Joint power control-allocation for green cognitive wireless networks using mean field theory

Hamidou Tembine *, Samson Lasaulce*, Marc Jungers†
* LSS (CNRS–Supélec-Paris 11), 3 rue Joliot-Curie, 91192 Gif-sur-Yvette Cedex, France
† CNRS CRAN ENSEM , 2 avenue de la foret de Haye 54516 Vandoeuvre Cedex, France

***Abstract*—The purpose of this paper is to show how key concepts from control theory, game theory, and mean field theory can be exploited to design joint control-allocation policies in cognitive wireless networks. One of the key features of the proposed approach is that transmitters (which are assumed to be cognitive and autonomous decisionnally speaking) have a certain knowledge of the channel evolution law and want to reach a certain transmission rate target while minimizing the energy consumed by the power supply and not the one corresponding to radio-frequency signals (which is known to be important to design green wireless networks). The optimal centralized policy is derived in order to have an upper bound on the performance of the decentralized system. Then, the Nash equilibrium of the cognitive network is determined by using recent results from mean field theory (MFT). In order to evaluate the performance gap between decentralized and centralized policies we introduce and evaluate the MFT-based asymptotic price of anarchy (APoA).**

I. Introduction

In cognitive wireless networks, transmitters are envisioned to choose freely their power control and power allocation policies. In such a framework, a transmitter acts as a free (and often selfish) decision-maker. Because transmitters interact through multiuser interference and using common resources (band, time-slots, space, etc), a relevant paradigm for analyzing such (distributed) problems is game theory (see e.g., [1]). While game theory is now well used in the wireless community to analyze distributed power control problems, key tools from control theory are rarely used. In the present work we propose to exploit control theory for essentially two reasons. First, it can allow one to take into account some propagation effects such as the shadowing effect [13]. Second, and more importantly regarding green wireless networks [14] it allows one to associate with a given power control policy a cost in terms of energy consumed by the power supply and not the radio-frequency part of the transmitter as it is usually the case. In control theory it is well-known that a smooth control policy generally requires less energy than a policy having a large peak-to-average ratio. The shift of paradigm towards green and cognitive wireless networks is one of the strong motivations for providing the theoretical analysis presented in this paper.

More specifically, we analyze the scenario of parallel multiple access channels with block fading links. The parallel multiple access channel consists of several transmitters having several orthogonal channels at their disposal to communicate with a common receiver. For example, the transmitters can be mobile stations and the receiver can be a set of base stations using non-overlapping bands and connected together by a radio network controller. As channels are assumed to be time-varying, transmitters have to allocate their total transmit power between the available channels and control their total transmit power over time as the channel gains vary. As it will be seen further, the assumption of cognitive transmitters is important to consider them as free decision-makers and able to acquire or learn some information about the communication system. In terms of information assumptions, one key feature of the work reported in this paper is that transmitters have a certain knowledge of the channel dynamics. The latter is exploited by the transmitters to optimize their joint control-allocation policy.

The structure of this paper is as follows. First, in Sec. II the system model is described in details. In Sec. III-C, the existence of an optimal centralized policy is shown by using the Hamilton-Jacobi-Bellman-Fleming (HJBF) condition for stochastic optimal control [12]. This leads to a Riccati equation, which can be solved easily. The same approach is adopted to tackle the problem of existence of decentralized or distributed policies. However, the determination of the individual and selfish control-allocation policies (namely a Nash equilibrium) requires solving a complex system of equations. Therefore, recent results from mean field limit [7], [3], [8], [10] are exploited to approximate the dynamic random process corresponding to multiuser interference in order to find an explicit solution for the selfish policies used by the cognitive transmitters. Sec. IV introduces and provides the asymptotic price of anarchy (APoA) for the problem under investigation. Summarizing comments and extensions are given in Sec. IV-C.

II. System model

A. Received signals

We consider networks which can be modeled by a parallel multiple access channel, which consists of several orthogonal multiple access channels. The sets of transmitters and channels are respectively denoted by $\mathcal{N}$ and $\mathcal{L}$. The cardinality of $\mathcal{N}$ is n, and the cardinality of $\mathcal{L}$ is L. The $L-$dimensional vector of received signals is denoted by $y = (y_1, \ldots, y_L)$ and defined by $y(t) = \sum_{j=1}^{n} H_j(t)s_j(t)+z(t)$, where t is the time index (the associated rate depends if fast fading or slow fading channels are considered), $\forall j \in \mathcal{N}$, $H_j(t)$ is the channel transfer matrix from the transmitter j to the receiver, $s_j(t)$ is the vector of

Digital Object Identifier: 10.4108/ICST.CROWNCOM2010.9219
http://dx.doi.org/10.4108/ICST.CROWNCOM2010.9219

symbols transmitted by transmitter j at time t, and the vector $z(t)$ represents the noise observed at the receiver, which is assumed to be centered, Gaussian, and distributed as $\mathbb{E}\left(zz^{\dagger}\right) = \mathrm{diag}\left(N_{0,1}, \ldots, N_{0,L}\right)$. We will exclusively deal with the scenario where $\forall j \in \mathcal{N}$, matrix H_j is an m-dimensional diagonal matrix, i.e., $H_j(t) = \mathrm{diag}\left(h_{j,1}(t), \ldots, h_{j,L}(t)\right)$, whose entries $h_{j,\ell}(t)$, $\forall \ell \in \mathcal{L}$, are i.i.d. complex Gaussian random variables with independent real and imaginary parts, each with zero-mean and variance $N_{0,\ell}$. This scenario models, for instance, the case of intersymbol interference channels which have been diagonalized in the frequency domain by using a pre-cyclic prefix and the Fourier transform, or the case of the uplink of a network with several access receivers operating over non-overlapping flat-fading channels. We denote by $h(t) = (h_1(t), \ldots, h_n(t))$ where $h_j(t) = (h_{j,1}(t), \ldots, h_{j,L}(t))$. The vector of transmitted symbols $s_j(t)$, $\forall j \in \mathcal{N}$ is characterized in terms of power by the covariance matrix $P_j(t) = \mathbb{E}\left(s_j(t)s_j^{\dagger}(t)\right) = \mathrm{diag}\left(p_{j,1}(t), \ldots, p_{j,L}(t)\right)$. As a matter of fact, $p_{j,\ell}(t)$, $\forall (j,\ell) \in \mathcal{N} \times \mathcal{L}$, represents the transmit power allocated by the transmitter j over the AP ℓ. Now, since transmitters are power-limited, we have that

$$\forall j \in \mathcal{N}, \quad \sum_{\ell=1}^{L} p_{j,\ell}(t) \leq p_{j,\max}(t). \tag{1}$$

We define a power allocation vector for transmitter $j \in \mathcal{N}$ as a vector $p_j(t) = (p_{j,1}(t), \ldots, p_{j,L}(t))$ with non-negative entries satisfying (1). We denote by $p(t) = (p_1(t), \ldots, p_n(t))$. At last, the band of channel $\ell \in \mathcal{L}$ will be denoted by W_ℓ and the one of the system by $W = \sum_{\ell=1}^{L} W_\ell$.

B. Channel dynamics

We use the change of variables $g_{j,\ell}^n(t) := e^{X_{j,\ell}^n(t)}$. The variable $X_{j,\ell}^n(t)$ will be called the individual (channel) state of transmitter j and the super vector $X^n(t) = (X_1^n, ..., X_n^n)$ with $X_j^n = (X_{j,1}^n, ...X_{j,L}^n)$ will be called the state of the system or network. As in [13] we assume that the state $X_{j,\ell}^n(t)$ follows the corresponding evolution or dynamics:

$$\mathrm{d}X_{j,\ell}^n(t) = -a_{j,\ell}^n(X_{j,\ell}^n(t) + b_{j,\ell}^n)\mathrm{d}t + \sigma_{j,\ell}^n \mathrm{d}\mathbb{W}_{j,\ell}(t),\ t \geq 0, \tag{2}$$

where $a_{j,\ell}^n, b_{j,\ell}^n, \sigma_{j,\ell}^n$ are reals, $\mathbb{W}_{j,\ell}(t)$ are mutually independent Wiener processes. This model is very well justified in [13]. In particular, let us point out that the deterministic part allows one to account for the path loss and shadowing effects and the random part can also model rapid channel variations or channel uncertainty (when channel gains are estimated). Now we introduce some notations and definitions to simplify the analysis of the problem.

Individual state evolution: Denote by

$$A_j^n = \mathrm{diag}(a_{j,\ell}^n, \ell \in \mathcal{L}) = \begin{pmatrix} a_{j,1}^n & 0 & \ldots & 0 \\ 0 & a_{j,2}^n & \ldots & 0 \\ \ldots & \ldots & \ddots & \vdots \\ 0 & 0 & 0 & a_{j,L}^n \end{pmatrix} \tag{3}$$

Similarly, $B_j^n = \mathrm{diag}(a_{j,\ell}^n b_{j,\ell}^n, \ell \in \mathcal{L})$, $\Gamma_j^n = \mathrm{diag}(\sigma_{j,\ell}^n,\ \ell \in \mathcal{L})$. Then the individual state evolution is described by $\mathrm{d}X_j^n(t) = -[A_j^n X_j^n(t) + B_j^n]\mathrm{d}t + \Gamma_j^n \mathrm{d}\mathbb{W}_j(t)$.

System state evolution: The system state evolution is obtained by combining all the individual state evolution. Denote by $A^n = \mathrm{diag}(A_1^n, \ldots, A_n^n)$ be the $nL \times nL$ dimensional matrix given by

$$\begin{pmatrix} A_1^n & 0 & \ldots & 0 \\ 0 & A_2^n & \ldots & 0 \\ \ldots & \ldots & \ddots & \vdots \\ 0 & 0 & 0 & A_n^n \end{pmatrix}. \tag{4}$$

Similarly, we define the matrices B^n and Γ^n. The system state evolution is then given by

$$\mathrm{d}X^n(t) = -[A^n X^n(t) + B^n]\mathrm{d}t + \Gamma^n \mathrm{d}\mathbb{W}(t). \tag{5}$$

Control signal of each transmitter: Usually, by control it is meant finding the vector valued function $P_j(t)$ for transmitter j, which represents the radio-frequency signal emitted by the transmitter. Here, we introduce an auxiliary function $u_{j,\ell}^n(t)$, which a signal controlling the transmitter, and represents the power amplifier supply. At each time t, transmitter j adjusts its power allocation scheme at each link l by using the control signal $u_{j,\ell}^n(t)$. This control signal is linked to the transmit power used for this link by the relation

$$\mathrm{d}p_{j,\ell}^n(t) = u_{j,\ell}^n(t)\mathrm{d}t. \tag{6}$$

We do not assume any restriction on the domain in which U^n will be in. By defining a cost function associated to each control, it is clear that any optimal control u which may depend on the state, will be almost surely in a bounded set. We restrict our attention to this simple class of strategies. Note that, in general differential games, there are more general classes of strategies (non-anticipative strategies with/out delay) under which the analysis can be investigated. At last, we define $\mathrm{d}P_j^n(t) = U_j^n(t)\mathrm{d}t$ where U_j^n is the vector with component l equal to $u_{j,\ell}^n$.

III. The joint power control-allocation problem in the finite case

A. Definition of the joint power control-allocation game

The joint power control-allocation game can be described under a strategic form. The set of players is $\mathcal{N}$, which is the set of transmitters. The strategy of a transmitter is the control-allocation vector $u_j^n(t) = (u_{j,1}^n(t), \ldots, u_{j,L}^n(t))$. The objective of a transmitter is to reach a transmission rate or signal-to-interference-plus-noise ratio (SINR) target with the minimum effort in terms of energy consumed by the power supply. When the receiver implement single-user decoding, the SINR of transmitter j on channel ℓ is given by

$$\mathrm{SINR}_{j,\ell}^n(t) = \frac{p_{j,\ell}^n(t) g_{j,\ell}^n(t)}{N_{0,\ell} + \sum_{i \in \mathcal{N} \setminus \{j\}} p_{i,\ell}^n(t) g_{i,\ell}^n(t)}. \tag{7}$$

The constraint for transmitter j is to insure a minimum requirement for the quality of service in terms of rate:

$$\forall j,\ \forall l,\ \forall t, \log_2(1 + \mathrm{SINR}_{j,\ell}^n(t)) \geq \bar{w}_{j,\ell}^n \tag{8}$$

which is equivalent to

$$\mathrm{SINR}_{j,\ell}^n(t) \geq 2^{\bar{w}_{j,\ell}^n} - 1 =: \gamma_{j,\ell}^n. \tag{9}$$

Digital Object Identifier: 10.4108/ICST.CROWNCOM2010.9219
http://dx.doi.org/10.4108/ICST.CROWNCOM2010.9219

This inequality is equivalent to

$$\left[e^{X^n_{j,\ell}(t)}p^n_{j,\ell}(t) - \gamma^n_{j,\ell}\left(N_{0,\ell} + \sum_{i\neq j} e^{X^n_{i,\ell}(t)}p^n_{i,\ell}(t)\right)\right] \geq 0 \quad (10)$$

In practice, this amounts to minimizing the following quadratic term in the power vectors: $\sum_{\ell\in\mathcal{L}}\left[e^{X^n_{j,\ell}(t)}p^n_{j,\ell}(t) - \gamma^n_{j,\ell}\left(N_{0,\ell} + \sum_{i\neq j} e^{X^n_{i,\ell}(t)}p^n_{i,\ell}(t)\right)\right]^2$. The objective of transmitter j is to minimize the above quantity with a minimum amount of energy which is translated by cost matrix $R_j = \text{diag}\left(\mathbb{E}[(u^n_{j,1}(t))^2], ..., \mathbb{E}[(u^n_{j,L}(t))^2]\right) = \text{diag}(r^n_{j1}, ..., r^n_{jL})$. The payoff function for transmitter j is therefore defined by:

$$\Phi_{j,\beta}(U) = \mathbb{E}\int_0^\infty e^{-\beta t}[-\sum_{l=1}^L \left(\gamma^n_{j,\ell} I^n_{-j,\ell}(t) - e^{X^n_{j,\ell}(t)}p^n_{j,\ell}(t)\right)^2 \\ -\langle U^n_j, R_j U^n_j\rangle]\ \mathrm{d}t$$

where β is a positive parameter allowing the transmitters to evaluate the current and future payoffs differently, $I^n_{-j,\ell}(t) = N_{0,\ell} + \sum_{i\neq j} e^{X^n_{i,\ell}(t)}p^n_{i,\ell}(t)$, and $U^n_j = (u^n_{j,1}, \ldots, u^n_{j,L})$.

B. Optimal centralized policies

The centralized power allocation problem is then given by the following dynamic optimization problem

$$\bar{V}^n(X,P) = \sup_{U^n}\ \mathbb{E}\left(\int_0^{+\infty} -e^{-\beta t}\mathbb{B}(X^n(t), P^n(t), U^n(t))\ dt \right. \\ \left. \mid\ X^n(0) = X, P^n(0) = P\right)$$

such that the individual state dynamics (ID) of all the transmitters are satisfied.

$$(ID)\begin{cases} \mathrm{d}X^n_{j,\ell}(t) = -a^n_{j,\ell}(X^n_{j,\ell} + b^n_{j,\ell})\mathrm{d}t + \sigma^n_{j,\ell}\mathrm{d}\mathbb{W}_{j,\ell},\ t \geq t_0, \\ \mathrm{d}p^n_{j,\ell}(t) = u^n_{j,\ell}(t)\mathrm{d}t \\ \forall j\in\mathcal{N},\ \forall \ell\in\mathcal{L} \end{cases}$$

where $U^n(t) = (U^n_1(t), \ldots, U^n_n(t))$, $\mathbb{B}(X,P,U) = \langle P, A_1(X)P\rangle + 2\langle A_2(X), P\rangle + A_3 + \langle U, RU\rangle$ is the total sum of payoffs of all the transmitters, obtained after expanding and collecting the polynomial term in P. The initial vector $X(0)$ is an element belonging to the set of feasible values along the optimal state trajectory at time 0. We assume that the vector (X,P) is given.

The basic objective is to maximize the total payoff function (the sum of all the transmitter payoffs). For large n, the exact globally optimal solution requires centralized information and this leads to high computational complexity. Using mean field limit under suitable assumptions, one can compute the optimal total payoff off-line; the resulting set of strategies is asymptotically not far to the global optimum as n tends to infinity. A general stochastic differential game approach can be found in [4].

The centralized problem can be written as

$$\begin{cases} \sup_{U^n}\ \mathbb{E}\int_0^{+\infty} e^{-\beta t}\left[-\sum_{j=1}^n\sum_{l=1}^L\left(\gamma^n_{j,\ell}I^n_{-j,\ell} - e^{X^n_{j,\ell}(t)}p^n_{j,\ell}(t)\right)^2\right. \\ \left. -\langle U^n(t), R.U^n(t)\rangle\right]\ dt \\ \mathrm{d}X^n(t) = -[A^nX^n(t) + B^n]\mathrm{d}t + \Gamma^n\mathrm{d}\mathbb{W}(t) \\ \mathrm{d}P^n(t) = U^n\mathrm{d}t \end{cases}$$

where $R = \text{diag}(R_j)_j$.

Since time is not explicitly involved in the above equation, the derived control u will be a function of X, P only. This problem is in the framework of linear-quadratic (LQ) stochastic optimal control. Hence one can obtain:

Proposition III-B1. *A set of controls $U_* = \eta(X,P)$ constitutes an optimal solution to the centralized problem if there exists continuously differentiable function $\bar{V}^n$ such that the Hamilton-Jacobi-Bellman-Fleming (HJBF) equation given by*

$$\beta\bar{V}^n(X,P) = \frac{1}{2}\sum_\ell\sum_j \sigma^{n,2}_{j,\ell}\frac{\partial^2}{\partial X^2_{j,\ell}}\bar{V}^n + \inf_U[r(X,P,U) + \langle D_X\bar{V}^n, f\rangle + \langle D_P\bar{V}^n, U\rangle] \quad (11)$$

$$= \frac{1}{2}\sum_\ell\sum_j \sigma^{n,2}_{j,\ell}\frac{\partial^2}{\partial X^2_{j,\ell}}\bar{V}^n + \left(r(X,P,\eta) + \langle D_X\bar{V}^n, f\rangle + \langle D_P\bar{V}^n, \eta\rangle\right) \quad (12)$$

is satisfied.

This implies that if there exists a regular function $\bar{V}^n$ such that the Hamilton-Jacobi-Bellman-Fleming (HJBF) optimality equation

$$\beta\bar{V}^n = \frac{1}{2}\sum_{j,\ell}\sigma^{n,2}_{j,\ell}\frac{\partial^2}{\partial X^2_{j,\ell}}\bar{V}^n + \sum_{j,\ell}a^n_{j,\ell}(X_{j,\ell} + b^n_{j,\ell})\frac{\partial}{\partial X_{j,\ell}}\bar{V}^n + \\ \langle P, A_1(X)P\rangle + 2\langle A_2(X), P\rangle + A_3 + \frac{1}{4}\langle D_P\bar{V}^n, R^{-1}D_P\bar{V}^n\rangle$$

is satisfied then the optimal control is

$$U^n_* = -\frac{1}{2}R^{-1}D_P\bar{V}^n. \quad (13)$$

To determine the optimal value $\bar{V}^n$ we write it as a quadratic function in P i.e there exists a symmetric matrix $\tilde{K}(X)$, a vectorial function $\tilde{S}(X)$ and mapping $\tilde{q}$ such that

$$\bar{V}^n(X,P) = \langle P, \tilde{K}(X)P\rangle + 2\langle P, \tilde{S}(X)\rangle + \tilde{q}(X)$$

By reincorporating these functions, we obtain the following partial differential systems

$$\beta\tilde{K} = \sum_{j,\ell}f_{j\ell}D_{X_{j\ell}}\tilde{K} - \tilde{K}R^{-1}\tilde{K} + A_1 + \frac{1}{2}\sum_j\sum_\ell\sigma^{n,2}_{j,\ell}\frac{\partial^2}{\partial X^2_{j,\ell}}\tilde{K} \quad (14)$$

$$\beta\tilde{S} = \sum_{j,\ell}f_{j,\ell}D_{X_{j,\ell}}\tilde{S} - \tilde{K}R^{-1}\tilde{S} + A_2 + \frac{1}{2}\sum_{j,\ell}\sigma^{n,2}_{j,\ell}\frac{\partial^2}{\partial X^2_{j,\ell}}\tilde{S} \quad (15)$$

$$\beta\tilde{q} = \langle f, D_X\tilde{q}\rangle - \langle\tilde{S}, R^{-1}\tilde{S}\rangle + A_3 + \frac{1}{2}\sum_{j,\ell}\sigma^{n,2}_{j,\ell}\frac{\partial^2}{\partial X^2_{j,\ell}}\tilde{q} \quad (16)$$

where $f_{j\ell} = -a^n_{j,\ell}(X^n_{j,\ell} + b^n_{j,\ell})$. By choosing the term $e^{X^n_{j\ell}(t)}p^n_{j\ell}(t)$, the first equation is known as standard Riccati equation. We conclude that $U^n_* = -R^{-1}[\tilde{K}(X^n)P^n + \tilde{S}(X^n)]$

$$u^n_{*,j\ell}(t) = -\frac{\tilde{K}_{jj\ell}(X^n)p^n_{j\ell}(t)}{r^n_{j\ell}} - \frac{1}{r^n_{j\ell}}\sum_{j\neq i}\tilde{K}_{ji\ell}(X^n)p^n_{i\ell}(t) - \frac{1}{r^n_{j,\ell}}\tilde{S}_{j\ell}(X^n).$$

Note that in the centralized scheme, each optimal control depends on its power levels but also on the weighted sum of the received power of the other users.

Digital Object Identifier: 10.4108/ICST.CROWNCOM2010.9219
http://dx.doi.org/10.4108/ICST.CROWNCOM2010.9219

C. Decentralized selfish policies

In this subsection, we focus on the decentralized power allocation problem. Each user solves its long-term payoff under individual state evolution constraint. We assume that each user knows the payoff functions $\Phi_\beta = (\Phi_{j,\beta})_{j\in\mathcal{N}}$ as well as the state dynamics.

$$\left\{\begin{array}{c} \sup_{U_j} \;\; \Phi_{j,\beta}(U) \\ \text{s.t} \\ \mathrm{d}X_j^n(t) = -[A_j^n X_j^n(t) + B_j^n]\mathrm{d}t + \Gamma_j^n \mathrm{d}\mathbb{W}_j(t) \\ \mathrm{d}P_j^n(t) = U_j^n(t)\mathrm{d}t \end{array}\right.$$

We define a feedback Nash equilibrium of the dynamic game. If the transmitters information structure is reduced to the initial profile X, P at time 0 and the trajectory generated, the strategies becomes functions of X, P and the current time t. A strategy $U_j(t)$ can be written as $U_j(t, X, P)$. Denote by $\mathcal{U}_j$ be the set of such strategies. Under this class of strategies, a feedback Nash equilibrium is a configuration in which none of the transmitters improves its long-term payoff by unilaterally changing its strategy. The following result gives the existence of feedback Nash equilibria in the stochastic dynamic power allocation game.

Proposition III-C1. *The stochastic dynamic power allocation has at least one feedback Nash equilibrium.*

A proof of the Proposition III-C1 can be obtained directly from [2], [5]. The result is obtained from the definition of Nash equilibrium and from the fact that by fixing all transmitters strategies, except the $j-$th one's, at their equilibrium choices (which are known to be feedback), we arrive at a fixed point of stochastic optimal control problem whose optimal solution is a feedback strategy.

2) Determination of Nash equilibria: A Nash equilibrium η is characterized as following: if there exists regular functions $\{V_j^n\}_j$ such that the systems of HJBF equations: for all $j \in \mathcal{N}$,

$$\begin{aligned} \beta V_j^n(X,P) &= \frac{1}{2}\sum_\ell\sum_j \sigma_{j,\ell}^{n,2}\frac{\partial^2}{\partial X_{j,\ell}^2}V_j^n \\ + \inf_U [r_j^n(X,P,\eta_j^U) &+ \langle D_X V_j^n, f\rangle + \langle D_P V_j^n, U\rangle] \\ &= \frac{1}{2}\sum_\ell\sum_j \sigma_{j,\ell}^{n,2}\frac{\partial^2}{\partial X_{j,\ell}^2}V_j^n + \\ [\tilde{r}_j^n(X,P,\eta) &+ \langle D_X V_j^n, f\rangle + \langle D_P V_j^n, \eta\rangle] \end{aligned}$$

where $\eta_j^U = (\eta_1, \ldots, \eta_{j-1}, U_j, \eta_{j+1}, \ldots, \eta_n)$. We use the HJBF equations in order to derive the Riccati equations. Thus, one gets a system of $3nL$ differential equations to solve to get the best response correspondence.

IV. Large-system analysis

A. Determination of the Nash equilibria

We now examine the non-cooperative decentralized mean field power allocation problem. When using HJBF optimality conditions, one gets a system of $3nL$ partial differential equations to solve. Moreover, fixed point equations and closed-loop parameters are in these equilibrium equations. In order to reduce the number of unknown parameters of these equations and to derive closed form expressions, we use mean field asymptotic approach [15]. We assume that the channel states evolutions described by (ID) satisfy asymptotically: $a_{j,\ell}^n \longrightarrow a, b_{j,\ell}^n \longrightarrow b_\ell, \sigma_{j,\ell}^n \longrightarrow \sigma_\ell, r_{j,\ell}^n \longrightarrow_\ell$. For each link ℓ, define the random measure $M_\ell^n = \frac{1}{n}\sum_{j=1}^n \delta_{z_{j,\ell}^n}$ with marginal $M_\ell^n(t) = \frac{1}{n}\sum_{j=1}^n \delta_{z_{j,\ell}^n(t)}$ where $z_{j,\ell}^n(t) = e^{X_{j,\ell}^n(t)}p_{j,\ell}^n(t)$. The long-term performance metric of the transmitter j is a function of the random measure $(M_\ell^n(t))_\ell$ and $z_j^n(t)$ i.e $\phi(z_j^n(t), M^n(t))$ where ϕ is continuous differentiable function (a quadratic function). The weak convergence of the occupancy measure to some deterministic measure $m^*(t)$ which evolves in time, combined with the weak convergence of the pair of processes $(z_j^n(t), M^n(t))$ determines the mean field convergence of the performance metric of user j.

We now use Kotelenez & Kurtz's result and the results in [7] on mean field convergence. Under independent Wiener process of individual state dynamics and the assumption that the states and the powers remain almost surely bounded (otherwise it is not optimal), one can use the theorem 3.1 and theorem 4.1 in Kotelenez & Kurtz (2008) which states that if the initial occupancy measure $M^n(0)$ is weakly convergent and the process $(z_{j,\ell}^n(0))_{j,\ell}$ are exchangeable, the mean field convergence holds. Using the asymptotic exchangeable assumptions of the Wiener processes $\mathbb{W}_{j,\ell}$ we conclude that the individual state evolution generated a random process $z_{j,\ell}^n(t)$ which has the mean field convergence of $M^n(t)$.

Proposition IV-A1. *If the process $(z_j^n(t))_j$ is asymptotically exchangeable, then $M^n(t)$ converge in distribution to a deterministic measure $m^*(t)$ (hence the convergence in probability holds because the limit is deterministic).*

Moreover, the pair of processes $(z_j^n(t), M^n(t))$ converges in law. Note that a similar result have been shown in [10] in the discrete state space under Markovian policies. Using this result, one has that the asymptotic of the term $\phi(z_j^n(t), M^n(t))$ is well-defined.

Note that the asymptotic exchangeability is a more general assumption in the sense that it does not require symmetry assumption as it is widely used the literature.

Using the above result, we have the mean field convergence of the process

$$\gamma_\ell\left(N_{0,\ell} + \frac{\alpha}{n-1}\sum_{i\neq j} z_{i,\ell}(t)\right)$$

to $m_\ell^*(t)$, where α term. For given mean field limit trajectory $\{m^*(t)\}_t$, the individual dynamic optimization problem reads

$$V_j(X_j, P_j, m^*) = \sup_{U_j}\mathbb{E}\left(\int_0^{+\infty} e^{-\beta t}\tilde{r}_j(X_j(t), P_j(t), U_j(t), m^*(t))\; dt \right.$$
$$\left. \mid X_j(0) = X_j, P_j(0) = P_j, m^*(0) = m^*\right)$$

such that

$$\left\{\begin{array}{c} \mathrm{d}X_{j,\ell}(t) = -(a_\ell X_{j,\ell}(t) + b_\ell)\mathrm{d}t + \sigma_\ell \mathrm{d}\mathbb{W}_{j,\ell},\;\; t\geq 0,\;\; \ell\in\mathcal{L} \\ \mathrm{d}p_{j,\ell} = u_{j,\ell}\mathrm{d}t,\;\; \ell\in\mathcal{L} \end{array}\right.$$

where $\tilde{r}_j = -\sum_{\ell\in\mathcal{L}}\left(|z_{j,\ell}(t) - m_\ell^*(t)|^2 + r_\ell u_{j,\ell}^2(t)\right)$, with $\mathbb{E}(z_{j,\ell}^2(0)) < \infty,\;\; z_{j,\ell}(t) = e^{X_{j,\ell}(t)}p_{j,\ell}(t)$.

Digital Object Identifier: 10.4108/ICST.CROWNCOM2010.9219
http://dx.doi.org/10.4108/ICST.CROWNCOM2010.9219

As in [9], for the single link case, the optimal control is determined by

$$U_{*,j}(t) = -\frac{b}{r}\left(\bar{K} z_j(t) + \bar{S}(t)\right)$$

where $\bar{K}$ positive and $\bar{S}$ bounded and continuous solution of the system

$$\begin{cases} (\beta - 2a)\bar{K} = -\frac{b^2}{r}\bar{K}^2 + 1 \\ (\beta - a)\bar{S} = \frac{d}{dt}\bar{S} - m^*(t) - \frac{b^2}{r}\bar{K}\bar{S} \end{cases}$$

Note that the optimal power allocation of user j at time t depends only its individual state $X_j(t)$, and the function $\bar{S}$ which depends on the mean field limit $m^*(t)$.

B. The asymptotic price of anarchy

It is natural to ask the gap between the global optimum payoff and the worse payoff obtained at the decentralized dynamic power allocation solutions. This gap is known as *price of anarchy* or inefficiency of anarchy [11]. In the context of mean field games, the sum of individual payoffs at the worse equilibrium can be infinite. Hence, we need to define the expected payoff of the infinite population or to consider the limiting of the price of anarchy with n users in which the global social welfare is normalized.

We define the *mean field price of anarchy* (APoA) as follows:

$$\text{APoA} = \parallel \lim_n \ \frac{1}{n}\sum_{j=1}^{n} \tilde{r}_j^n(GO) - \tilde{r}_j(U_*, m^*(t)) \parallel \qquad (17)$$

which measures the gap between the limit of global optimum payoff per individual $\lim_n \ \frac{1}{n}\sum_{j=1}^{n} \tilde{r}_j^n(GO)$ (the $\liminf$ is taken when the limit does not exists) and the payoff at the mean field equilibrium. The reason to consider the global optimum payoff per individual instead of the global optimum payoff is that the term $\sum_{j=1}^{n} \tilde{r}_j^n(GO)$ can be infinite when n grows to infinity.

Proposition IV-B1. *The mean field price of anarchy is in order of $O(\alpha \frac{L}{\beta} \max_{\ell \in \mathcal{L}} \theta_\ell)$ where θ_ℓ is the Lipschitz constant of ϕ_ℓ.*

The proof of this result follows from the mean field convergence and the fact that the integrand is Lipschitz continuous with Lipschitz constant less than $\alpha L \max_\ell \gamma_\ell$.

C. Remark

Gaussian distribution are widely used as channel state law. One of the reasons of that choice is that the Gaussian is the worse case for the error estimation, namely the quadratic gap between $s(t)$ and $y(t)$. Once we move from this objective by choosing for example $\tilde{r}_j^n(X^n(t), P^n(t), U^n(t))$, the Gaussian distribution will not be a minimizer. To model this case, we introduce *Nature* as a player who choose the distribution of the individual state (ID) at each time with a fixed expectation and fixed variance. It is now clear that Nature need not to follow a Gaussian distribution. A (mixed) strategy of Nature is to choose a distribution among all the possible distributions with fixed first two moments. In order to compute at the worse case for the users, we say that Nature minimizes the users payoff among all the distributions with zero expectation and variance equal to Γ. In this power allocation problem, this change does not affect the optimal strategy of the users because the other moments of the random variable does not affect the Hamilton-Jacobi-Bellman-Fleming equations. In particular, our result holds for time-varying non-Gaussian state evolution.

V. Concluding remarks

In this paper we have proposed a novel approach to the problem of joint power control-allocation in green cognitive networks. The proposed approach allows the network designer to design policies which are energy-efficient from the power supply point of view and not from the radio-frequency part point of view. We have shown how to use tools from control theory to find such policies. Additionally we have shown how to exploit mean field theory to approximate the SINR and facilitate the search of selfish optimal policies. The gap between the latter and the centralized policies is assessed by introducing and evaluating the asymptotic price of anarchy. An interesting question that we leave for future works is the computation of Nash equilibria in more general class of strategies and an efficient learning scheme for the mean field limit.

References

[1] S. Lasaulce, M. Debbah and E. Altman, Methodologies for analyzing equilibria in wireless games, IEEE Signal Processing Magazine, 26,5, 41–52, Sept. 2009.
[2] I. Gikhman and A. V. Skorohod, Stochastic Differential Equations, Springer-Verlag, Berlin, 1972.
[3] J. M. Lasry and P. L. Lions, *Mean field games*, Jpn. J. Math., vol. 2 (1), 2007, pp 229-260.
[4] D. W. K. Yeung and L. A. Petrosyan, Cooperative stochastic differential games, Springer-Verlag, New York, 2006.
[5] T. Basar, G. Jan Olsder, Dynamic Noncooperative Game Theory, 2nd edition, Classics in Applied Mathematics, SIAM, Philadelphia, 1999.
[6] M. Debbah, S. Lasaulce, M. Le Treust, H. Tembine, Dynamic Power Control in Wireless Communications, in Proc. ROADEF 2010.
[7] Y. Tanabe, The propagation of chaos for interacting individuals in a large population, Mathematical Social Sciences Volume 51, Issue 2, March 2006, Pages 125-152.
[8] P. M. Kotelenez, T. G. Kurtz, *Macroscopic limits for stochastic partial differential equations of McKean-Vlasov type*, Quart. Appl. Math. 66, 539-564, 2008.
[9] M. Y. Huang, P. E. Caines and R. P. Malhamé, "Large Population Cost-Coupled LQG Problem with non-uniform Agents: Individual-Mass Behaviour and Decentralized $\epsilon-$Nash equilibria, IEEE Trans. on Automatic Control, Vol. 52, No. 9, Sept. 2007, pp 1560-1571.
[10] H. Tembine, J. Y. Le Boudec, R. ElAzouzi and E. Altman, Mean Field Asymptotic of Markov Decision Evolutionary Games and Teams, in the Proc. of GameNets, May 2009.
[11] North D., Structure anc change in economic history, New York, 1981.
[12] W. H. Fleming and H. M. Soner, Controlled Markov Processes and Viscosity Solutions, Springer Verlag, 1992.
[13] Olama M. M., Djouadi S. M., Charalambous C. D. , Stochastic Power Control for Time-Varying Long-Term Fading Wireless Networks, EURASIP Journal on Applied Signal Processing, 2006, ID 89864, pp. 1-13.
[14] J. Palicot, Cognitive radio: an enabling technology for the green radio communications concept, proc. International Conference on Wireless Communications and Mobile Computing, 2009.
[15] H. Tembine, J. Y. Le Boudec, R. ElAouzi, E. Altman, From mean field interaction to evolutionary game dynamics,Proc. 7th international conference on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks, 2009.

Digital Object Identifier: 10.4108/ICST.CROWNCOM2010.9219
http://dx.doi.org/10.4108/ICST.CROWNCOM2010.9219