View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by Universidad Carlos III de Madrid e-Archivo

Universidad Carlos III de Madrid

e-Archivo
Institutional Repository

This document is published in:

Signal Processing Conference (EUSIPCO), 2014 Proceedings of the 22nd European (2014) pp. 401 - 405.

© 2014 IEEE. Personal use of this material is permitted. Permission from IEEE must be obtained for all other uses, in any current or future media, including reprinting/republishing this material for advertising or promotional purposes, creating new collective works, for resale or redistribution to servers or lists, or reuse of any copyrighted component of this work in other works.

A CONTROL THEORETIC APPROACH TO SOLVE A CONSTRAINED UPLINK POWER DYNAMIC GAME

Santiago Zazo[*], *Javier Zazo*[†]

Universidad Politécnica de Madrid
Signals, Systems & Radiocommunications Dept.
Madrid, Spain

Matilde Sánchez-Fernández[‡]

Universidad Carlos III de Madrid
Signal Theory & Communictions Dept.
Leganés, Spain

ABSTRACT

This paper addresses an uplink power control dynamic game where we assume that each user battery represents the system state that changes with time following a discrete-time version of a differential game. To overcome the complexity of the analysis of a dynamic game approach we focus on the concept of Dynamic Potential Games showing that the game can be solved as an equivalent Multivariate Optimum Control Problem. The solution of this problem is quite interesting because different users split the activity in time, avoiding higher interferences and providing a long term fairness.

1. INTRODUCTION

Dynamic Games and in particular Differential Games play a very important role in Game Theory (GT) in many applications as economic and management science [1]. As already stated, solutions are very complicated and just few simplistic examples are known to have closed expressions while in most of the cases, only approximate solutions by discretizing state / actions spaces or parameterizing value / utility functions are affordable. On the other hand, there are few publications on Electrical Engineering [2]-[4] that consider some dynamic effect on standard scenarios, and few tools to characterize the solutions, such as uniqueness of Nash equilibrium, or algorithms for finding them. In this work we consider an uplink scenario where a set of independent users need to define individually the power that they are going to allocate at each discrete time instant in order to maximize its achievable rate. It should be noted that if a subset of users decide to transmit at time t they are going to interfere to each other, thus significantly decreasing the achievable rate of every user. Also, each user has a limited battery for the transmission over time. This scenario describes a dynamic game where users try to transmit, the remaining battery represents the "state" of each user, and the state together with the achievable rate define the "utility" of each user. Finally the power that each user decides to allocate at each time t is the "action". The discrete time domain can be assumed in this scenario without loss of generality since the users are not going to change their power in a continuous way given that most communications system define time intervals where the transmitted power need to be fixed, for example time symbol or more generally frame duration. In our case, this allows us to get closer relationships between the problem formulation and the algorithmic implementation. Furthermore we consider an infinite (discounted) horizon problem because our scenario is not constrained in time, being the remaining energy of the battery the limitation that will finish the game.

Our approach to solve the game is based on reformulating the game as an equivalent Multivariate Optimum Control Problem (MOCP). This procedure, known as Dynamic Potential Games [5], follows the same spirit as Static Potential Games where the objective is to find an optimization control problem whose solution coincides with the Nash equilibria of the game. In many cases, it also provides information, under certain hypothesis, about the uniqueness of the solution. In dynamic scenarios, although the idea is similar, the verification of the conditions needed by the game is not a simple issue. The idea of Dynamic Potential Games has been very recently formulated in [5] although the basic principles are much older and originally developed by Dechert [6]-[9]. The rest of the manuscript is organized as follows. Sec. 2 introduces the dynamic game framework and presents the game as a MOCP. Sec. 3 formulates our uplink power game and provides the potential function to solve the game as a MOCP. Sec. 4, presents different approaches to solve the uplink power MOCP and finally Sec. 5 and 6 show the simulation results and the conclusions.

2. GAME AS A CONTROL THEORETIC PROBLEM

In a dynamic game we have a set of players $\mathcal{Q} = \{1 \dots Q\}$, where for each player i its utility function at discrete time t is given by $\pi^i(\mathbf{x}_t, \mathbf{u}_t)$, and it depends both on the system state $\mathbf{x}_t = (x_{1t} \dots x_{it} \dots, x_{Qt})$, with $x_{it} \in \mathcal{X}_i$, and on

[*] This work was supported in part by the program CONSOLIDER-INGENIO 2010 under the grant CSD2008-00010 COMONSENS

[†] This work was supported in part by the Spanish Ministry of Science and Innovation under an FPU doctoral grant.

[‡] This work has been partly funded with the project GRE3N (TEC 2011-29006-C03-01/02/03) and in the program CONSOLIDER-INGENIO 2010 under project COMONSENS (CSD 2008-00010).

the set of actions of all players, denoted in vector form as $\mathbf{u}_t = (u_{1t} \dots u_{it} \dots u_{Qt})$ with $u_{it} \in \mathcal{U}_i$. The discrete time Dynamic Game can be represented as:

$$\begin{aligned} & V^i(\mathbf{x}_0) = \max_{\{u_{it}\}} \sum_{t=0}^{\infty} \beta^t \pi^i(\mathbf{x}_t, \mathbf{u}_t) \qquad \forall i \\ & \text{s.t. } \mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t),\ g_i(\mathbf{x}_t, \mathbf{u}_t) \le 0 \end{aligned} \tag{1}$$

Each user i intends to find the optimum sequence of actions $\{u_{it}\}$ that maximizes its value function $V^i(\mathbf{x}_0)$ expressed in terms of its own current and future (discounted) utility function $\pi^i(\mathbf{x}_t, \mathbf{u}_t)$. Parameter $\beta < 1$ is the discount factor. Very importantly, there is one constraint related to the time evolution of the sequence of states $\{\mathbf{x}_t\}$ typically depending on the previous state and current actions (Markovian model). Also, some extra constraints are included $g_i(\mathbf{x}_t, \mathbf{u}_t) \le 0$ because in most of the applications, states and actions are constrained. Solving these problems requires finding the sequence of actions $\{\mathbf{u}_t^*\}$ that provide a Nash equilibrium. Using similar concepts as in static games:

$$\sum_{t=0}^{\infty} \beta^t \pi^i\left(\mathbf{x}_t, \mathbf{u}_{-it}^*, u_{it}^*\right) \ge \sum_{t=0}^{\infty} \beta^t \pi^i\left(\mathbf{x}_t, \mathbf{u}_{-it}^*, u_{it}\right) \forall i$$

where the equation must hold $\forall u_{it} \in \mathcal{U}_i$ and where $u_{it}^* \in \mathcal{U}_i$ represents the optimum action of user i at time t and $\mathbf{u}_{-it}^*$ is the same concept for all users except i . We will see next that in practice, this optimization procedure is very complicated because each user has to solve a constrained optimum control problem where several coupled differential (or difference) equations are involved. Typically, for open loop solutions, that is, $u_{it}^* = \vartheta(t)$ can be solved using the Maximum (or Pontryagin) Principle and for the closed loop (Feedback Markovian $u_{it}^* = \phi(\mathbf{x}(t))$) by solving the Euler Equation. We should note that $\vartheta(t)$ and $\phi(\mathbf{x}(t))$ are precisely the optimal trajectories (actions) to be determined.

We could solve the dynamic game in (1) by defining the Lagrangian for each agent and optimizing. Then for the i-th player, including the corresponding multipliers we have:

$$\begin{aligned} \mathcal{L}^i\left(\mathbf{x}_t, \mathbf{u}_t, p_t^i, \lambda_t^i\right) = \textstyle\sum_{t=0}^{\infty} \beta^t \left(\pi^i(\mathbf{x}_t, \mathbf{u}_t) + \right. \\ \left. + p_t^i\left(f(\mathbf{x}_t, \mathbf{u}_t) - \mathbf{x}_{t+1}\right) + \lambda_t^i g_i(\mathbf{x}_t, \mathbf{u}_t)\right) \end{aligned} \tag{2}$$

The first order condition for the optimization is given $\forall t$:

$$\begin{aligned} \tfrac{\partial}{\partial u_i} \pi^i(\mathbf{x}_t, \mathbf{u}_t) + p_t^i \tfrac{\partial}{\partial u_i}\left(f(\mathbf{x}_t, \mathbf{u}_t) - \mathbf{x}_{t+1}\right) \\ + \lambda_t^i \tfrac{\partial}{\partial u_i} g(\mathbf{x}_t, \mathbf{u}_t) = 0 \end{aligned} \tag{3}$$

and getting the dynamical equations by taking the Lagrangian derivatives with respect p_t^i and $\mathbf{x}_{t+1}$:

$$\begin{aligned} \mathbf{x}_{t+1} = &\ f(\mathbf{x}_t, \mathbf{u}_t) \\ \beta^{-1} p_t^i = &\ \tfrac{\partial}{\partial \mathbf{x}_{t+1}} \pi^i(\mathbf{x}_{t+1}, \mathbf{u}_{t+1}) \\ & + p_t^i \tfrac{\partial}{\partial \mathbf{x}_{t+1}}\left(f(\mathbf{x}_{t+1}, \mathbf{u}_{t+1}) - \mathbf{x}_{t+2}\right) \\ & + \lambda_t^i \tfrac{\partial}{\partial \mathbf{x}_{t+1}} g_i(\mathbf{x}_{t+1}, \mathbf{u}_{t+1}) \end{aligned} \tag{4}$$

including the complementary slackness condition and the positiveness of multipliers λ_t^i. It should be noted that the way to solve the game through the Lagrangian is by solving eqs. (3), (4) $\forall t$ and $\forall i$.

An alternative to this approach is to define the game as an equivalent Multivariate Optimum Control Problem. For this, we consider the following control problem for a yet unspecified function $\Pi(\mathbf{x}_t, \mathbf{u}_t)$:

$$\begin{aligned} & \max_{\{\mathbf{u}_t\}} \sum_{t=0}^{\infty} \beta^t \Pi(\mathbf{x}_t, \mathbf{u}_t) \\ & s.t.: \mathbf{x}_{t+1} = f(\mathbf{x}_t, \mathbf{u}_t), g(\mathbf{x}_t, \mathbf{u}_t) \le 0 \end{aligned} \tag{5}$$

Similarly to what is done in the game, we can find the optimal solution from the Lagrangian

$$\begin{aligned} \mathcal{J}(\mathbf{x}_t, \mathbf{u}_t, p_t, \lambda_t) = \textstyle\sum_{t=0}^{\infty} \beta^t \left(\Pi(\mathbf{x}_t, \mathbf{u}_t) + \right. \\ \left. + p_t\left(f(\mathbf{x}_t, \mathbf{u}_t) - \mathbf{x}_{t+1}\right) + \lambda_t g(\mathbf{x}_t, \mathbf{u}_t)\right) \end{aligned} \tag{6}$$

by getting the partial derivatives in the Lagrangian variables. In order for the game in (1) to be equivalent to (5), functions π^i must satisfy the following conditions $\forall i, j$ [6, 7]:

$$\text{C1. } \frac{\partial^2 \pi^i}{\partial u_i \partial x_j} = \frac{\partial^2 \pi^j}{\partial x_j \partial u_i}$$

$$\text{C2.a } \frac{\partial^2 \pi^i}{\partial u_i \partial u_j} = \frac{\partial^2 \pi^j}{\partial u_j \partial u_i}, \text{C2.b } \frac{\partial^2 \pi^i}{\partial x_i \partial x_j} = \frac{\partial^2 \pi^j}{\partial x_j \partial x_i}$$

$$\text{C3.a } \frac{\partial}{\partial x_j} \int \sum_{i=1}^{Q} \frac{\partial \pi^i}{\partial u_i} du_i = \int \sum_{i=1}^{Q} \frac{\partial^2 \pi^i}{\partial u_i \partial x_j} du_i$$

$$\text{C3.b } \frac{\partial}{\partial u_j} \int \sum_{i=1}^{Q} \frac{\partial \pi^i}{\partial x_i} dx_i = \int \sum_{i=1}^{Q} \frac{\partial^2 \pi^i}{\partial x_i \partial u_j} dx_i$$

If these conditions are fulfilled, the equivalent control problem is finally given by:

$$\Pi(\mathbf{x}, \mathbf{u}) = \int \sum_{i=1}^{Q} \frac{\partial \pi^i}{\partial x_i} dx_i + \int \sum_{i=1}^{Q} \frac{\partial \pi^i}{\partial u_i} du_i \tag{7}$$

including the same set of constraints as in the original problem. Note that $\Pi(\mathbf{x}, \mathbf{u})$ is expressed in terms of line integrals.

Solving both the optimization control problem (5) or the system of equations (3), (4) would yield the solution of the game, but its complexity may still be high enough to prevent obtaining it through the Lagrangian in both cases. However, the MOCP presents an easier structure to exploit, allowing a representation through Dynamic programming, as we will see in section 4.2.

3. UPLINK POWER DYNAMIC GAME AS A POTENTIAL DYNAMIC GAME

Let us now analyze the following particular game in the form of equation (1):

$$\max_{\{u_{it}\}} \sum_{t=0}^{\infty} \beta^t \left(\log \left(1 + \frac{\left|h_i\right|^2 u_{it}}{1 + \sum\limits_{j \neq i} \left|h_j\right|^2 u_{jt}} \right) + \alpha x_{it} \right) \tag{8}$$

$$\begin{aligned} s.t. :& x_{it+1} = x_{it} - u_{it},\ x_{i0} = X_i^{\max}, \\ & 0 \leq u_{it} \leq U_i^{\max},\ 0 \leq x_{it} \leq X_i^{\max},\ \forall i \end{aligned}$$

assuming t as an integer variable. It can be noticed that the first term corresponds to maximizing a capacity term associated to each user, where h_i is the channel coefficient of user i and u_{it} is the power used by user i at time t. The second term (parameter α is just a scaling parameter as a degree of freedom to weight properly both terms) corresponds to the state of that user defined as the battery level measured as the remaining power/energy left in the battery. Clearly, at every time step, the more power is used, the less energy is left. Also, standard constraints on the instantaneous power and energy level apply. This problem is inspired by [10] but adding in this case one time-varying term representing energy consumption.

In Annex A we show that (8) fulfills the conditions to be solved as a MOCP where:

$$\Pi\left(\mathbf{x}_t, \mathbf{u}_t\right) = \log\left(1 + \sum_{m=1}^{Q} \left|h_m\right|^2 u_{mt}\right) + \alpha \sum_{i=1}^{Q} x_{it} \tag{9}$$

Thus the equivalent control problem to the game in (8) is given by:

$$\begin{aligned} &\max_{\{\mathbf{u}_t\}} \sum_{t=0}^{\infty} \beta^t \left(\log\left(1 + \sum_{m=1}^{Q} \left|h_m\right|^2 u_{mt}\right) + \alpha \sum_{i=1}^{Q} x_{it} \right) \\ &s.t. : x_{it+1} = x_{it} - u_{it},\ x_{i0} = X_i^{\max}, \\ &\qquad 0 \leq u_{it} \leq U_i^{\max},\ 0 \leq x_{it} \leq X_i^{\max} \forall i \end{aligned} \tag{10}$$

Therefore, we have shown that if we are able to solve the MOCP given by (10), we can guarantee that its solution is a Nash equilibrium of the original problem given by equation (8). We will solve this problem with a centralized algorithm and with perfect channel state information.

4. SOLVING THE GAME

We will show next two approaches to solve the control problem avoiding the Lagrange approach explained in Sec. 2.

Algorithm 1 Multi-level Waterfilling Algorithm

```
1. Initialize $u_{it}$ for all $i \in \mathcal{Q}$ and $t \in \{0 \ldots N-1\}$, $k \leftarrow 0$ and tolerance $\varepsilon$.
2. For each $i \in \mathcal{Q}$, do
   (a) Compute $S_t^i = \frac{1+\sum_{j \neq i} |h_j|^2 u_{jt}}{|h_i|^2}, t \in \{0 \ldots N-1\}$
   (b) Set $\overline{\mu}_0^i = \max_t \left\{U_i^{\max} + S_t^i\right\}$ and $\underline{\mu}_0^i = 0$
       i.   Calculate $\mu_0^i\left(k\right) = \frac{\overline{\mu}_0^i + \underline{\mu}_0^i}{2}$ and determine $\mu_{t+1}^i\left(k\right)$ for all $t \in \{0 \ldots N-2\}$.
       ii.  Apply the waterfilling rule
            $u_{it}\left(k\right) = \left[\mu_t^i\left(k\right) - S_t^i\right]_0^{U_i^{\max}} \quad t \in \{0 \ldots N-1\}$
       iii. If $\sum_{t=0}^{N-1} u_{it} \geq X_i^{\max}$ set $\overline{\mu}_0^i \leftarrow \mu_t^i\left(k\right)$, otherwise $\underline{\mu}_0^i \leftarrow \mu_t^i\left(k\right)$.
       iv.  Set $k \leftarrow k+1$
       v.   Repeat from step (2(b)i) until $\overline{\mu}_0^i - \underline{\mu}_0^i \leq \varepsilon$
   (c) Set $u_{it} \leftarrow u_{it}(k)$ for player $i$
3. Repeat from step (2) until stopping criteria is met.
```

4.1. Waterfilling Algorithm

We propose first to solve the MOCP for a finite horizon for a sufficiently large time limit, allowing this way an efficient solution. Forming the Lagrangian from problem (10) with finite time horizon yields

$$\begin{aligned} \mathcal{L}\left(\mathbf{x}_t, \mathbf{u}_t, \mathbf{p}_t\right) = \sum_{t=0}^{N-1} \beta^t \Bigg(\log\left(1 + \sum_{m=1}^{Q} \left|h_m\right|^2 u_{mt}\right) + \\ + \alpha \sum_{i=1}^{Q} x_{it} + \sum_{i=1}^{Q} p_t^i \left(x_{it+1} - x_{it} + u_{it}\right) \Bigg) \\ s.t. \quad 0 \leq u_{it} \leq U_i^{\max}, \quad 0 \leq x_{it} \leq X_i^{\max},\ \forall i, t \end{aligned}$$

and solving for the control variables with the KKT condition $\partial \mathcal{L} / \partial u_{it} = 0$ for all $i \in \mathcal{Q}$ results into

$$u_{it} = \left[\frac{1}{p_t^i} - \frac{1 + \sum_{j \neq i} \left|h_j\right|^2 u_{jt}}{\left|h_i\right|^2} \right]_0^{U_i^{\max}} \tag{11}$$

where $[z]_a^b := \min\left\{\max\left\{z, a\right\}, b\right\}$, $t \in \{0 \ldots N-1\}$ and where p_t^i represents the inverse of a water level to be determined. To calculate these dual variables, the KKT condition $\partial \mathcal{L} / \partial x_{it+1} = -\beta^{t+1}\left(\alpha + p_{t+1}^i\right) + \beta^t p_t^i = 0$ of the state variables provide the relation $p_{t+1}^i = \frac{1}{\beta} p_t^i - \alpha, \quad t \in \{0 \ldots N-2\}$ which forms a recursive set of equations, for a given first value p_0^i. The interpretation behind this result is graphically understood with water levels which are different from each other at all instants, but that are related with the adjacent slots as the steps of a staircase would be, where the position of every step is given by p_{t+1}^{i}. This multi-level waterfilling is novel in its result, and can be formalized in Algorithm 1, where we have introduced water level variables $\mu_t^i = \frac{1}{p_t^i}$ for numerical reasons, that transform p_{t+1}^i into $\mu_{t+1}^i = \beta\mu_t^i/1 - \alpha\beta\mu_t^i$. We have solved the MOCP by applying a Gauss-Seidel sequence of updates, where each control variable determines the interference parameters on step (2a) with the last known control variables from previously optimized players. In addition to this, we have used a bisection algorithm to determine the unknown water levels μ_0^i and recursively solved for $t \in \{0 \dots N-2\}$. The convergence of Algorithm 1 is guaranteed due to the strict concavity of the MOCP and boundness of the variable sets.

4.2. Iterative solutions based on Dynamic Programming

The control problem in (10) can be rewritten in a more compact way following Dynamic Programming Principles as:

$$\begin{aligned} V(\mathbf{x}_0) &= \max_{\mathbf{u}_0} \Pi(\mathbf{x}_0, \mathbf{u}_0) + \beta V(f(\mathbf{x}_0, \mathbf{u}_0)) \\ s.t.:\ & g(\mathbf{x}_t, \mathbf{u}_t) \leq 0 \end{aligned} \tag{12}$$

where $\Pi(\mathbf{x}_t, \mathbf{u}_t)$ follows the definition in (9) and $g(\mathbf{x}_t, \mathbf{u}_t)$ represent the boundary restrictions in (10). We can solve the previous problem iteratively by following a value function iterative approach assuming that we are able to solve the right hand side of equation (12) obtaining the optimal trajectory $\mathbf{u}_0^* = \phi(\mathbf{x}_0)$ and substituting back in (12) we get:

$$V(\mathbf{x}_0) = \Pi(\mathbf{x}_0, \phi(\mathbf{x}_0)) + \beta V(f(\mathbf{x}_0, \phi(\mathbf{x}_0))) \tag{13}$$

Given that $V(\cdot)$ is unknown in a first stage, we propose to iterate as shown in Algorithm 2. In practice, the loop ends when a certain condition on the stability of the solution is fulfilled. Policy iteration procedures can also be applied in a very similar way. Note the solution is stationary, and that for every state there is an optimal strategy.

Algorithm 2 Value function iterative algorithm

```
1. Initialize $V_0(\boldsymbol{x}) = 0$
2. for $k = 0$ to $\infty$, do
   (a) find $\phi_k(\mathbf{x}) = \text{argmax}_{\mathbf{u}} \Pi(\mathbf{x}, \mathbf{u}) + \beta V_k(g(\mathbf{x}, \mathbf{u}))$
   (b) $V_{k+1}(\mathbf{x}) = \Pi(\mathbf{x}, \phi_k(\mathbf{x})) + \beta V_k(g(\mathbf{x}, \phi_k(\mathbf{x})))$
```

5. RESULTS

In our simulation scenario we have considered $Q = 4$ players, a maximum transmit power level of $U_i^{\max} = 5$, total battery power $X_i^{\max} = 33$, forgetting value $\beta = 0.95$ and weighting value $\alpha = 0.001$. We have simulated both alternatives given in previous section and they provide similar results, for that reason we just show here results for the waterfilling algorithm with time horizon $N = 100$. Channels are randomly obtained with zero mean gaussian complex distribution. We can observe in Figure 1 that players transmit in strict order where users that have better channels transmit first, and then the rest. It seems, that this division in time avoids interference to other users, and allows to achieve the highest value in the potential function, and equivalently, in the game. This is the case, because the utility considers that running out of battery is detrimental towards the device's performance and so users decide to save battery until the channel becomes empty, coordinating to avoid collisions in time. Note that the delay shown among users appears naturally in the game and has not been introduced in any equation. This is a feature that could be taken into account with a different formulation in the future.

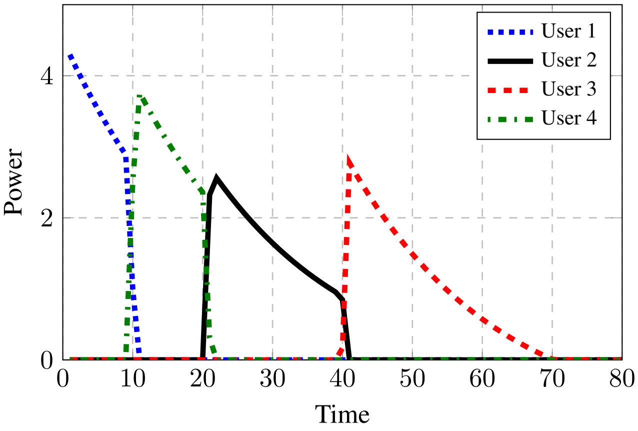

Fig. 1. Power Allocation of players

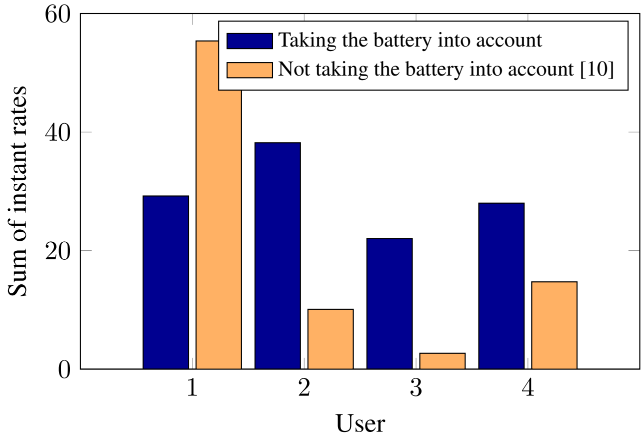

Fig. 2. Comparison according to battery considerations

Finally in Figure 2 we have plotted the total sum of instant capacities of each player as solved by Algorithm 1 (in blue) vs. the proposed algorithm in [10] that does not consider the state of the battery in its formulation (in orange):

$$u_{it} = \underset{0 \leq u_{it} \leq U_i^{\max}}{\text{argmax}} \log\left(1 + \frac{|h_i|^2 u_{it}}{1 + \sum_{j \neq i} |h_j|^2 u_{jt}}\right) - \gamma |h_i|^2 u_{it}$$

with $\gamma = 0.1$. With this comparison we simply state, that having into consideration the battery life of the devices allows to transmit more information for the life duration of the network.

6. CONCLUSION

This work formulates an uplink power control scenario as a dynamic game, given that the battery of each user decreases as they assign power in different time slots. To further solve the dynamic game, first the game is reformulated as a control problem, and latter waterfilling and iterative approaches are proposed to solve the control problem. The results show an interesting behavior where the users transmit in strict order following a channel quality criteria.

A. ANNEX

We will prove next that $\pi^i(\mathbf{x}_t, \mathbf{u}_t)$ in (8) fullfils C1-C3 and then solve the line integral in (7) to get the potencial function $\Pi(\mathbf{x}_t, \mathbf{u}_t)$. C1 is trivial $\frac{\partial^2 \pi^i}{\partial u_i \partial x_j} = \frac{\partial^2 \pi^j}{\partial x_j \partial u_i} = 0 \quad \forall i, j$. To validate C2.a we proceed as follows:

$$\frac{\partial^2 \pi^i}{\partial u_i \partial u_j} = -\frac{|h_i|^2 |h_j|^2}{\left(1 + \sum_m |h_m|^2 u_{mt}\right)^2} \tag{14}$$

and due to the symmetric structure in (14), it is straightforward to show that C2.a is satisfied. Identically for C2.b $\frac{\partial^2 \pi^i}{\partial x_i \partial x_j} = \frac{\partial^2 \pi^j}{\partial x_j \partial x_i} = 0$. And finally, for C3.a and C3.b we have:

$$\begin{aligned} \frac{\partial}{\partial x_j} \int \sum_{i=1}^{Q} \frac{\partial \pi^i}{\partial u_i} du_i = \int \sum_{i=1}^{Q} \frac{\partial^2 \pi^i}{\partial u_i \partial x_j} du_i = 0 \\ \frac{\partial}{\partial u_j} \int \sum_{i=1}^{Q} \frac{\partial \pi^i}{\partial x_i} dx_i = \int \sum_{i=1}^{Q} \frac{\partial^2 \pi^i}{\partial x_i \partial u_j} dx_i = 0 \end{aligned} \tag{15}$$

We solve now (7) in order to obtain the corresponding equivalent optimal control problem. Let us analyze each term individually by defining $\boldsymbol{\xi} : [0,1] \to (\mathcal{U}_1 \times \ldots \times \mathcal{U}_Q)$ like a piecewise continuously differentiable path in the utility domain with $\xi_i(0) = 0$ and $\xi_i(1) = u_i$ and $\boldsymbol{\eta} : [0,1] \to (\mathcal{X}_1 \times \ldots \times \mathcal{X}_Q)$ like a piecewise continuously differentiable path in the state domain with $\eta_i(0) = 0$ and $\eta_i(1) = x_i$. We must recall that in this case, initial state conditions would not be null because batteries start from a full level, but we have simplified the expression removing a constant term that is also considered when defining the constraints:

$$\begin{aligned} &\int \sum_{i=1}^{Q} \frac{\partial \pi^i}{\partial u_i} du_i = \int_0^1 \sum_{i=1}^{Q} \frac{\partial \pi^i(\mathbf{x}, \boldsymbol{\xi})}{\partial u_i} \frac{d\xi_i(\lambda)}{d\lambda} d\lambda = \\ &= \int_0^1 \frac{\sum_{i=1}^{Q} |h_i|^2}{1 + \sum_m |h_m|^2 \xi_m(\lambda)} \frac{d\xi_i(\lambda)}{d\lambda} d\lambda = \\ &= \log\left(1 + \sum_m |h_m|^2 \xi_m(1)\right) - \log\left(1 + \sum_m |h_m|^2 \xi_m(0)\right) \end{aligned}$$

The third equality results from the fact that $\sum_{i=1}^{Q} |h_i|^2 \xi_i'$ is the derivative of the sum term in the denominator of the integral. For the state term:

$$\begin{aligned} &\int \sum_{i=1}^{Q} \frac{\partial \pi^i}{\partial x_i} dx_i = \int_0^1 \sum_{i=1}^{Q} \frac{\partial \pi^i(\boldsymbol{\eta}, \mathbf{u})}{\partial x_i} \frac{d\eta_i(\lambda)}{d\lambda} d\lambda = \\ &= \alpha \int_0^1 \sum_{i=1}^{Q} \frac{d\eta_i(\lambda)}{d\lambda} d\lambda = \alpha \sum_{i=1}^{Q} \eta_i(1) - \alpha \sum_{i=1}^{Q} \eta_i(0) \end{aligned}$$

Finally with the initial conditions defined before for $\xi_i(\cdot)$ and $\eta_i(\cdot)$ and introducing the time reference we get:

$$\Pi(\mathbf{x}_t, \mathbf{u}_t) = \log\left(1 + \sum_m |h_m|^2 u_{mt}\right) + \alpha \sum_{i=1}^{Q} x_{it}$$

REFERENCES

[1] E. Dockner, *Differential Games in Economics and Management Science*, Cambridge University Press, Nov. 2000.

[2] T. Alpcan, T. Başar, R. Srikant, and E. Altman, “CDMA uplink power control as a noncooperative game,” *Wireless Networks*, vol. 8, no. 6, pp. 659–670, Nov. 2002.

[3] A. M. Mishra, “Bandwidth allocation with dynamic selection in wireless networks,” in *ICCCE 2012*, April 2012.

[4] K. Zhu, D. Niyato, P. Wang, and Z. Han, “Dynamic spectrum leasing and service selection in spectrum secondary market of cognitive radio networks,” *IEEE Trans. Wireless Commun.*, vol. 11, no. 3, pp. 1136–1145, 2012.

[5] O. Hernandez-Lerma and D. G. Sanchez, *Discrete Time Stochastic Control and Dynamic Potential Games: The Euler Equation Approach*, Springer International Publishing, Aug. 2013.

[6] W.D. Dechert, “Optimal control problems from second-order difference equations,” *Journal of Economic Theory*, vol. 19, no. 1, pp. 50–63, Oct. 1978.

[7] W.D. Dechert, “Non cooperative dynamic games: a control theoretic approach,” *Unpublished. Available on request to the author*, 1997.

[8] W. D. Dechert and W. A. Brock, “The lake game,” *Economic and Ecological Modelling, The Beijer Institute, The Royal Academy of Sciences*, , no. 24, 2000.

[9] W.D. Dechert and S.I. O’Donnell, “The stochastic lake game: A numerical solution,” *Journal of Economic Dynamics and Control*, vol. 30, pp. 1569–1587, Sept. 2006.

[10] G. Scutari, S. Barbarossa, and D.P. Palomar, “Potential games: A framework for vector power control problems with coupled constraints,” in *ICASSP 2006 Proceedings*, 2006, vol. 4, pp. IV–IV.