}} \{r \leq R_{k,j}^{\mathbf{C-I-NO}}(\mathcal{P}, \mathcal{H})\}. \quad (6)$$

If the asymptotic system described above chooses transmission rates from a finite set, the BS always transmits at the rate:

$$R_{k,j}^{\mathbf{A-D-ZF}}(\mathcal{P}, \mathcal{H}) = \inf_{r \in \mathbb{R}} \{r \leq R_{k,j}^{\mathbf{A-I-ZF}}(\mathcal{P}, \mathcal{H})\}, \quad (7)$$

4) *Rate adaptation Without CSIT* : It is once again not realistic to assume the knowledge of complete CSIT. There are many schemes that estimate the rate blindly or using some partial CSIT (eg. [33]). The algorithm proposed in the next section is a very general algorithm and works with all those systems which satisfy assumptions **A.1-3** mentioned in the next section. These are quite simple assumptions and most of the systems will be able to satisfy these and hence the algorithm works for majority of the blind/partial CSIT systems.

We explain one such blind system wherein, the BS estimates the transmission rates without knowledge of CSIT. Each time, the BS begins by attempting at the highest available rate r_1 . If the data is not received correctly (information obtained via a feedback from the receiver), the BS sends some more information about the same data packet so that the overall rate now is the second highest r_2 . This procedure repeats until the two agree upon the correct rate. We assume that this rate adaptation system is always successful, i.e, it can estimate the instantaneous rates without errors. Such a system does not require CSIT, however the final rate at which the transmission takes place depends upon the current channel state in exactly the same way as in the case of **C-D** (or **A-D** for large

antenna and user case and note there is no channel coding in this case as there is no CSIT) and hence,

$$R_{k,j}^{\text{N-RA-NO}}(\mathcal{P}, \mathcal{H}) = R_{k,j}^{\text{C-D-NO}}(\mathcal{P}, \mathcal{H})$$

5) *Rate Adaptation with local CSIT*: All the base stations have local CSIT, i.e., BS j knows the $\mathbf{H}_{j,j}$ part of CS. However they can't estimate the current rates just based on local CSIT. So, they once again use rate adaptation technique as in the system (4). They can however design a better system by using for example a zero forcing precoder. In this case, as in system (4) the rate will be adapted to the actual underlying rate and hence will be same as that in **C-D-ZF**:

$$R_{k,j}^{\text{L-RA-ZF}}(\mathcal{P}, \mathcal{H}) = R_{k,j}^{\text{C-D-ZF}}(\mathcal{P}, \mathcal{H})$$

6) *Rate Adaptation with errors* : There can be some errors in rate adaptation algorithm of system (4) or (5). In this case

$$\begin{aligned} R_{k,j}^{\text{N-RAE-NO}}(\mathcal{P}, \mathcal{H}) &= R_{k,j}^{\text{N-RA-NO}}(\mathcal{P}, \mathcal{H}) - E_{k,j}(R_{k,j}^{\text{N-RA-NO}}(\mathcal{P}, \mathcal{H})) \\ R_{k,j}^{\text{L-RAE-ZF}}(\mathcal{P}, \mathcal{H}) &= R_{k,j}^{\text{L-RA-ZF}}(\mathcal{P}, \mathcal{H}) - E_{k,j}(R_{k,j}^{\text{L-RA-ZF}}(\mathcal{P}, \mathcal{H})) \end{aligned}$$

where (assuming independent errors) $E_{k,j}(\bar{r})$ can take values in the subset $\mathbb{R} \cap \{r \leq \bar{r}\}$ with a given probability distribution.

System Description	$R_{k,j}^{\text{sys}}(\mathcal{P}, \mathcal{H})$
A-I-ZF : Large no. of antennas and users. We assume $M > K$. Asymptotic rates of [31] approximate the instantaneous rates for almost all CS. Transmission takes place at the instantaneous (ideal) rates itself. Zero forcing (ZF) TX precoder.	$\log \left(1 + \frac{p_{k,j}}{\frac{1}{\beta-1} \sum_{\bar{j}=1, \bar{j} \neq j}^J \gamma_{\bar{j}} \frac{\text{tr}(\mathbf{P}_{\bar{j}})}{K} + \sigma_{k,j}^2} \right)$
A-D-ZF : Similar to A-I-ZF, but TX rate allocation from discrete set \mathbb{R}	$\inf_{r \in \mathbb{R}} \{r \leq R_{k,j}^{\text{A-I-ZF}}(\mathcal{P}, \mathcal{H})\}$
C-I-ZF : Number of antennae/users is not large enough. Asymptotic results are not accurate. Every BS has access to CSIT and hence can compute the theoretical limit rate. TX possible at ideal rates. ZF TX precoder.	$\log \left(1 + \frac{p_{k,j}}{\sum_{\bar{j}=1, \bar{j} \neq j}^J \gamma_{\bar{j}} \frac{\text{tr}(\mathbf{H}_{j,j}^t Q_{\bar{j}} \mathbf{H}_{j,j}^t)}{K} + \sigma_{k,j}^2} \right)$ $Q_{\bar{j}} = \mathbf{H}_{\bar{j}}^t (\mathbf{H}_{\bar{j}}^t \mathbf{H}_{\bar{j}}^t)^{-1} \mathbf{P}_{\bar{j}}^t (\mathbf{H}_{\bar{j}}^t \mathbf{H}_{\bar{j}}^t)^{-1} \mathbf{H}_{\bar{j}}^t$
C-D-ZF : Similar to C-I-ZF, but TX rate allocation from discrete set \mathbb{R} .	$\inf_{r \in \mathbb{R}} \{r \leq R_{k,j}^{\text{C-I-ZF}}(\mathcal{P}, \mathcal{H})\}$
C-I-NO : Similar to C-I-ZF, but without Precoder.	$\log \left(1 + \frac{\mathbf{E}_{s_j} [u_{k,j} ^2]}{\sum_{\bar{j}} \mathbf{E}_{s_{\bar{j}}} [z_{k,j,\bar{j}} ^2] + \sigma_{k,j}^2} \right)$
C-D-NO : Similar to C-I-NO, but TX rate allocation from discrete set \mathbb{R} .	$\inf_{r \in \mathbb{R}} \{r \leq R_{k,j}^{\text{C-I-NO}}(\mathcal{P}, \mathcal{H})\}$
N-RA-NO : Rate adaptation w/o CSIT. Uses blind methods to adapt to the correct rate as long as the underlying channel can support the same. No TX precoder.	$R_{k,j}^{\text{C-D-NO}}(\mathcal{P}, \mathcal{H})$
L-RA-ZF : Rate adaptation with local CSIT. BS has local CS, Uses blind methods to assign rates (as in N-RA-NO) and local CS for precoding.	$R_{k,j}^{\text{C-D-ZF}}(\mathcal{P}, \mathcal{H})$
N-RAE-NO : Similar to N-RA-NO, but with rate estimation errors, $E_{k,j}(r)$.	$R_{k,j}^{\text{N-RA-NO}}(\mathcal{P}, \mathcal{H}) - E_{k,j}(R_{k,j}^{\text{N-RA-NO}}(\mathcal{P}, \mathcal{H}))$
L-RAE-ZF : Similar to L-RA-ZF, but with rate estimation errors, $E_{k,j}(r)$.	$R_{k,j}^{\text{L-RA-ZF}}(\mathcal{P}, \mathcal{H}) - E_{k,j}(R_{k,j}^{\text{L-RA-ZF}}(\mathcal{P}, \mathcal{H}))$

Table 2: Some example Systems. Right column gives the rate at which data is transmitted when CS is \mathcal{H} and when system uses the profile \mathcal{P}

A. Game theoretic formulation

As the base stations influence each other, the problem can best be captured using a game theoretic formulation. We begin by introducing the components of the game. The calligraphic letters (for example \mathcal{P}) represent the ensemble of either vectors, matrices or scalars for all the base stations.

Power profile, $\mathcal{P} := \{p_{k,j}\}_{k \leq K, j \leq J}$, represents the vector comprising of the powers used at all the base stations and for all the users. Recall, $p_{k,j}$ represents the power used by the BS of cell j for user k in cell j .

Channel State (CS), $\mathcal{H} := \{\mathbf{H}_{1,1}, \mathbf{H}_{1,2}, \dots, \mathbf{H}_{J,J}\}$, arranged as a matrix of dimension $KJ \times MJ$, represents the channel state of the entire system. If the transmitter has the knowledge of CS, we refer to it as CSIT (Channel State information at Transmitter).

Rate for a given power profile and system, $R_{k,j}^{sys}(\mathcal{P}, \mathcal{H})$, represents the rate at which the data is transmitted in system represented by sys (eg. **N-RAE-NO**) when the base stations use powers \mathcal{P} and when the CS is \mathcal{H} . These rates were discussed in the beginning of this section for some example systems and the same are also listed in the right column of the Table 2.

Average Rate for a given system and power profile, is the rate that is achieved on average when a given system uses the power profile \mathcal{P} is:

$$R_{avg,k,j}^{sys}(\mathcal{P}) = \mathbf{E}_{\mathcal{H}}[R_{k,j}^{sys}(\mathcal{P}, \mathcal{H})]. \quad \text{Let } R_{avg}^{sys} := \{R_{avg,k,j}^{sys}\}_{k,j}.$$

Power constraint ($\mathcal{P} \leq \bar{\mathcal{P}}$) We use \leq in a special manner to facilitate defining the power constraint. We say a power profile \mathcal{P} is "less than or equal to" and hence satisfies the constraint defined in terms of another power profile $\bar{\mathcal{P}}$ if the two profiles satisfy the constraints for each base station in the following manner:

$$\sum_k p_{k,j} \leq \sum_k \bar{p}_{k,j} \text{ for all } j \leq N.$$

System Specific Capacity Region for any given power profile constraint $\bar{\mathcal{P}}$ and a system sys is defined as the collection of all possible tuple of average rates while using powers that satisfy the constraints defined in terms of $\bar{\mathcal{P}}$, i.e.,

$$\mathcal{C}^{sys}(\bar{\mathcal{P}}) := \left\{ \{R_{k,j}\} \in \mathbf{R}^{NK} : \text{for all } k,j \ R_{k,j} = R_{avg,k,j}^{sys}(\mathcal{P}) \text{ for some } \mathcal{P} \text{ with } \mathcal{P} \leq \bar{\mathcal{P}} \right\}.$$

This region is different for different systems. For a system with ideal rates the capacity region coincides with the theoretical one. A system with discrete rates cannot achieve the maximum possible rate and hence its capacity region shrinks. It further depends upon \mathbb{R} , the set of supported rates. Similarly, if the system has estimation errors, the capacity region shrinks further.

Utilities and Players : Each BS j is a player and its strategy is K -dimensional power vector, $\mathbf{p}_j := [p_{1,j}, \dots, p_{K,j}]$. Note that $\mathcal{P} = [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_J]$. Define the utility of player j (i.e., that of the BS j) as,

$$U_j^{sys}(\mathbf{p}_j, \mathcal{P}_{-j}) := \sum_k \min \left\{ R_{avg,k,j}^{sys}(\mathbf{p}_j, \mathcal{P}_{-j}), r_{k,j} \right\} \text{ with } \mathcal{P}_{-j} := [\mathbf{p}_1, \mathbf{p}_2, \dots, \mathbf{p}_{j-1}, \mathbf{p}_{j+1}, \dots, \mathbf{p}_K]. \quad (8)$$

In the above, \mathcal{P}_{-j} is the power vector profile excluding only the power of BS of cell j and $r_{k,j}$ is the demand of user k belonging to cell j . very system with given power constraint $\bar{\mathcal{P}}$, demand vectors $\{r_{k,j}\}$ defines an N -player non cooperative strategic form game:

$$([1, 2 \dots N], \{U_j^{sys}\}_{j \leq N}).$$

The Nash equilibrium (NE) of this game is a power profile \mathcal{P}^* that satisfies,

$$\mathbf{p}_j^* \in \arg \max_{\mathcal{P} \leq \bar{\mathcal{P}}} U_j^{sys}(\mathbf{p}_j, \mathcal{P}_{-j}^*) \text{ for all } j.$$

NE Characterization : From all the above definitions, it is evident that,

Lemma 1: For any given system and power constraints $\bar{\mathcal{P}}$, if the vector of the demands $\{r_{k,j}\}_{k,j}$ is in the

corresponding capacity region $\mathbb{C}^{sys}(\bar{\mathcal{P}})$, then there exists a $\mathcal{P}^* \leq \bar{\mathcal{P}}$, which is a NE satisfying all the demands:

$$R_{avg,k,j}^{sys}(\mathcal{P}^*) = r_{k,j} \text{ for all } k, j. \quad \diamond$$

Thus, when all the base stations use the NE power profile \mathcal{P}^* of Lemma 1, all the users in each cell achieve an average rate which equals their demand, i.e., will be able to receive the information at the demand rate on average. The main aim of this paper is to obtain this NE (time) asymptotically (if required in a completely distributed way) for any given system when the demands satisfy Lemma 1. This NE depends on the system considered (for example higher amount of power may be required if one uses discrete rates in the place of ideal rates) even if the power constraint and demands are same. The proposed algorithm is a general iterative algorithm which works irrespective of the system considered, i.e, the proposed algorithm converges to the system specific NE.

Remark on hypothesis of Lemma 1: It requires that the demands equal one of the average rates of the capacity region. In the next section, we come up with a easily verifiable assumption which ensures this for almost all the systems considered. The only exception is the system with large number of antennas/users and with discrete rates, i.e., systems **A-D-ZF** or **A-D-NO**. In contrast to the other systems, the average rates of these system are also discrete and hence an interesting phenomenon happens with these systems. The algorithm in this case does not converge to a single power profile. Asymptotically it reaches a limit cycle, wherein, the power levels changes in a periodic manner. However the time average of the average rate achieved still equals the demand rates. We could not obtain a proof for this, but is established using simulations. It is interesting to note that the variations in the power levels in the limit cycle are very small (see figure 5), but are sufficient enough to toggle the rates between the two discrete levels neighboring the demand rate.

Set of demand meeting NE, $\mathbb{L}^{sys} \subset \mathbb{C}^{sys}(\bar{\mathcal{P}})$, is the set of NE which meet the demands as in Lemma 1.

In the next section, we present the Decentralized Demand achieving algorithm for Power constrained Multi Cell Networks (DDPMCN).

IV. DECENTRALIZED ALGORITHM : DDPMCN

We consider a quasi-static channel and obtain the NE of Lemma 1 asymptotically by iteratively updating the power profile at the beginning of every slot, during which the CS is assumed constant. Lemma 2 (provided at the end of the section) gives the conditions under which the hypothesis of Lemma 1 are satisfied.

Basic idea⁴: Each BS j in every time slot knows the rates at which data is transmitted to its users, $\{R_{k,j}^{sys}(\mathcal{P}, \mathcal{H})\}_k$. The characterization of these rates is provided for various examples in Table 2. An iterative algorithm can find the average value of it. One can then update the power vectors to force this average towards the demands $\{r_{k,j}\}$.

Let $d_{k,j}^{t+1}$ represent the number of bytes of data transmitted successfully in time slot $t+1$ by the j^{th} base station to its user k divided by the duration of the time slot. This ratio depends upon the power profile of the entire system in the previous slot (\mathcal{P}^t) and the entire CS in the current slot (\mathcal{H}^{t+1}), but ($\mathcal{P}^t, \mathcal{H}^{t+1}$) are only partially known at the base stations. However $d_{k,j}^{t+1}$ is still known at base station j as it is the source that pumps out the data. Infact, it will be precisely equal to $d_{k,j}^{t+1} = R_{k,j}^{sys}(\mathcal{P}^t, \mathcal{H}^{t+1})$ of Table 2 by definition. Let $\{\mu^t\}$ represent the step sizes.

A. DDPMCN algorithm

With $\Pi_{\mathbb{A}}$ representing the projection in to the set \mathbb{A}

$$p_{k,j}^{t+1} = \Pi_{\mathbb{A}_j} [p_{k,j}^t - \mu^t (d_{k,j}^t - r_{k,j})] \text{ with } \mathbb{A}_j := \left\{ \mathbf{p} \in \mathcal{R}^K : \sum_k p_k \leq \bar{P}_j \right\} \text{ and let } \mathcal{A} := \mathbb{A}_1 \times \mathbb{A}_2 \cdots \times \mathbb{A}_N. \quad (9)$$

⁴Most of the cases stochastic approximation algorithms obtain optimum of a function as the zero of its derivative. In contrast, this algorithm obtains the profile that satisfies the demands, as the zero of the function given by the average rate minus demand.

B. Analysis

We obtain the asymptotic analysis of the algorithm using the ordinary differential equations (ODE) approach of [1]. We establish Theorem 1 given below, under the following hypothesis:

A.1 The step sizes, $\{\mu^t\}$, change slowly in the sense that there is a sequence of integers $\alpha_t \rightarrow \infty$ such that

$$\lim_t \sup_{0 \leq i \leq \alpha_t} \frac{\mu^{t+i}}{\mu^t} = 0.$$

A.2 The channel state $\{\mathcal{H}^t\}$ is an independent and identically distributed (iid) sequence with finite mean and variance.

A.3 The instantaneous rates (rates for a given CS) are bounded by the same constant, i.e., $|R_{k,j}^{sys}(\mathcal{P}, \mathcal{H})| \leq B$ for all k, j, \mathcal{P} and \mathcal{H}

A.4 The average rate function $R_{avg,k,j}^{sys}$ is continuous in \mathcal{P} for all k, j .

As a first step, we rewrite the algorithm as in [1]:

$$\begin{aligned} Y_{k,j}^t &:= r_{k,j} - d_{k,j}^t \\ p_{k,j}^{t+1} &= \Pi_{A_j} [p_{k,j}^t + \mu^t Y_{k,j}^t] \end{aligned}$$

Define the filtration,

$$\mathcal{F}_t := \sigma \left(\mathcal{P}^\tau, \{Y_{k,j}^{\tau-1}\}_{k,j}, \text{ for all } \tau \leq t \right)$$

and let \mathbb{E}_t represent the expectation w.r.t. \mathcal{F}_t . Under the assumptions **A.1** and **A.2** clearly the condition expectation

$$\mathbb{E}_t[Y_{k,j}^t] = g_{k,j}^p(\mathcal{P}^t) := r_{k,j} - R_{avg,k,j}(\mathcal{P}^t)$$

for all k, j and t . We will show that the DDPMC trajectory (9) can be approximated by the solution $(\mathcal{P}(t))$ of the following ODE (to be precise a differential inclusion).

$$\dot{p}_{k,j} = g_{k,j}^p(\mathcal{P}) + z_{k,j}(\mathcal{P}) \text{ for all } k, j \quad (10)$$

here $z_{k,j}(\mathcal{P})$ represents the projection term. Define the limit set of this ODE :

$$\mathbb{L}^{ODE} := \lim_{t \rightarrow \infty} \cup_{\mathcal{P} \in \mathcal{A}} \{\mathcal{P}(s) : s \geq t \text{ and } \mathcal{P}(0) = \mathcal{P}\}.$$

The δ -neighborhood of this set is defined as:

$$\mathbb{B}_\delta(\mathbb{L}^{ODE}) := \{\mathcal{P} : |\mathcal{P} - \bar{\mathcal{P}}| \leq \delta \text{ for some } \bar{\mathcal{P}} \in \mathbb{L}^{ODE}\}$$

Theorem 1 establishes that the trajectory ultimately spends time in this limit set. We first establish the theorem and later study the systems of previous section using this Limit set. The proof is given in Appendix.

Theorem 1: Assume **A.1-4**. Then for every $\delta > 0$, the fraction of time the tail of the algorithm (for any initial power profile with $\tilde{\mathcal{P}} < \bar{\mathcal{P}}$)

$$\{\mathcal{P}^\tau\}_{\tau \geq t} \text{ with initializations } \mathcal{P}^t = \tilde{\mathcal{P}}$$

spends in the δ -neighborhood of the limit set $\mathbb{B}_\delta(\mathbb{L}^{ODE})$ tends to one (in probability) as $t \rightarrow \infty$.

By the above theorem DDPMC trajectory ultimately spends most of its time in the limit set of the ODE. If further, for the system under consideration, the power constraints $\bar{\mathcal{P}}$ and the vector of the demands $\{r_{k,j}\}_{k,j}$ satisfy the hypothesis of Lemma 1, i.e, if there exists a $\mathcal{P}^* \leq \bar{\mathcal{P}}$ with $R_{avg,k,j}(\mathcal{P}^*) = r_{k,j}$ for all k, j , then \mathcal{P}^* is a

stationary point of the ODE (10) (as projection term z is zero inside the constraint set). It will be in the limit set if further we can show that it is an attractor. In that case, the algorithm spends most of its time in these attractors or in other words the *DDPMCN algorithm asymptotically meets the demands of all the users*.

C. Analysis of the specific systems

Most of the systems considered in this paper (for example, **C-D-ZF**) transmit at one of rates from a discrete set \mathbb{R} depending on the instantaneous CS and for these one need to explicitly prove the continuity of the average rates. This is achieved in the following:

Lemma 2: *Assume A.1 and A.2. Then, for all the systems considered in Table 2, assumptions A.3 and A.4 are satisfied, Theorem 1 applies and hence for these systems the DDPMCN trajectory (9) asymptotically spends most of its time in the limit set, \mathbb{L}^{ODE} . Further, the demand meeting NE set, \mathbb{L}^{sys} , is non empty and these form the stationary points of the ODE (10), whenever for all k, j the demands satisfy*

$$r_{k,j} \leq \sup_{\mathcal{P} \leq \bar{\mathcal{P}}} R_{avg,k,j}(\mathcal{P}).$$

Proof : One can easily see that all the systems considered satisfy the assumption A.3. The boundedness is direct for discrete rate systems and is also true for ideal rate systems as seen from the formulas. The ideal rates are point wise continuous and are bounded and hence by bounded convergence theorem satisfy the continuous assumption A.4. The same for the discrete rates is given by the Lemma 3 given in Appendix.

The continuity assumption A.4 now also holds for the rate adaptation system with errors, **L-RAE-NO**, whenever the statistics of the errors $\{E_{k,j}\}$ are independent of the power profile or when they are continuous in \mathcal{P} . Thus for all the systems considered in this paper (except for the **A-D-ZF** system) Theorem 1 applies and hence for these systems the trajectory asymptotically spends most of its time in the limit set.

For all the systems considered so far, the hypothesis of Lemma 1 is satisfied, i.e., \mathbb{L}^{NE} is non empty whenever the power constraints are sufficient to cater the demand rates. This fact is established by the continuity of the average rates w.r.t. the power profile, i.e., the establishment of the assumption A.4. To be precise Lemma 1 is satisfied, i.e., \mathbb{L}^{NE} is non empty whenever for all k, j

$$r_{k,j} \leq \sup_{\mathcal{P} \leq \bar{\mathcal{P}}} R_{avg,k,j}(\mathcal{P}). \quad \diamond$$

For further analysis, one needs to study the limit set of ODE (10). A limit set of a ODE usually contains limit cycles or attractors. Right now, we can only say that, every stationary point of ODE (10) is a demand meeting NE and any attractor of the ODE must be a stationary point. We will show via numerical examples in the next section that the algorithm indeed converges to a demand meeting NE for all the systems considered in this paper, except for the **A-D-ZF** system. For the asymptotic discrete rate system, **A-D-ZF** we could not even establish Theorem 1, however we notice in simulations that the algorithm converges to a limit cycle.

D. Extensions to DDPMCN

The DDPMCN algorithm works under the basic assumption that the BS always has sufficient data to transmit. But in reality, data often arrives in real time and hence there can be situations when the BS can transmit at a higher rate but does not have sufficient data. In this case we propose the following extension to DDPMCN:

$$b_{k,j}^{t+1} = b_{k,j}^t + B_{k,j}^{t+1} - \min\{d_{k,j}^t, b_{k,j}^t\} \quad \text{and} \quad p_{k,j}^{t+1} = \Pi_{\mathbb{A}_j} [p_{k,j}^t + \mu_t (\min\{d_{k,j}^t, b_{k,j}^t\} - r_{k,j})], \quad (11)$$

	Type	Intf BS	Tx Ant	Users
L1	Linear	2	16	8
L2	Linear	2	32	8
H1	Hexagon	6	16	8
H2	Hexagon	6	2	2

Table 3: Network configurations

S1	Asymptotic Ideal with ZF Precoder (A-I-ZF)
S2	Rate adaptation with local CSIT and ZF precoder (L-RA-ZF)
S3	Rate adaptation with local CSIT, ZF Precoder and with estimation errors (L-RAE-ZF)
S4	Rate adaptation without CSIT (N-RA-NO)
S5	Rate adaptation without CSIT and estimation errors (N-RAE-NO).
S6	Asymptotic Discrete with ZF Precoder (A-D-ZF)

Table 4: System configurations

where $b_{k,j}^t$ represents the remaining (accumulating) bytes of data to be transmitted by BS j to the user k at the beginning of time slot t and $B_{k,j}^t$ represents the fresh sample of data added to the corresponding buffer.

V. SIMULATION

We consider two types of cellular networks in our simulations (Table 3). The first one is a Hexagonal network, where users in each cell experience interference from BS transmissions of surrounding cells (typically assumed to be from the 1st tier of surrounding 6 cells (see for example figure (1)). The second one is a linear network, where users in each cell experience interference from BS transmissions of adjacent cells (typically two adjacent neighbors). The system configurations are summarized in Table 4. Each BS equipped with M transmit antennas is serving K users in its cell. In all the simulations we also compute the average rates via the following iteration:

$$\phi_{k,j}^{t+1} = \phi_{k,j}^t + \mu^t \left(d_{k,j}^t - \phi_{k,j}^t \right) \text{ for all } k, j.$$

This iteration is only a measurement procedure that is used for the purpose of calculating average rates of the system for the numerical examples considered. That it represents the average rate can be understood by noticing that $\phi_{k,j}^t$ is actually a weighted average of all the instantaneous rates $\{d_{k,j}^\tau; \tau \leq t\}$ up to time t . These average rates are used to illustrate that systems considered in these examples, asymptotically (as time progresses) satisfy the user's demands on average.

The power limit on each BS is set to 1 unit. Interference factor γ_j from each interfering BS, j , is set to 0.5. For the simulations considered here, we choose the demand rate vector (to lie within the capacity region and is common for all the base stations) as: $\mathbf{r} = [.066 \ .132 \ .198 \ .264 \ .330 \ .396 \ .462 \ .528]$.

In the first set of simulations, we consider the hexagonal network (H1). We study the convergence behavior of the algorithm for systems S1, S2 and S3. We look at rate and power convergence as a function of number of iterations. Figure 2 shows the average rate convergence, while, figure 3 shows the convergence of power profile to a demand meeting NE. Figure 5 shows the limit cycle behavior explained in section III-A for system S6. We observe that :

1) The algorithm converges to the demand meeting NE: we see in Figure 2 that for all the systems, the average rate achieved asymptotically converges towards the demand rates.

2) As discussed in the previous sections, we notice from figure 3, that the converged power profile (demand meeting NE) is system specific. S3 is a system with errors, the proposed algorithm still satisfies the demands

asymptotically, however, the converged power profile has higher power levels in comparison with the error free systems S2 and S1.

3) Asymptotic results could be used in systems with moderate number of antennas/users (like $M = 16$ and $K = 8$) to decide the current transmission rates even in the absence of complete CSIT (shown in [31]). We observe from figure 3, that even the demand satisfying power profiles of the asymptotic system is close to that of the complete CSIT system (power profile converges to nearly the same values for systems S1 and S2). Thus, one can also compute the demand satisfying power profile in the absence of CSIT (via our proposed algorithm) using the asymptotic rates. Thus, complete CSIT is not required for either determining the current transmission rates or for computing the demand meeting power profile. However local part of the CSIT will be required for precoding purposes. This suggests that *with moderate number of antennas/users and with local CSIT, one can achieve good performance which will be close to that of a system with complete CSIT*. Further, the demand meeting power profile can be estimated faster with an asymptotic system (as in this case, the proposed algorithm is no longer stochastic in nature but rather is a deterministic Newton's zero finding algorithm).

4) As the discrete levels increases, the power profile decrease and finally converges to that of ideal rate. This is tabulated in Table 5.

5) For system S6, we observe the Limit cycle.

In the second set of simulations, for a given demand rate vector, we compare the algorithm behavior for different network configurations L1, L2 and H1 with system S2. All other parameters including the demand vector remains the same. Figure 4 plots power profile to which the DDPMC algorithm converged for various network configurations, which basically represent system specific demand meeting NE for satisfying the same demands, but, in different scenarios. We observe that: 1) To satisfy a given demand rate, the base stations in L2 expend the least power, followed by L1 and then H1. This is inferred from figure 4 as shown.

2) L2 performs better than L1, as the number of transmit antennas are doubled and so is the diversity. The interference also increases, however, this has lesser effect because the interference factor $\gamma < 1$.

3) H1 is the worst among the three due to increased number of interfering BS.

In the final set of simulations, for network configuration H2, we consider the least informed (No CSIT) systems, the rate adaptation systems S4 and S5. We choose the common demand rate vector as $[0.1694 \ 0.1936]$. Figures 6 and 7 illustrate the average rate and power profile convergence. As CSIT (even local) is not available at the base stations, they cannot use any precoders. Thus, it is a totally interference dominated system and hence the achievable capacity region is small. But the DDPMC algorithm works even for this least informed system: it asymptotically satisfies the demand rates, albeit using a higher power profile. Further, the rate estimation errors in system S5 demand higher power levels in comparison with the error free rate adaptation system S4 to achieve the same demands.

Further, we observe that the convergence to the demand meeting NE is quicker in those systems where base stations have more information regarding their neighbors CS (see for eg. figure 3).

VI. CONCLUSIONS

Mobile broadband users demand certain rates depending on the end application and QoS requirements. The base station serving these users has to allocate power to satisfy user demand operating within its own total power budget. Intra-cell and inter-cell interference diminish the available rates in multicell networks. Neighboring base stations can co-operate to exchange some form of channel state information depending on backhaul capacity and processing power to alleviate interference and thus enhance achievable rates. Further, system specific components like modulation, coding, rate allocation, channel estimation and synchronization impacts the achievable rates and

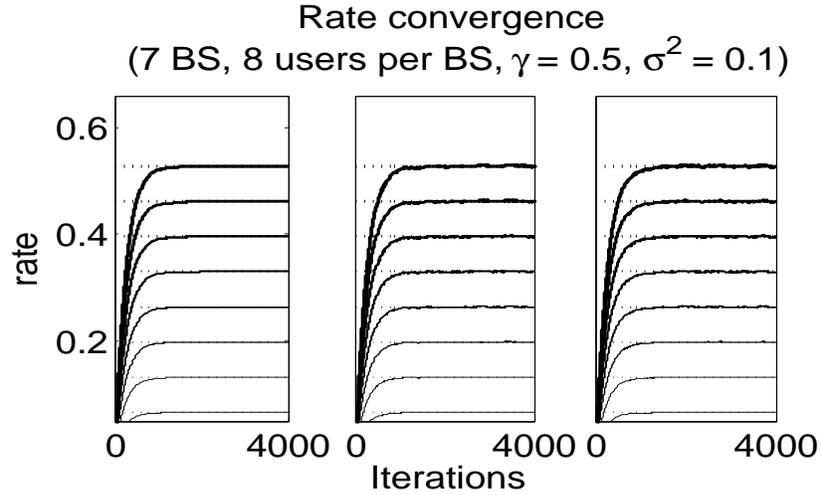


Fig. 2. Rate convergence (a) Asymptotic discrete with ZF Precoder (S1) (b) Rate adaptation with local CSIT and ZF Precoder (S2) (c) Rate adaptation with local CSIT, ZF Precoder and with estimation errors (S3)

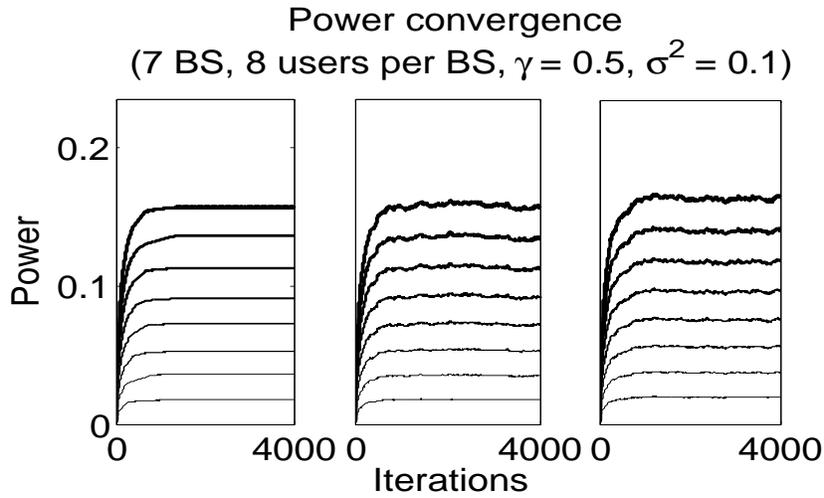


Fig. 3. Power convergence (a) Asymptotic discrete with ZF Precoder (S1) (b) Rate adaptation with local CSIT and ZF Precoder (S2) (c) Rate adaptation with local CSIT, ZF Precoder and with estimation errors (S3)

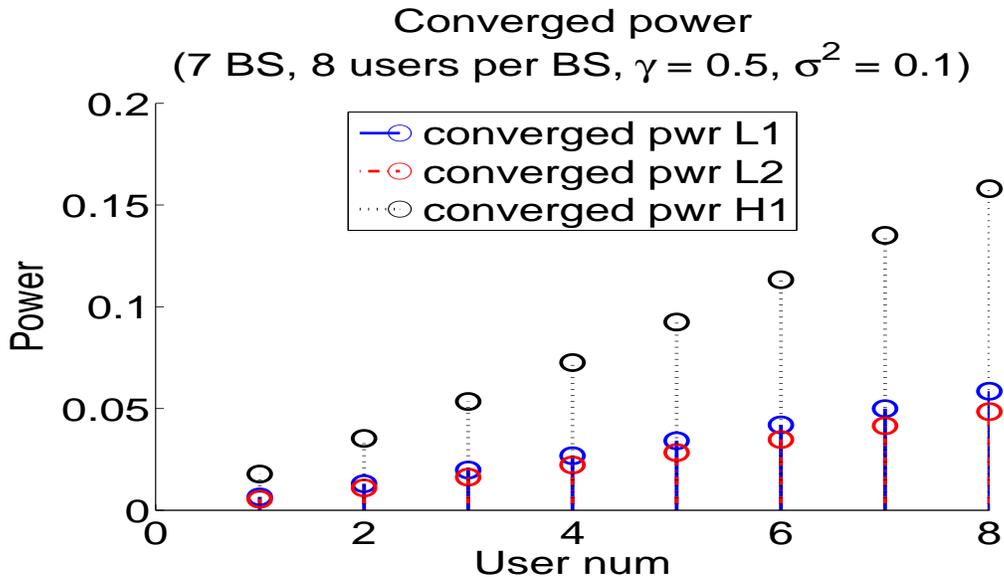


Fig. 4. Demand satisfying NE for Rate adaptation with local CSIT and ZF Precoder (S2) for linear L1, L2 and hexagonal H1 networks

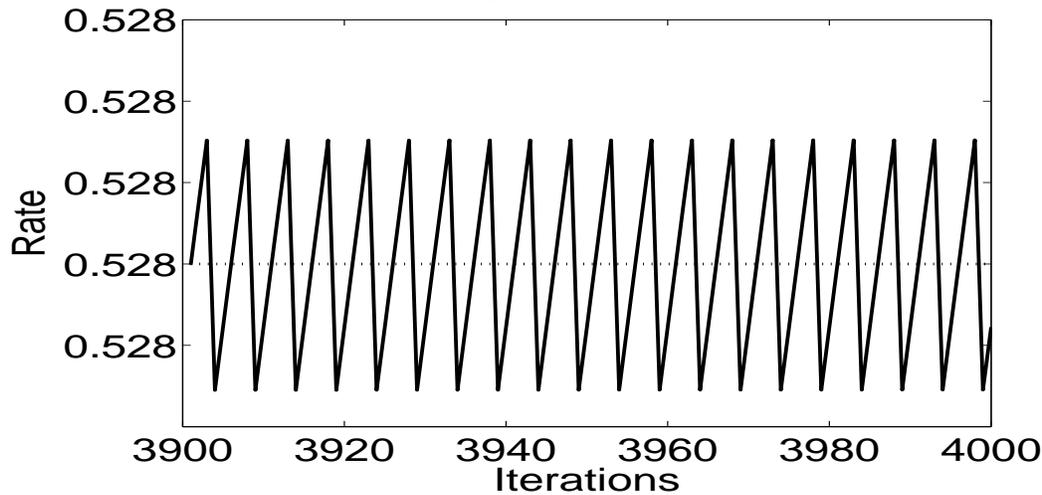


Fig. 5. Limit cycle behavior of the Asymptotic system (S6)

hence the power allocation. In our paper, we propose an universal power allocation algorithm which works in this setting. The stochastic approximation based decentralized power allocation algorithm runs at each BS, independently and simultaneously to meet the user demands as long as the demands are achievable. The power allocation is formulated as a game problem. A system specific capacity region is defined and the proposed algorithm is analyzed with an ODE framework. The proposed algorithm works well in a multitude of system configurations as demonstrated via simulations and analysis.

Our algorithm assumes that the serving BS always has sufficient amount of data to transmit. However, in many applications, the data is available in real time. We mentioned a possible extension of the same in the paper.

REFERENCES

- [1] H. J. Kushner and G. Yin, "Stochastic Approximation and Recursive Algorithms and Applications" second edition, Springer, 2003
- [2] G. J. Foschini, K. Karakayali, and R. A. Valenzuela, "Coordinating multiple antenna cellular networks to achieve enormous spectral efficiency", *IEE Proc. Commun.*, vol. 153, no. 4, pp. 548555, Aug. 2006
- [3] S. Shamai, O. Somekh, O. Simeone, A. Sanderovich, B. M. Zaidel, and H. V. Poor, "Cooperative Multi-Cell Networks: Impact of Limited-Capacity Backhaul and Inter-Users Links", in *Proc. Joint Workshop on Coding and Communications (JWCC07)*, Oct. 2007, pp. 1416.
- [4] D. Gesbert, S. Hanly, H. Huang, S. Shamai, O. Simeone, and W. Yu, "Multi-cell MIMO Cooperative Networks: A New Look at Interference", *IEEE J. Sel. Areas Commun.*, 2010, to appear. ul 2004.
- [5] "Beyond the base station router", Alcatel-Lucent technical note available at <http://innovationdays.alcatel-lucent.com/2008/documents/Beyond%20BSR.pdf>
- [6] Sivarama Venkatesan, Angel Lozano, Reinaldo Valenzuela, "Network MIMO: Overcoming Intercell Interference in Indoor Wireless Systems", *Asilomar 2007*
- [7] Sheng Jing, David N. C. Tse, Joseph B. Soriaga I, Jilei Hou, John E. Smeets and Roberto Padovani, "Downlink Macro-Diversity in Cellular Networks", *ISIT 2007*
- [8] Mung Chiang, Prashanth Hande, Tian Lan and Chee Wei Tan, "Power Control in Wireless Cellular Networks", now publishers, 2008
- [9] G. J. Foschini and Z. Miljanic, "A Simple Distributed Autonomous Power Control Algorithm and its Convergence", *IEEE Trans. on Vehicular Tech.* Vol. 42, No. 4, Nov 1993
- [10] C. Wu, D. P. Bertsekas, "Distributed Power Control Algorithms for Wireless Networks", *VTC*, Mar 2001
- [11] T. Holliday, A. Goldsmith, P. Glynn, and N. Bambos, "Distributed power and admission control for time varying wireless networks", *Globecom*, Nov 2004.
- [12] T. Javidi, "Decentralized rate assignments in a multi-sector CDMA network", *Globecom*, December 2003.
- [13] D. O'Neill, D. Julian, and S. Boyd, "Adaptive management of network resources", *IEEE VTC*, October 2003.
- [14] P. Hande, S. Rangan, and M. Chiang, "Distributed uplink power control for optimal SIR assignment in cellular data networks", *INFOCOM*, April 2006.
- [15] J.-W. Lee, R. R. Mazumdar, and N. B. Shroff, "Joint opportunistic power scheduling and rate control for wireless ad hoc networks", *IEEE Transactions on Vehicular Technology*, vol. 56, no. 2, pp. 801809, March 2007

System	Demand satisfying NE (converged power)								
ID	.015	.032	.049	.067	.085	.105	.126	.147	
D1	.017	.035	.053	.072	.092	.113	.135	.158	
D2	.031	.049	.071	.097	.121	.146	.174	.203	

Table 5: System S2

ID: Ideal, D1: Discrete - 100 levels, D2: Discrete - 20 levels

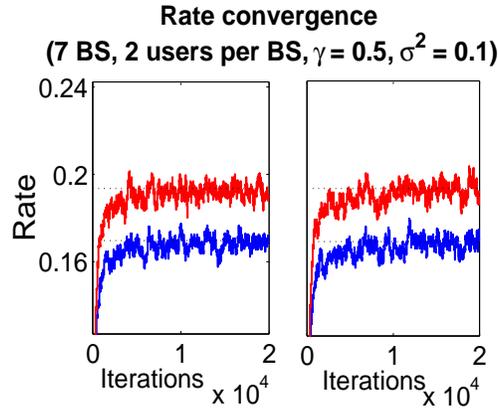


Fig. 6. Rate convergence (a) Rate adaptation without CSIT (S4) (b) Rate adaptation without CSIT and with estimation errors (S5)

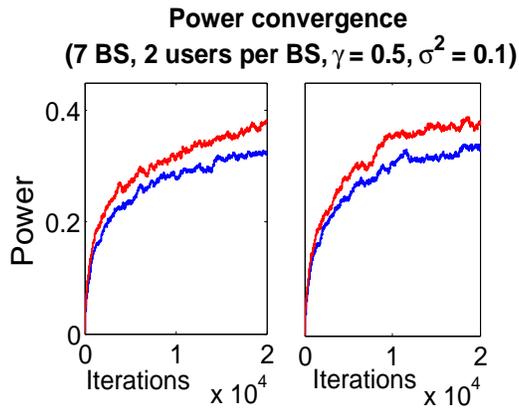


Fig. 7. Power convergence (a) Rate adaptation without CSIT (S4) (b) Rate adaptation without CSIT and with estimation errors (S5)

- [16] R. Agarwal, V. Abhishek, R. Vannithamby, and J. Cioffi, Opportunistic feedback protocol for minimizing power in uplink with heterogeneous traffic, *IEEE VTC*, Sep 2007
- [17] E. Altman and Z. Altman, S-modular games and power control in wireless networks, *IEEE Transactions on Automatic Control*, vol. 48, no. 5, pp. 839842, 2003.
- [18] A. B. MacKenzie and S. B. Wicker, Game theory in communications: Motivation, explanation, and application to power control, *Globecom*, Nov 2001.
- [19] S. Ulukus and R. Yates, Stochastic power control for cellular radio systems, *IEEE Transactions on Communications*, vol. 46, no. 6, pp. 784798, June 1998.
- [20] H. Viswanathan, S. Venkatesan, and H. Huang, Downlink capacity evaluation of cellular networks with known-interference cancellation, *IEEE Journal of Selected Areas in Communications*, vol. 21, no. 6, pp. 802811, June 2003.
- [21] S. Hanly, An algorithm for combined cell-site selection and power control to maximize cellular spread spectrum capacity, *IEEE Journal of Selected Areas in Communications*, vol. 13, no. 7, pp. 27962815, September 1995 [32, 33, 194, 196]
- [22] T. ElBatt and A. Ephremides, Joint scheduling and power control for wireless ad hoc networks, *IEEE Transactions on Wireless Communications*, vol. 3, no. 1, pp. 7485, January 2004.
- [23] R. Zakhour and D. Gesbert, Coordination on the MISO Interference Channel Using the Virtual SINR Framework, *International ITG Workshop on Smart Antennas*, 2009.
- [24] E. Bjornson, R. Zakhour, D. Gesbert, and B. Ottersten, Distributed Multicell and Multiantenna Precoding: Characterization and Performance Evaluation, *Globecom*, 2009
- [25] E. A. Jorswieck, E. G. Larsson, and D. Danev, Complete Characterization of the Pareto Boundary for the MISO Interference Channel”, *IEEE Trans. Sig Proc*, vol. 56, no. 10-2, pp. 52925296, 2008.
- [26] Federico Boccardi, Howard Huang ”Optimum power allocation for the MIMO-BC zero-forcing precoder with per-antenna power constraints”, 40th Annual Conference on Information Sciences and Systems, 2006
- [27] Zhu Han, Zhu Ji, and K. J. Ray Liu, ”Power Minimization for Multi-Cell OFDM Networks Using Distributed Non-cooperative Game Approach” *IEEE Globecom* 2004
- [28] H. Dahrouj and W. Yu, Coordinated beamforming for the multicell multi-antenna wireless system, 42nd Annual Conference on Information Sciences and Systems, 2008
- [29] Subhash Lakshminaryana, Jakob Hoydis, Merouane Debbah and Mohamad Assaad, ”Asymptotic Analysis of Distributed Multi-cell Beamforming”, *PIMRC* 2010
- [30] Aaron D. Wyner, ”Shannon-Theoretic Approach to a Gaussian Cellular Multiple-Access Channel”, *IEEE Trans. on Information Theory*, vol. 40, pp. 1713 - 1727, Nov. 1994
- [31] S. Ramanath, M. Debbah, E. Altman and V. Kumar, ”Asymptotic analysis of precoded small cell networks”, *InfoCom*, Mar 2010
- [32] D. L. Goeckel, Adaptive coding for time-varying channels using outdated fading estimates,” *IEEE Trans. Comm.*, vol. 47, pp. 844-855, June 1999.
- [33] K. Balachandran, S. R. Kadaba, and S. Nanda, Channel quality estimation and rate adaptation for cellular mobile radio,” *IEEE J. Select. Areas Comm.*, vol. 17, pp. 1244-1256, July 1999.
- [34] D. V. Djonin, A. K. Karmokar, and V. K. Bhargava, Joint rate and power adaptation for type-I hybrid ARQ systems over correlated fading channels under different buffer-cost constraints,” *IEEE Trans. Veh. Tech.*, vol. 57, pp. 421-435, Jan. 2008.

APPENDIX

Proof of 1: For every j , the constrain set \mathbb{A}_j satisfies the assumption (A3.2), page 107 of [1]. By assumption A.3 $\{Y_{k,j}^t; t\}$ is uniformly integrable and hence satisfies assumption A.2.1, pp. 258 [1]. They also satisfy the assumption A.2.3 to A.2.7 of pages 258, 259 [1] with $g_t = \bar{g} = g^p$ and with $\beta_t = 0 \xi_t = 0$ for all time t . Assumption A.2.2, pp 258, [1] is satisfied because of our Assumption A.4. Let \mathbf{z}_j represent the projection or constraint term, the minimum force needed to keep the vector \mathbf{p}_j in \mathbb{A}_j . Then by Theorem 2.3, pp. 259, [1] the DDPMCN algorithm trajectory \mathcal{P}^t converges weekly to the trajectory of the solution of the ODE (10) (in the sense as explained in [1]). Further by the same theorem of [1], for any $\delta > 0$, the fraction of time that the tail sequence $\{\mathcal{P}^\tau\}_{\tau \geq t}$, with initializations $p_{k,j}^t = \tilde{p}_{k,j}$ spends in the δ -neighborhood of the limit of set of the above ODE (10), $\mathbb{B}_\delta(\mathbb{L}^{ODE})$, goes to one (in probability) as $t \rightarrow \infty$. \diamond

Lemma 3: The average rates $R_{avg,k,j}^{sys}$ for systems **C-D-ZF** and **C-D-NO** are continuous w.r.t. power profile \mathcal{P} for all k, j .

Proof : Let $R_{k,j}(\mathcal{P}, \mathcal{H})$ represent the corresponding ideal rate (the rate before discretization) for the given CS \mathcal{H} . From all the rate formulas in this paper, we can see that these rates bounded and are continuous in \mathcal{P} , for all \mathcal{H} . For the discretized systems, the average rates can be written as,

$$R_{avg,k,j}(\mathcal{P}) = \sum_{i \leq N} q(i, \mathcal{P}) r_i \text{ where}$$

$$q(i, \mathcal{P}) := \int 1_{\{r_{i-1} \leq R_{k,j}(\mathcal{P}, \mathcal{H}) \leq r_i\}} d\Gamma(\mathcal{H}) \quad (12)$$

with $d\Gamma$ representing the Gaussian measure. For a given \mathcal{P} , the probability of the sets of the type (the boundaries of the sets used while defining the indicators in (12))

$$\Gamma(\{\mathcal{H} : R_{k,j}(\mathcal{P}, \mathcal{H}) = r_i\}) = 0,$$

because of the continuity of the Gaussian measure. Hence, the point wise functions of integral (12) are continuous w.r.t. to \mathcal{P} for almost all \mathcal{H} . Thus the lemma follows by bounded convergence theorem. \diamond