Distributed Discrete Power Control For Bursty Transmissions over Wireless Data Networks

Yiping Xing and R. Chandramouli
Multimedia System, Networking, and Communications (MSyNC) Laboratory
Department of Electrical and Computer Engineering
Stevens Institute of Technology, Hoboken, NJ 07030
Email:{yxing , mouli}@stevens.edu

***Abstract*— Distributed power control is an important issue in wireless networks. Recently, due to the bursty nature of data communication, packet switching is used in cellular systems. In addition, majority of previous power control algorithm assume that the transmitter power level can take values in a continuous domain. However, recent trends such as the GSM standard and QUALCOMM's proposal to the IS-95 standard use a finite number of discretized power levels. These motivate the need to investigate solutions for distributed discrete power control for bursty transmission. We first note that, by simply discretizing the previously proposed continuous power adaptation techniques will not suffice. This is because, a simple discretization does not guarantee convergence and uniqueness. On the other hand, the conventional analytical model based on mean values may be too optimistic and the analysis assuming that the data subscribers are always transmitting may be too pessimistic to evaluate the system performance. Therefore, we propose a probabilistic power adaptation algorithm and analyze its theoretical properties along with the numerical behavior for bursty transmission. We approximate the discrete power control iterations by an equivalent ordinary differential equation (ODE) to prove that the proposed stochastic learning power control algorithm converges to a stable Nash equilibrium. Conditions when more than one stable Nash equilibrium may exist are also studied. Experimental results are presented for several cases and the impact of data burstiness on the proposed algorithm is also concerned.**

I. Introduction

Due to the increased demand for wireless services, efficient use of available resources is important. Power control which mitigates unnecessary interference and saves the battery life of mobile users is a useful technique. Naturally, on one hand mobile users prefer to transmit at a lower power for a fixed signal to interference ratios (SIR) and on the other hand, for a given transmitter power, they prefer to obtain a better SIR. This observation motivates a reformulation of the power control problem using concepts from microeconomics and game theory. The framework was originally proposed in [1] for voice traffic, where a utility function is defined for each mobile user. The utility function reflects the user's preference regarding the SIR and the transmitter power. This concept is extended for power control in wireless data networks in [2] and [3], where throughput per battery life is chosen as the utility function, which seems to be a practical metric.

In most of the previous studies, the transmitter power level can assume any continuous value in a domain. However, in digital cellular system or future PCS systems, the power level is quantized into discrete values [4]. Therefore it is not clear how to apply those power control algorithms into a practical power-quantized system. Some previous work in this direction can be found in [5] and [6]. Discrete power control algorithms were developed based on conventional continuous framework. *However, it is shown in [6] that by simply "discretizing" the continuous power control algorithm, the convergence and uniqueness of the continuous power control are lost.* Hence, discrete transmission power control may need separate analysis on *convergence* and *uniqueness* issues. Without loss of generality, in this paper we focus on one of the game theory framework for wireless data networks proposed in [3]. In order to alleviate some of the problems in using previews, noncooperative power control strategies in the quantized-power levels set-up, we propose a discrete stochastic learning power control.

In traditional telephony systems, the calls have been circuit-switched, meaning that the call occupies one channel throughout the call. Due to the bursty nature of data communication this is inefficient, and instead packet switching can be used, where the data is divided into packets and transmitted independently. On the source level, data traffic is commonly modelled by an ON/OFF source model. Different distributions of ON/OFF-periods (e.g. exponential or heavy-tailed) have been adopted for data communications. Some pioneering work dealing with bursty data transmissions is reported in [7] and [8]. Because of the stochastic nature of our proposed algorithm, it can adapt well to the bursty data transmissions.

II. Problem Formulation

A. Utility Function and Noncooperative Power Control Game

We first define the utility function for the power control game. Let the power vector $\mathbf{y} = (y_1, ..., y_N) \in Y$ denote the selected power levels of all the users, where $y_j \in Y_j$ is the power level selected by user j, Y_j is the strategy set for user j and Y is the set of all power vectors. The resulting utility value for the j^{th} user is $u_j(\mathbf{y})$. To emphasize that the j^{th} user has control over its own power y_j only, we use an alternative notation $u_j(y_j, \mathbf{y}_{-j})$, where $\mathbf{y}_{-j}$ denotes the vector consisting of elements of $\mathbf{y}$ other than the j^{th} element. The utility of user j obtained by expending power y_j can be expressed as [3],

$$u_j(y_j, \mathbf{y}_{-j}) = \frac{LR}{My_j} f(\gamma_j) \frac{bits}{joule}, \quad (1)$$

where L is the information bits in frame (packets) of $M > L$ bits at a rate R b/s using y_j Watts of power. The efficiency function $f(\gamma)$ is defined as $f(\gamma) = (1 - 2P_e)^M$, where P_e is the bit error rate (BER). For example, if non-coherent FSK modulation scheme is used, then we have $P_e = 0.5e^{-\frac{\gamma}{2}}$, and γ_j (SIR of user j) is defined as,

$$\gamma_j = \frac{W}{R} \frac{h_j y_j}{\sum_{i \neq j} h_i y_i X_i + \sigma^2}, \quad (2)$$

and we assume that transmissions are slotted and mobiles have two modes of transmissions: on and off. let X_i be the activity indicator for mobile i, i.e., $X_i \in \{0, 1\}$ and $X_i = 1$ if and only if the mobile is active at a point in time. W is the available spread-spectrum bandwidth [Hz], σ^2 is the AWGN power at the receiver [Watts], and $\{h_j\}$ is the set of path gains from the mobile to the base station.

In the noncooperative power control game each user maximizes its own utility in a distributed fashion. Formally, the NPG is expressed as:

$$(\textbf{NPG}) \; max_{y_j \in Y_j} u_j(y_j, \mathbf{y}_{-j}) \; for \; all \; j \in N. \quad (3)$$

The solution of this NPG is given in the sense of the *Nash equilibrium* [10]. The existence and uniqueness of the NPG equilibrium has been shown in [3] for continuous power space for mobile users.

B. Discrete Stochastic Learning Power Control Game

In the stochastic learning game, the mobile users act as players or learning agents who participate in the power control game. The objective of each player is to maximize its expected payoff, which reflects the satisfaction of the players. And the payoff is measured in utility (e.g. Eq. (1)). The game is played repeatedly to learn the optimal strategies. Each individual automata (or mobile user) may not be aware of the number of mobile users participating in the game, the strategies available to the other users, and the responses for each possible play. The only information a player knows is its payoff after each play, based on which, the player learns the optimum strategy. The user's strategy is defined in probability. A strategy for player i is defined to be a probability vector $\mathbf{p}_i = [p_{i1}, ..., p_{im}]^t$, where player i chooses action j (or power level y_{ij}) with probability p_{ij}. Because each mobile user can only choose a power level from a finite discrete set, $y_{ij} \in Y_i$ should be a finite discrete set with dimension m_i. Then we can define the expected payoff for player i as g^i given by,

$$g^i(\mathbf{p}_1, ..., \mathbf{p}_N) = E[u_i | j^{th} \; player \; employs \; strategy \; \mathbf{p}_j, 1 \leq j \leq N] = \sum_{j_1, ..., j_N} d^i(j_1, ..., j_N) \prod_{s=1}^{N} p_{sj_s}, \quad (4)$$

where $d^i(j_1, ..., j_N) = E[u_i | player \; s \; chose \; action \; j_s, j_s \in Y_s, 1 \leq s \leq N]$.

Definition 1 The N-tuple of strategies $(\mathbf{p}_1^o, ..., \mathbf{p}_N^o)$ is said to be a *Nash equilibrium*, if for each i, $1 \leq i \leq N$, we have

$$\begin{aligned} & g^i(\mathbf{p}_1^o, ..., \mathbf{p}_{i-1}^o, \mathbf{p}_i^o, \mathbf{p}_{i+1}^o, ..., \mathbf{p}_N^o) \\ & \geq g^i(\mathbf{p}_1^o, ..., \mathbf{p}_{i-1}^o, \mathbf{p}_i, \mathbf{p}_{i+1}^o, ..., \mathbf{p}_N^o) \\ & \forall \; probability \; vector \; \mathbf{p}_i \in [0, 1]^{m_i}. \end{aligned} \quad (5)$$

In general, each $\mathbf{p}_i^o$ above will be a mixed strategy and we refer to $(\mathbf{p}_1^o, ..., \mathbf{p}_N^o)$ satisfying Eq. (5) as a Nash equilibrium in *mixed* strategies. With this definition, when there is no pure equilibrium as maybe the case in the discrete power control game, we can come to a mixed Nash equilibrium. It is well known that *every finite strategic-form game has a mixed strategy equilibrium* [10]. With this theory, there always exists a Nash equilibrium in our formulation. A Nash equilibrium is said to be in *pure* strategies if $(\mathbf{p}_1^o, ..., \mathbf{p}_N^o)$ is a Nash equilibrium with each $\mathbf{p}_i^o$ being a unit probability vector.

The proposed algorithm used by each of the user is as given below:

1) Set the initial probability vector $\mathbf{p}_i(0)$.
2) At every time step k, each active user chooses a power according to its action probability vector $\mathbf{p}_i$. Thus, the i^{th} player chooses action a_i at instant k, based on the probability distribution $\mathbf{p}_i(k)$.
3) Each active player obtains a payoff based on the set of all actions. The payoff to player i is $u_i(k)$, which is normalized.
4) Each active player (i) updates its action probability according to the rule:

$$\begin{aligned} p_{ij}(k+1) &= p_{ij}(k) - b u_i(k) p_{ij}(k) \quad a(k) \neq y_j, \\ p_{ij}(k+1) &= p_{ij}(k) + b u_i \sum_{s \neq j} p_{is}(k) \quad a(k) = y_j, \\ i &= 1, ..., N, \; j = 1, ..., m_i. \end{aligned} \quad (6)$$

   where $0 < b < 1$ is the step size, and u_i is normalized to lie in the interval $(0, 1)$.
5) If $\mathbf{p}_i$ converges, stop. Otherwise, go to step 2).

This update is known as linear reward-inaction (L_{R-I}) [9]. Let $P(k) = (\mathbf{p}_1(k), ..., \mathbf{p}_N(k))$ denote the state of the power strategies at instant k. Under this learning algorithm, $\{P(k), k \geq 1\}$ is a Markov process. L_{R-I} scheme is known to be ϵ-optimal ($\epsilon > 0$), i.e., upon convergence, the solution discovered by this scheme will produce a value for the objective function that is within ϵ of the optimal value.

We normalize the payoff $\{u_i\}$ as $u_i = \frac{\bar{u}_i - A}{B - A}$ $i = 1, 2, ..., N$, where $A = min_i\{\bar{u}_i\}$ and $B = max_i\{\bar{u}_i\}$. $\bar{u}_i$ is the utility of user i. In our system, A can just be set to 0 since $0 \leq \bar{u}_i, \forall i$, but it is not realistic to know B in advance. So what we do for the normalization is that we update the maximum value B dynamically. That is, first we initialize $B = 0$, then at instance k, if $\bar{u}_i(k) > B$, let $B = \bar{u}_i(k)$, otherwise keep B unchanged. We note that this normalization does not affect our theoretical results. This B is not the actual maximum value of $\bar{u}_i$, but is an overestimate value. But this is perhaps the best we can do without advance information. Alternatively, we could choose B to be a large constant. But as one can imagine, choosing

B too large compared to $\bar{u}_i$ will significantly decrease the convergence speed of the proposed algorithm since all u_i will be small values. Although we introduced this dynamic normalization in our algorithm, based on the above comments, we assume in our theoretical analysis that the normalization has already been carried out.

III. Theoretical Analysis Of Discrete Stochastic Learning Power Control

The proposed power control algorithm is an iterative stochastic process, so we need to characterize the long term behavior of this process. Our analysis resorts to an ordinary differential equation (ODE) whose solution approximates the asymptotic behavior of $P(k)$ if the step size parameter b used in Eq. (6) is sufficiently small. We can represent the learning algorithm given by Eq. (6) as $P(k+1) = P(k) + bG(P(k), a(k), u(k))$, where $a(k) = (a_1(k), ..., a_N(k))$, and $u(k) = (u_1(k), ..., u_N(k))$. Then we define a function f as $f(P) = E[G(P(k), a(k), u(k))|P(k) = P]$. With initial vector $P(0) = P_0$, the sequence $\{P(k)\}$ will converge weakly, as $b \to 0$, to P' which is the solution of the ODE [12]

$$\frac{dP}{dt} = f(P), P(0) = P_0. \tag{7}$$

First we analyze a relatively simple two-player two-power level data transmission stochastic power control game. The game is defined by the following pair of game matrices (general sum game):

$$\mathbf{R} = \begin{bmatrix} r_{11} & r_{12} \\ r_{21} & r_{22} \end{bmatrix}; \ \mathbf{C} = \begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix} \tag{8}$$

specifying the utility payoffs for the row player (user 1) and the column player (user 2) respectively. If the row player chooses power (action) $i \in \{1, 2\}$ and the column player chooses power $j \in \{1, 2\}$ the payoff to the row player is r_{ij} and the payoff to the column player is c_{ij}. In addition, when only user 1 is active the payoff is r_{i0}, and only user 2 is active the payoff is c_{0j}. Let $0 \leq \alpha \leq 1$ denote the probability of the row player picking action 1 and let $0 \leq \beta \leq 1$ denote the probability of the column player picking action 1. And let p_1^{on}, p_2^{on} denote the probability of user 1 and user 2 are active respectively. Then, we can express the differential equation in Eq. (7) whose solution characterizes the long term behavior of the proposed algorithm explicitly as:

$$\begin{aligned} f_{11} = \frac{dp_{11}}{dt} = \frac{d\alpha}{dt} &= p_1^{on}(\alpha(1-\alpha)E[u_1|P, a_1 = y_{11}] + \\ &\quad (1-\alpha)(-\alpha)E[u_1|P, a_1 = y_{12}]) \\ = p_1^{on}(\alpha(1-\alpha)&[p_2^{on}(\beta r_{11} + (1-\beta)r_{12}) + (1-p_2^{on})r_{10}] \\ +(1-\alpha)(-\alpha)&[p_2^{on}(\beta r_{21} + (1-\beta)r_{22}) + (1-p_2^{on})r_{20}]), \end{aligned} \tag{9}$$

$$\begin{aligned} f_{21} &= \frac{dp_{21}}{dt} = \frac{d\beta}{dt} \\ &= p_2^{on}(\beta(1-\beta)[p_1^{on}(\alpha c_{11} + (1-\alpha)c_{21}) + (1-p_1^{on})c_{01}] \\ &+(1-\beta)(-\beta)[p_1^{on}(\alpha c_{12} + (1-\alpha)c_{22}) + (1-p_1^{on})c_{02}]). \end{aligned} \tag{10}$$

By setting the left hand side of Eq. (9) and Eq. (10) to zero, we can solve for the *critical point* (e.g. (α^*, β^*)) of the system, and the corresponding constant solution $\alpha(t) \equiv \alpha^*, \beta(t) \equiv \beta^*$, which is called an equilibrium solution. We can easily see that *all pure strategies here are equilibrium solutions*, since at a pure strategy, either $\alpha = 0$ or $\alpha = 1$, which makes $f_{11} = 0$. Similarly, $\beta = 0$ or $\beta = 1$ is also an equilibrium solution from Eq. (10).

A critical point (x_o, y_o) is stable if any trajectory that begins near (within δ-neighborhood of) the point (x_o, y_o) remains near (within ϵ-neighborhood of) this point. If all the trajectories that start near a stable critical point actually approach it as $t \to \infty$, then the critical point is asymptotically stable. Thus if the ODE has an asymptotically stable stationary point (critical point) P^o, then for all initial conditions sufficiently close to it, the proposed algorithm essentially converges to P^o.

Remark 3.1 We know that [11] (a) all stationary points (critical points) that are not Nash equilibrium are unstable, and (b) all pure strategies that are strict Nash equilibria are asymptotically stable, which can be proved through *Lyapunov's stability theorem* [14].

Theorem 1: The proposed algorithm will never converge to a point which is not a Nash equilibrium.

Proof sketch: Suppose that the proposed algorithm eventually converges to a point ((i)which implies that point is stable), that is not a Nash equilibrium. The equilibrium solutions of the ODE (Eq. (9) and Eq. (10)) which characterize the long term behavior of the proposed algorithm are by definition stationary points. This implies that: (ii). the proposed algorithm will only converge to the stationary points. (i) and (ii) imply that stationary points that are not Nash equilibrium are stable. This contradicts *Remark 3.1 (a)*. Therefore, the proposed algorithm will never converge to a point which is not a Nash equilibrium.■

A sketch of the critical point set for a system, along with representative integral curves and their trajectories with arrows indicating the flow, is called a phase plane diagram. We use these diagrams to obtain qualitative information about the solution of the system, which is referred to as phase plane analysis [13]. As Eq. (9) and Eq. (10) are nonlinear differential equations, it makes it difficult to solve for the $(\alpha(t), \beta(t))$ trajectories explicitly. But we can obtain the trajectories using numerical techniques. As we show later, the phase plane diagram for $(\alpha(t), \beta(t))$ shows that, when there is no limit cycle, the proposed algorithm will eventually converge to one of the asymptotically stable points.

IV. Numerical Results

In our simulation experiments, we use a single cell CDMA system. The path gains are obtained using the simple path loss model $h_j = K/d_j^4$ where $K = 0.097$ is a constant. We also use $M = 80 bits$, $L = 64 bits$, $W = 1MHz$, $\delta^2 = 5{\times}10^{-15} Watts$, $\hat{p} = 1Watt$, and $R = 10^4 bits/sec$. The utility function given in Eq. (1) was used. First, we consider the proposed algorithm in a discrete time environment with two possible power levels for two mobile users. We use this

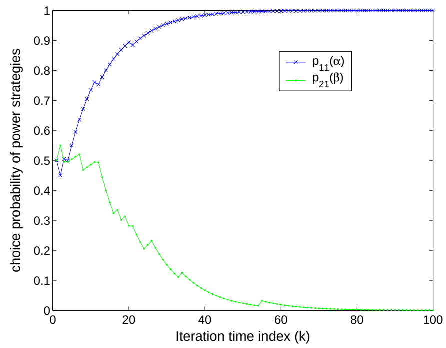

Fig. 1. Choice probability of the power strategies vs. number of iterations for two mobile users continuous transmission case with two power levels $\{0.01Watts, 0.1Watts\}$. Convergence of power strategies is seen to converge to the optimal values.

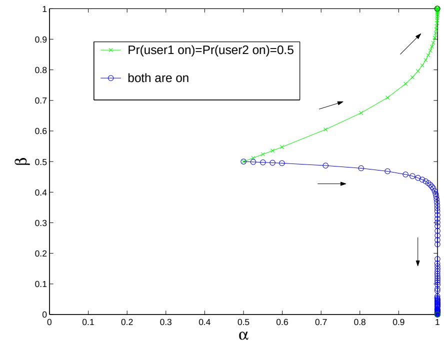

Fig. 2. The trajectory of ODE with $\alpha(0) = \beta(0) = 0.5$.

example to illustrate how the proposed algorithm works and how the derived ODE's (Eq. (9) and Eq. (10)) trajectories evolve. The two terminals are located at $\mathbf{d} = [300, 600]$ meters from the base station. Both users have two pure power strategies $Y = \{0.01Watts, 0.1Watts\}$. The probability of user 1 choosing power 0.01Watts at time slot k is $\alpha(k) = p_{11}(k)$, and choosing power 0.1Watts is $1 - \alpha(k) = p_{12}(k)$. In the same fashion, the probability of user 2 choosing power 0.01Watts is $\beta(k) = p_{21}(k)$, and choosing power 0.1Watts is $1 - \beta(k) = p_{22}(k)$. We set the initial probabilities for $\alpha(0)$ and $\beta(0)$ equal to 0.5, and the step size parameter $b = 0.1$. For continuous transmission, where $p_1^{on} = p_2^{on} = 1$, we observe that the proposed algorithm converges to the solution $y_1 = 0.01$ and $y_2 = 0.1$, both with probability 1, as depicted in Fig. 1, where $\alpha \to 1$ and $\beta \to 0$. And when for bursty transmission, where $p_1^{on} = p_2^{on} = 1$, the proposed algorithm converges to the solution $y_1 = 0.01$ and $y_2 = 0.01$, both with probability 1. By solving theoretically for the Nash equilibrium using game matrices (as given in Eq. (8)) for this power set, we see that there is only a unique pure Nash equilibrium,because $(r_{11} - r_{21})(r_{12} - r_{22}) > 0$. Thus, in this two-user two-power level case the Nash equilibrium in pure strategy is discovered by the proposed algorithm. The trajectory of the solutions of the ODE (Eq. (9) and Eq. (10)) is shown in Fig 2, which is obtained by numerical methods. The $r's$ and $c's$ in the ODE were calculated using the utility function given by Eq. (1). The arrows show the flowing direction of the trajectory. This trajectory of the ODE characterizes the long term behavior of the stochastic learning process. We can notice that the curve starts at the initialization point $(\alpha(0) = 0.5, \beta(0) = 0.5)$ and ultimately is attracted to the point $(\alpha = 1, \beta = 0)$ and point $(\alpha = 1, \beta = 1)$ respectively for continuous transmission $(p_1^{on} = p_2^{on} = 1)$ and bursty transmission $(p_1^{on} = p_2^{on} = 0.5)$. This is the behavior we observe in our simulation results also as seen in Fig. 1.

As noted previously, when the power levels are discrete there could be more than one equilibrium consisting of both pure and mixed strategies. For the two-user two-power level case, the game equilibrium can be classified into 3 categories. For multiple user $(N > 2)$, multiple power level case, the existence of multiple equilibrium will be pronounced. If the deterministic iterative algorithm proposed in [3] is used to solve the discrete power level NPG, then depending on the initialization of the power levels, the algorithm may converge to one of the pure equilibria or could oscillate between them without convergence. This is because that deterministic algorithms work on the objective function space. On the other hand, the proposed algorithm operates on the probability space and for an unbiased initialization in probability, converges to the best optimal equilibrium strategy. To illustrate this, we construct an example with $p_1^{on} = p_2^{on} = 1$ as follows. Let two terminals be located at $\mathbf{d} = [300, 600]$ meters from the base station. Both users have two pure power strategies $Y = \{0.003Watts, 0.975Watts\}$. let the probability of user 1 choosing power 0.003Watts be $\alpha(k) = p_{11}(k)$, and choosing power 0.975Watts be $1 - \alpha(k)$. In the same fashion, let the probability of user 2 choosing power 0.003Watts be $\beta(k) = p_{21}(k)$, and choosing power 0.975Watts be $1-\beta(k)$. The game matrices for this game are given below,

$$\mathbf{R} = \begin{bmatrix} 2.6667 & 0.0019 \\ 0.0082 & 0.0082 \end{bmatrix} \times 10^6, \mathbf{C} = \begin{bmatrix} 1.3018 & 0.8205 \\ 0.0000 & 0.0224 \end{bmatrix} \times 10^4.$$

Since $(r_{11} - r_{21})(r_{12} - r_{22}) < 0$, $(c_{11} - c_{12})(c_{21} - c_{22}) < 0$, and $(r_{11} - r_{21})(c_{11} - c_{12}) > 0$, there are 2 pure equilibria and 1 mixed equilibrium. As we can see, there are two pure Nash equilibria pair $(2.667 \times 10^6, 1.3018 \times 10^4)$ and $(0.0082 \times 10^6, 0.0224 \times 10^4)$ with power strategy pair $(0.003Watts, 0.003Watts)$ and $(0.975Watts, 0.975Watts)$ respectively. It is obvious that the first equilibrium is better than the second one. So we would like the power control algorithm to converge to the first one. Suppose we use the iterative power control algorithm (NPG) proposed in [3] with initial power strategy pair $(0.003Watts, 0.003Watts)$ which means initially player one chooses power level $0.003Watts$ and player two chooses power level $0.003Watts$, it will converge to the first equilibrium, but with strategy pair $(0.975Watts, 0.975Watts)$ it will converge to the second equilibrium. When with $(0.975Watts, 0.003Watts)$ or $(0.003Watts, 0.975Watts)$ it will even not converge but

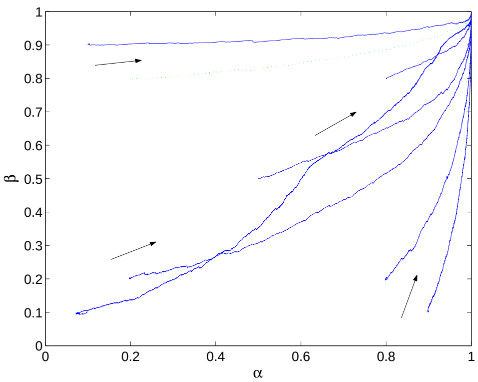

Fig. 3. Evolution of power strategies (probabilities) for different initial values of $\alpha(0)$ and $\beta(0)$. Convergence is seen to be robust to initializations.

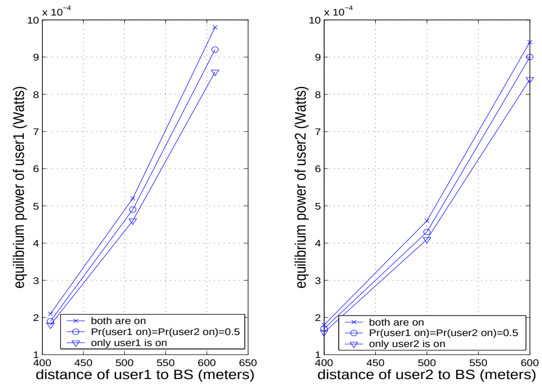

Fig. 4. Equilibrium transmitted power of user1 and user2

oscillate between two possible strategies. While using the proposed algorithm, except for the extreme case where the initial power strategy probabilities are biased very much towards the second equilibrium, it will converge to the first (better) equilibrium. This is illustrated in Fig. 3. More specifically, this figure shows the trajectories of $\alpha(t)$ and $\beta(t)$ for different initializations, $\alpha(0)$ and $\beta(0)$. Here, $\alpha(t)$ is the probability of user 1 choosing its first power strategy $(0.003Watts)$ and $\beta(t)$ is the corresponding value for the second user. As you can see all the sample paths of the proposed learning process converge to the optimal strategy $(\alpha = 1, \beta = 1)$ independent of their initial probability.

In the next experiment we evaluate the outcomes of games with bursty transmissions. Each user has 100 discrete power levels. The probability of transmissions of user1 and 2 in a slot is denoted by (Probability of user1 is on, Probability of user2 is on) and we test four cases: $(1, 1), (0.5, 0.5), (1, 0), (0, 1)$. The first case: $(1, 1)$ is the complete information game when both users continuously transmit data and is identical to previous experiment. The $(0.5, 0.5)$ case is the incomplete information game that both users may transmit data with a 50% chance. The last two cases are the cases when a single user transmits data. Additionally, we vary the distance of mobile users to base station. The results are present in Fig. 4. As one would expect, when a single user is transmitting, it has no interference from other users, it requires the lowest equilibrium power. When both users have a 50% chance of transmission, the equilibrium transmit power is in between the case where both users are always transmitting (e.g. $(1, 1)$) and the user is by itself (e.g. $(1, 0)$). Similar results were reported in [8], where the solution is calculated by centralized Newton's method which is impractical in implementation. In addition, for a more general case of 9 users, each with 50 power levels, our simulation shows the proposed algorithm will also converge to the optimal equilibrium.

V. Conclusion

A distributed discrete stochastic learning based power control algorithm for wireless data network is proposed. The proposed method evolves in probability space, thus may not be attracted by local maximum solutions. At each iteration, for an individual mobile user only one feedback payoff from the base station is used to update the power strategies, which significantly reduces the communication between base station and mobile users, and simplifies the control signaling. The proposed algorithm is shown to converge to the Nash equilibrium, robust to initial values and can handle bursty as well as continuous data transmission.

References

[1] Ji H, Huang CY, "Non-cooperative uplink power control in cellular radio systems," *Wireless Networks,* 1998; 4(2): 233-240.

[2] D. Famolari, N.B. Mandayam, D.J. Goodman and V. Shah, "A new framework for power control in wireless data networks: games, utitlity and pricing," *Proc. 36th Annual Allerton Conference on Communication, Control, and Computing.* Monticello, IL, pp. 289-310 1998

[3] C.U. Saraydar, N.B. Mandayam and D.J. Goodman, "Efficient power control via pricing in wireless data networks," *IEEE Trans. Commun.,* vol. 50, pp. 291-303. Feb. 2002

[4] "An overview of the application of Code Dvision Multiple Access (CDMA) to digital cellular systems and personal cellular networks," Document Number EX60-10010, QUALCOMM Inc., May 1992

[5] Chi Wan Sung, K. W. Shum, "Evaluation on the stability of discrete power control algorithms," *IEEE Personal, Indoor and Mobile Radio Communications.,* vol. 3, pp. 1097-1101. Sept. 2002

[6] M. Andersin, Z. Rosberg and J. Zander, "Distributed discrete power control in cellular PCS," *Wireless Personal Comm.,* vol. 6, pp. 211-231, 1998.

[7] Debasis Mitra and John A. Morrison, "A distributed power control algorithm for bursty transmissions on cellular, spread spectrum wireless networks," *Wireless Information Networks.,* J.M. Holtzman (editor), Kluwer, 1996, pp. 201-212.

[8] W. Teerapabkajorndet and P. Krishnamurthy, "A game theoretic model for power control in multi-rate mobile data networks.," *Proc. ICC.,* Alaska, 2003

[9] K.S. Narendra and M.A.L. Thathachar, "Learning Automata: An Introduction," Englewood Cliffs: Prentice Hall, 1989

[10] D. Fudenberg, J.Tirole, *Game Theory,* the MIT Press, Cambridge, 1992

[11] P.S Sastry, V.V. Phansalkar, and M.A.L. Thathachar, "Decentralized Learning of Nash Equilibria in Multi-Person Stochastic Games With Incomplete Information," *IEEE Trans Systems, Man, and Cybernetics.,* vol. 24. pp. 769-777, May 1994

[12] H. J. Kushner, *Approximation and Weak Convergence Methods for Random Process*, Cambridge, MA: MIT Press, 1984

[13] R. Kent Nagle, Edward B. Saff, *Fundamentals of Differential Equations and Boundary Value Problem*, Addison-Wesley Publishing Company, 1996

[14] K. S. Narendra, A. Annaswamy, "Stable Adaptive System," Englewood Cliffs: Prentice Hall, 1989