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

Sreenath Ramanath[1], Veeraruna Kavitha[1,2], Merouane Debbah[3]
[1]INRIA, Sophia-Antipolis, France; [2]Universite d'Avignon, Avignon, France;
[3]SUPELEC, 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 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 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 or other forms of BS co-operation enable sharing complete or statistical knowledge of channel states (CS) amongst neighbors via back-haul links to alleviate interference and offer better rates to users. When back-haul is not available, each BS may estimate the local channel state information and use the same for better performance. In some cases, a low rate feedback from the receiver indicating the QoS of the current transmissions is utilized, while in the worst case the transceivers are designed with no CS information. Thus we have a variety of systems with varying degrees of the information about the interfering channels. However the goal in each is the same: satisfy the demands of all the users. We may require higher power profiles to satisfy the same demands when working with lesser information. 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 or using some partial information. The following are the contributions of this paper:

1) A system specific game theoretic problem formulation using the system specific capacity region.

2) A *Stochastic Approximation* based universal power allocation algorithm in an interference limited multi-cell network.

3) Various properties (eg., convergence) of the proposed algorithm is analyzed using an ODE framework.

4) Simulation results demonstrate the effectiveness of the proposed algorithm for a variety of systems.

Related Work: For an excellent survey on power control in wireless networks, the reader is referred to [2] and the references there-in (eg. [3], [4], [5], [6], [7]). In recent years, several authors have addressed distributed power control strategies with various levels of co-operation for a given system configuration (eg. [8]). Typically, the design objective 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. But our algorithm does not optimize, it only meets the demands (in the form of average transmission rates) on average asymptotically[1]. This relaxation helps us in proposing an algorithm that requires minimal information (hence has minimal complexity) 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. Hence this algorithm does not require any extra information and this can be exploited in many more ways. For example, one can probably use this algorithm in networks with heterogeneous cells, i.e., when each cell has a system configuration that can be different from the other cells.

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. Appendix contains example systems and proofs.

Notations: Boldface lower-case symbols represent vectors, capital boldface symbols denote matrices ($\mathbf{I}_N$ is the $N \times N$ identity matrix). Hermitian transpose is denoted $(\cdot)^{\mathsf{H}}$ while $\mathbf{tr}[\mathbf{X}]$ 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 $\mathbf{E}_{\mathbf{s}}$ is expectation w.r.t to $\mathbf{s}$ when conditioned (if any) on the other random variables.

II. System model

We consider a multi-cell 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 $\mathbf{H}_{j,l}$ represent the $K \times M$ channel matrix, when the users in cell j receive signals from the BS of cell l 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 $\gamma_l \in [0,1]$ be the interference factor, representative of the level of interference from cell l. For example, as base stations become denser, interference increases and hence $\gamma_l \to 1$. The signal vector (of length K) received by users in cell j is given by,

$$\mathbf{y}_j = \mathbf{H}_{j,j}\mathbf{x}_j + \sum_{l=1,l\neq j}^{N} \gamma_l \mathbf{H}_{j,l}\mathbf{x}_l + \mathbf{n}_j \text{ for all } j \leq N. \tag{1}$$

In the above 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 $\mathbf{tr}(\mathbf{E}[\mathbf{x}_j\mathbf{x}_j^H]) \leq \bar{P}_j$ 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,

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

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,

$$\begin{aligned} y_{k,j} &= \mathbf{h}_{k,j,j}^{\mathsf{H}}\mathbf{x}_j + \sum_{i=1,i\neq k}^{K}\mathbf{h}_{i,j,j}^{\mathsf{H}}\mathbf{x}_j + \sum_{l=1,l\neq j}^{N}\sum_{i=1}^{K}\gamma_l\mathbf{h}_{i,j,l}^{\mathsf{H}}\mathbf{x}_l + n_{k,j} \\ &= u_{k,j} + i_{k,j,j} + \sum_{l\neq j} i_{k,j,l} + n_{k,j} \end{aligned} \tag{2}$$

where $\mathbf{h}_{k,j,l}$, the k^{th} row of matrix $\mathbf{H_{j,l}}$, represents the M length channel vector for user k of cell j as received from the BS of cell l. In the above, $u_{k,j}$, $i_{j,j,k}$ and $i_{k,j,l}$ 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 multi-cell 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 $\mathbf{x}_j = (\sqrt{\mathbf{p}_j}.\mathbf{s}_j)$. 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:

$$\begin{aligned} \mathbf{E}_{\mathbf{s}_j,1\leq j\leq N}[|u_{k,j}|^2] &= p_{k,j}|h_{k,j,j,k}|^2, \\ \mathbf{E}_{\mathbf{s}_j,1\leq j\leq N}[|i_{k,j,j}|^2] &= \sum_{\bar{k}\neq k} p_{\bar{k},j}|h_{k,j,j,\bar{k}}|^2 \text{ and} \\ \mathbf{E}_{\mathbf{s}_j,1\leq j\leq N}[|i_{k,j,l}|^2] &= \sum_{\bar{k}} \gamma_l p_{\bar{k},l}|h_{k,j,l,\bar{k}}|^2 \end{aligned} \tag{3}$$

[1]We show that the demand meeting power profile to be a NE of a 'leaky' game. We call this game 'leaky', because the utility of the game is upper bounded by the demands (see definition (5), section III-A). In summary our aim is to provide a channel, to each one of the users, whose (system specific) capacity is more than or equal to the user's demand.

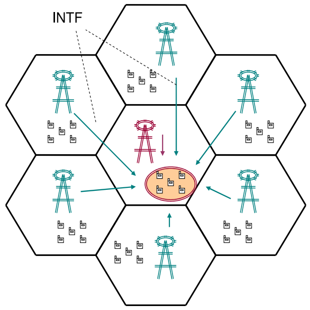

Fig. 1. 2D Wyner model

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

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

where, $h_{k,j,l,\bar{k}}$ is the $(k,\bar{k})^{\text{th}}$ component of the matrix $\mathbf{H_{j,l}}$. 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 its users demand rates represented by $\mathbf{r}_j := \{r_{k,j}, k \le K\}$. 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 universal power allocation algorithm which runs independently and simultaneously at all the base stations and tunes the power levels to achieve the user demands using minimal information. The power levels depend upon the system configuration (for example channel precoding scheme, rate allocation scheme). We consider some interesting example systems briefed in Table **2** and described in Appendix A. These systems are referred using a three part code, I-II-III, as explained below:

1) The first part (I) represents the **availability of channel state information at transmitter**[2]: a) **A** represents an asymptotic (large number of antennas/users) system, where achievable rates for all most all CS are approximated by a constant (see [10] and references there in), b) **C** for systems with 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 CSIT.

2) The second part (II) represents **the transmission rates used at the system**[3]: 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 $\mathbb{R} = \{r_1, r_2, \cdots r_{N_R}\}$ (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. [11]) without CSIT, d) **RAE** when there are estimation errors in the rate adaptation algorithm.

3) The third part (III) represents the **precoder**[4]: a) **ZF** for zero forcing precoding, b) **NO** for no channel precoding.

[2]One can also consider systems which have an estimate of the CS.

[3]We illustrate these concepts using simple rate allocation schemes. One can extend it to other rate allocations, for eg. schemes that incorporate fairness.

[4]One can also consider other types of precoders (eg. MMSE). Our analysis and proofs hold for these configurations as long as they satisfy the assumptions **A**.1 to 4 (refer Section IV).

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, $\mathcal{P} := \{p_{k,j}\}_{k \le K, \ \ j \le N}$, 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), $\mathcal{H} := \{\mathbf{H}_{1,1}, \mathbf{H}_{1,2}, \cdots, \mathbf{H}_{N,N}\}$, arranged as a matrix of dimension $KN \times MN$, represents the channel state of the entire system.

Rate for a given power profile and system, $R^{sys}_{k,j}(\mathcal{P}, \mathcal{H})$, represents the transmission rates allocated, to the user k by the base station j, in system represented by sys (eg. **N-RAE-NO** in Table **2**) when the base stations use powers $\mathcal{P}$ and when the CS is $\mathcal{H}$. These rates are given in the right column of the Table **2** for various example systems, whose detailed descriptions are provided in Appendix A.

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

Power constraint ($\mathcal{P} \le \bar{\mathcal{P}}$) We use $\le$ in a special manner to facilitate defining the power constraint. We say a power profile $\mathcal{P}$ is "less that 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 as: $\sum_k p_{k,j} \le \sum_k \bar{p}_{k,j}$ for all $j \le N$.

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.,

$$\mathbb{C}^{sys}(\bar{\mathcal{P}}) := \Big\{ \ \{R_{k,j}\} \in \mathcal{R}^{NK} : \text{ for all } k, j \\ R_{k,j} = R^{sys}_{avg,k,j}(\mathcal{P}) \text{ for some } \mathcal{P} \text{ with } \mathcal{P} \le \bar{\mathcal{P}} \Big\}. \quad (4)$$

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 always achieve the maximum possible rate and hence its capacity region shrinks. It further depends upon $\mathbb{R}$, the set of supported rates. 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, $\mathbf{p}_j := [p_{1,j}, \cdots, p_{K,j}]$. Note that $\mathcal{P} = [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_N]$. Define the utility of player j as[5],

$$U^{sys}_j(\mathbf{p}_j, \mathcal{P}_{-j}) \ := \ \sum_k \min\Big\{R^{sys}_{avg,k,j}(\mathbf{p}_j, \mathcal{P}_{-j}), \ \ r_{k,j}\Big\}$$
$$\text{with } \mathcal{P}_{-j} \ := \ [\mathbf{p}_1, \mathbf{p}_2, \cdots, \mathbf{p}_{j-1}, \mathbf{p}_{j+1}, \cdots, \mathbf{p}_N]. \quad (5)$$

In the above, $\mathcal{P}_{-j}$ is the power vector profile excluding only

[5]The utility of an user is the average rate at which its data is transferred. The user k of cell j requires transmission at maximum at its demand rate $r_{k,j}$ and hence his utility is upper bounded by the same.

System Description (more details in Appendix A)	$R_{k,j}^{sys}(\mathcal{P},\mathcal{H})$
A-I-ZF: Large no. of antennas and users with $M > K$. Asymptotic rates of [10] approximate the instantaneous rates for almost all CS and transmission at ideal rates. Zero forcing precoder.	$\log\left(1+\frac{p_{k,j}}{\frac{1}{\beta-1}\sum_{l=1,l\neq j}^{N}\gamma_l\ \frac{\mathrm{tr}(\mathbf{P}_l)}{K}+\sigma_{k,j}^2}\right)$
C-I-ZF : Number of antennae/users not large enough. Asymptotic results not accurate. Every BS has CSIT, computes theoretical rates and transmits at ideal rates. ZF precoder.	$\log\left(1+\frac{p_{k,j}^t}{\sum_{l=1,l\neq j}^{N}\gamma_l\ \frac{\mathrm{tr}\left(\mathbf{H}_{j,l}^t Q_l \mathbf{H}_{j,l}^{t\ \mathsf{H}}\right)}{K}+\sigma_{k,j}^2}\right)$
C-D-ZF: Similar to C-I-ZF, but TX rate allocation from discrete set $\mathbb{R}$.	$\inf_{r\in\mathbb{R}}\{r \le R_{k,j}^{\mathbf{C-I-ZF}}(\mathcal{P},\mathcal{H})\}$
C-I-NO: Similar to C-I-ZF, but without Precoder.	$\log\left(1+\frac{\mathbf{E}_{\mathbf{s}_j}[\lvert u_{k,j}\rvert^2]}{\sum_l \mathbf{E}_{\mathbf{s}_l}[\lvert i_{k,j,l}\rvert^2]+\sigma_{k,j}^2}\right)$
C-D-NO: Similar to C-I-NO, but TX rate allocation from discrete set $\mathbb{R}$.	$\inf_{r\in\mathbb{R}}\{r \le R_{k,j}^{\mathbf{C-I-NO}}(\mathcal{P},\mathcal{H})\}$
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.	$R_{k,j}^{\mathbf{C-D-NO}}(\mathcal{P},\mathcal{H})$
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.	$R_{k,j}^{\mathbf{C-D-ZF}}(\mathcal{P},\mathcal{H})$
N-RAE-NO: Similar to N-RA-NO, but with rate estimation errors, $E_{k,j}(r)$.	$R_{k,j}^{\mathbf{N-RA-NO}}(\mathcal{P},\mathcal{H}) - E_{k,j}(R_{k,j}^{\mathbf{N-RA-NO}}(\mathcal{P},\mathcal{H}))$
L-RAE-ZF: Similar to L-RA-ZF, but with rate estimation errors, $E_{k,j}(r)$.	$R_{k,j}^{\mathbf{L-RA-ZF}}(\mathcal{P},\mathcal{H}) - E_{k,j}(R_{k,j}^{\mathbf{L-RA-ZF}}(\mathcal{P},\mathcal{H}))$

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

the powers of BS of cell j and $r_{k,j}$ is the demand of user k of cell j. Every system with given power constraint $\bar{\mathcal{P}}$ and demand vectors $\{r_{k,j}\}$ defines an N-player non cooperative strategic form game: $([1,2\cdots N],\quad \{U_j^{sys}\}_{j\le N})$. The Nash equilibrium (NE) of this game is a power profile $\mathcal{P}^*$ that satisfies,

$$\mathbf{p}_j^* \in \arg\max_{\mathcal{P}\le\bar{\mathcal{P}}} U_j^{sys}(\mathbf{p}_j,\mathcal{P}_{-j}^*) \text{ for all } j. \tag{6}$$

From 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 $\mathcal{P}^*\le\bar{\mathcal{P}}$, which is a NE satisfying all the demands:*

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

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. Lemma 2 of the next section gives an easily verifiable assumption which ensures this hypothesis of Lemma 1.

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.

We now present the Universal Power Allocation algorithm for power constrained Multi Cell Networks (UPAMCN).

IV. Universal Algorithm : UPAMCN

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.

Basic idea[6] **:** 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 some 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. UPAMCN algorithm

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

$$p_{k,j}^{t+1} = \Pi_{\mathbb{A}_j}\left[p_{k,j}^t - \mu^t\left(d_{k,j}^t - r_{k,j}\right)\right] \text{ with}$$
$$\mathbb{A}_j := \left\{\mathbf{p}\in\mathcal{R}^K : \sum_k p_k \le \bar{P}_j\right\};\quad \mathbb{A} := \mathbb{A}_1\times\mathbb{A}_2\cdots\times\mathbb{A}_N. \tag{7}$$

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:

[6]Most 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.*

	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**).

Table 4: System configurations

A.1 There exists a sequence $\alpha_t \to \infty$ with $\lim_t \sup_{0 \leq i \leq \alpha_t} \mu^{t+i}/\mu^t = 0$.

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 are bounded by the same constant, i.e., $|R^{sys}_{k,j}(\mathcal{P}, \mathcal{H})| \leq B$ for all k, j, $\mathcal{P}$ and $\mathcal{H}$.

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

We will show that the UPAMCN trajectory (7) can be approximated by the solution $(\mathcal{P}(t))$ of the following ODE (to be precise a differential inclusion).

$$\dot{p}_{k,j} \quad = \quad r_{k,j} - R_{avg,k,j}(\mathcal{P}) + z_{k,j}(\mathcal{P}) \text{ for all } k, j \quad (8)$$

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

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

The δ-neighborhood of this set is defined as:

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

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.

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 $\tilde{\mathcal{P}} < \bar{\mathcal{P}}$)*

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

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

Proof: Refer Appendix B.

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 (proof in Appendix B).

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 the UPAMCN trajectory (7) 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 (8), whenever for all k, j the demands satisfy

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

For further analysis, one needs to study the limit set of ODE (8). A limit set of a ODE usually contains limit cycles or attractors. The demand meeting NE of $\mathbb{L}^{sys}$ would be in the limit set if further we could show that they are attractors. In that case, the algorithm spends most of its time in these attractors or in other words the *UPAMCN algorithm asymptotically meets the demands of all the users.* Right now, we can only say that, every stationary point of ODE (8) 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.

D. Extensions to UPAMCN

The UPAMCN 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 UPAMCN:

$$\begin{aligned} b^{t+1}_{k,j} &= b^t_{k,j} + B^{t+1}_{k,j} - \min\left\{d^t_{k,j}, b^t_{k,j}\right\} \text{ and} \\ p^{t+1}_{k,j} &= \Pi_{\mathbb{A}_j}\left[p^t_{k,j} + \mu_t\left(\min\{d^t_{k,j}, b^t_{k,j}\} - r_{k,j}\right)\right] \quad (9) \end{aligned}$$

where $b^t_{k,j}$ represents the remaining (accumulating) bytes of data to be transmitted by BS j to the user k at the begining of time slot t and $B^t_{k,j}$ 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 1^{st} 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: $\phi^{t+1}_{k,j} = \phi^t_{k,j} + \mu^t\left(d^t_{k,j} - \phi^t_{k,j}\right)$ 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 $\phi^t_{k,j}$ is actually a weighted average of all the instantaneous rates $\{d^\tau_{k,j}; \tau \leq t\}$ 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_l from each interfering BS, l, is set

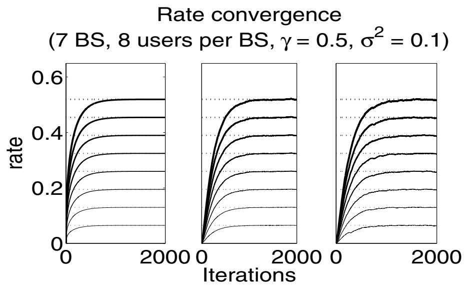

Fig. 2. Rate convergence for Systems S1, S2 and S3 (H1 Network)

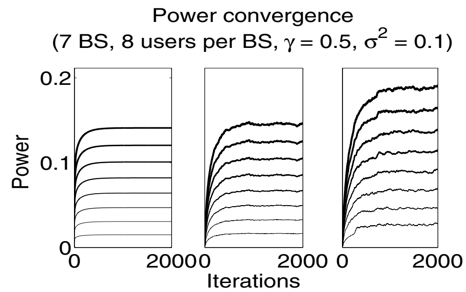

Fig. 3. Power convergence for Systems S1, S2 and S3 (H1 Network)

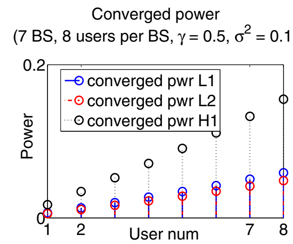

Fig. 4. Demand satisfying NE. System S2. (L1, L2 & H1 networks)

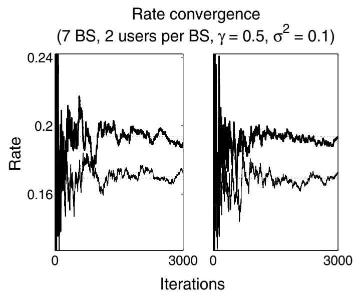

Fig. 5. Rate convergence for System S4 and S5 (H2 Network)

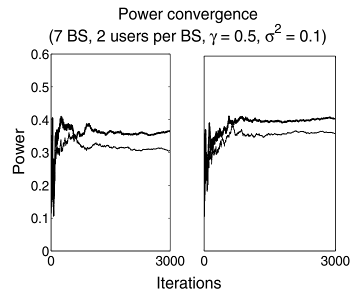

Fig. 6. Power convergence for Systems S4 and S5 (H2 Network)

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

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: $\mathbf{r} = [.065\ \ .130\ \ .195\ \ .260\ \ .325\ \ .389\ \ .454\ \ .520]$.

In the first set of simulations, we consider the hexagonal network (H1). The rate and the power convergence behavior of the algorithm for systems S1, S2 and S3 is plotted in figure 2 and 3, respectively. 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)** Note that S2 can also represents **C-D-ZF**, a complete CSIT system (see details on Table **2** and Appendix A). From figure 3, we observe that the converged power profile of **C-D-ZF** (S2) is close to that of **A-I-ZF** (S1) system. Thus the demand meeting power profile of systems with large number of transmit antennas and or users and large number of discrete levels in $\mathbb{R}$ is close to that of the asymptotic ideal rate system. Further convergence is faster with S1 system. Thus for such systems, UPAMCN algorithm can be used to estimate (approximately) the demand meeting power profile, much faster, using the asymptotic rate expressions in place of instantaneous transmit rates allocated, $\{d_{k,j}^t\}$. Note that this further avoids the need of complete CSIT, as we need only local CSIT for precoding. **(4)** As the discrete levels increase, the power profile decreases and finally converges to that of the ideal rate. This is tabulated in Table **5**.

In the second set of simulations, for given demand rates, we compare the algorithm behavior for different network configurations L1, L2 and H1 with system S2. We observe that: to satisfy the same demands, the base stations in L2 expend the least power, followed by L1 and then H1. L2 performs better than L1 due to improved transmit diversity. H1 is the worst (larger number of interfering base stations).

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 IV-D and IV-D 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 UPAMCN 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 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 (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 demands 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 hence the power allocation. In our paper, we propose an universal power allocation algorithm which works in this setting. The stochastic approximation based universal 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.

ACKNOWLEDGMENT

This research is carried out in the framework of the INRIA and Alcatel-Lucent Bell Labs Joint Research Lab on Self Organized Networks and the Alcatel-Lucent chair on Flexible Radio. The work of V. Kavitha is supported by CEFIPRA.

REFERENCES

[1] H. J. Kushner and G. Yin, "Stochastic Approximation and Recursive Algorithms and Applications", Second edition, Springer, 2003.
[2] M. Chiang, P. Hande, T. Lan and C. W. Tan, "Power Control in Wireless Cellular Networks", now publishers, 2008.
[3] 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.
[4] C. Wu, D. P. Bertsekas, Distributed Power Control Algorithms for Wireless Networks, IEEE VTC, Mar 2001.
[5] T. Holliday, A. Goldsmith, P. Glynn, and N. Bambos, Distributed power and admission control for time varying wireless networks, Globecom, Nov 2004.
[6] S. Ulukus and R. Yates, Stochastic power control for cellular radio systems, IEEE Trans. on Comm., vol. 46, no. 6, pp. 784-798, Jun 1998.
[7] T. ElBatt and A. Ephremides, Joint scheduling and power control for wireless ad hoc networks, IEEE Trans. on Wireless Comm., vol. 3, no. 1, pp. 74-85, Jan 2004.
[8] R. Zakhour and D. Gesbert, Coordination on the MISO Interference Channel Using the Virtual SINR Framework, International ITG Workshop on Smart Antennas, 2009.
[9] A. D. Wyner, "Shannon-Theoretic Approach to a Gaussian Cellular Multiple-Access Channel", IEEE Trans. on Information Theory, vol. 40, pp. 1713-1727, Nov 1994.
[10] S. Ramanath, M. Debbah, E. Altman and V. Kumar, "Asymptotic analysis of precoded small cell networks", Infocom, Mar 2010.
[11] K. Balachandran, S. R. Kadaba, and S. Nanda, Channel quality estimation and rate adaptation for cellular mobile radio," IEEE JSAC, vol. 17, pp. 1244-1256, July 1999.

APPENDIX A: EXAMPLE SYSTEMS

1) *Asymptotic Ideal Rate system :* 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 [10] 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 (10). Further, we consider a system in which, the base stations 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 ([10]) (when $M > K$):

$$R_{k,j}^{\mathbf{A-I-ZF}}(\mathcal{P},\mathcal{H}) \approx \log\left(1+ \frac{p_{k,j}}{\frac{1}{\beta-1}\sum_{l=1,l\neq j}^{N}\gamma_l\,\frac{\mathrm{tr}(\mathbf{P}_l)}{K} + \sigma_{k,j}^2}\right) \quad (10)$$

where, $\beta = M/K$ is the ratio of number of transmit antennas on the BS to the number of users and $\gamma_l \in (0,1)$ represents the interference from cell l. This rate is same for almost all CS $\mathcal{H}$ as it is an asymptotic rate. Similar expression is available for the case with $M = K$ in [10].

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:

$$R_{k,j}^{\mathbf{C-I-ZF}}(\mathcal{P},\mathcal{H}) = \log\left(1+ \frac{p_{k,j}^t}{\sum_{l=1,l\neq j}^{N}\gamma_l\,\frac{\mathrm{tr}\left(\mathbf{H}_{j,l}^t Q_l {\mathbf{H}_{j,l}^t}^{\mathsf{H}}\right)}{K} + \sigma_{k,j}^2}\right)$$
$$Q_l := {\mathbf{H}_l^t}^{\mathsf{H}}(\mathbf{H}_l^t{\mathbf{H}_l^t}^{\mathsf{H}})^{-1}\mathbf{P}_l^t(\mathbf{H}_l^t{\mathbf{H}_l^t}^{\mathsf{H}})^{-1}\mathbf{H}_l^t$$

For the same configuration, but without transmitter precoding, the the instantaneous transmission rate (as obtained using Shannon's capacity expression), from equation (2) is:

$$\begin{aligned} R_{k,j}^{\mathbf{C-I-NO}}(\mathcal{P},\mathcal{H}) &= \log(1+\eta_{k,j}) \text{ where} \quad (11)\\ \text{SINR, } \eta_{k,j} &:= \frac{\mathbf{E}[|u_{k,j}|^2]}{\rho_{k,j}} \text{ with}\\ \text{noise + interference, } \rho_{k,j} &:= \sum_l \mathbf{E}[|i_{k,j,l}|^2] + \sigma_{k,j}^2. \end{aligned}$$

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:

$$R_{k,j}^{\mathbf{C-D-ZF}}(\mathcal{P},\mathcal{H}) = \inf_{r\in\mathbb{R}}\{r \leq R_{k,j}^{\mathbf{C-I-ZF}}(\mathcal{P},\mathcal{H})\}, \quad (12)$$
$$R_{k,j}^{\mathbf{C-D-NO}}(\mathcal{P},\mathcal{H}) = \inf_{r\in\mathbb{R}}\{r \leq R_{k,j}^{\mathbf{C-I-NO}}(\mathcal{P},\mathcal{H})\}. \quad (13)$$

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. [11]). The UPAMCN algorithm is a very general algorithm and works with all those systems which satisfy assumptions **A.**1-4. These are quite simple assumptions and most of the systems can 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 actual 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,

$$R_{k,j}^{\mathbf{N-RA-NO}}(\mathcal{P},\mathcal{H}) = R_{k,j}^{\mathbf{C-D-NO}}(\mathcal{P},\mathcal{H}) \tag{14}$$

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**:

$$R^{\mathbf{L-RA-ZF}}(\mathcal{P},\mathcal{H}) = R_{k,j}^{\mathbf{C-D-ZF}}(\mathcal{P},\mathcal{H}) \tag{15}$$

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

$$\begin{aligned} & R_{k,j}^{\mathbf{N-RAE-NO}}(\mathcal{P},\mathcal{H}) \\ & \quad = R_{k,j}^{\mathbf{N-RA-NO}}(\mathcal{P},\mathcal{H}) - E_{k,j}(R_{k,j}^{\mathbf{N-RA-NO}}(\mathcal{P},\mathcal{H})) \qquad (16) \\ & R_{k,j}^{\mathbf{L-RAE-ZF}}(\mathcal{P},\mathcal{H}) \\ & \quad = R_{k,j}^{\mathbf{L-RA-ZF}}(\mathcal{P},\mathcal{H}) - E_{k,j}(R_{k,j}^{\mathbf{L-RA-ZF}}(\mathcal{P},\mathcal{H})) \qquad (17) \end{aligned}$$

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

Appendix B: Proofs

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

$$Y_{k,j}^t := r_{k,j} - d_{k,j}^t, \qquad p_{k,j}^{t+1} = \Pi_{\mathbb{A}_j}\left[p_{k,j}^t + \mu^t Y_{k,j}^t\right].$$

Define, $\mathcal{F}_t := \sigma\left(\mathcal{P}^\tau, \{Y_{k,j}^{\tau-1}\}_{k,j}, \text{ for all } \tau \le t\right)$ and let $\mathbb{E}_t$ represent the expectation w.r.t. $\mathcal{F}_t$, the filtration. Under the assumptions **A**.2 and **A**.3 clearly the condition expectation

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

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 UPAMCN algorithm trajectory $\mathcal{P}^t$ converges weekly to the trajectory of the solution of the ODE (8) (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 $\{\mathcal{P}^\tau\}_{\tau \ge t}$, with initializations $p_{k,j}^t = \tilde{p}_{k,j}$ for every (k,j), spends in the δ-neighborhood of the limit of set of the above ODE (8), $\mathbb{B}_\delta(\mathbb{L}^{ODE})$, goes to one (in probability) as $t \to \infty$. $\diamond$

Proof of Lemma 2: The boundedness assumption **A**.3 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 Lemma 3.
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 Theorem 1 applies.*
Conditions for existence of demand meeting NE : 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 to 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 $r_{k,j} \le \sup_{\mathcal{P} \le \bar{\mathcal{P}}} R_{avg,k,j}(\mathcal{P})$. $\diamond$

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,

$$\begin{aligned} R_{avg,k,j}(\mathcal{P}) &= \sum_{i \le N} q(i,\mathcal{P}) r_i \text{ where} \\ q(i,\mathcal{P}) &:= \int 1_{\{r_{i-1} \le R_{k,j}(\mathcal{P},\mathcal{H}) \le r_i\}} d\Gamma(\mathcal{H}) \end{aligned} \tag{18}$$

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 (18))

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

because of the continuity of the Gaussian measure. Hence, the point wise functions of integral (18) are continuous w.r.t. to $\mathcal{P}$ for almost all $\mathcal{H}$. Thus the lemma follows by bounded convergence theorem. $\diamond$