More about Base Station Location Games

François Mériaux, Samson Lasaulce
L2S - CNRS - SUPELEC - Univ Paris-Sud
3 rue Joliot-Curie
F-91192 Gif-sur-Yvette, France
{meriaux,lasaulce}@lss.supelec.fr

Michel Kieffer
L2S - CNRS - SUPELEC - Univ Paris-Sud
on leave at LTCI - CNRS - Telecom ParisTech
46 rue Barrault
F-75013 Paris, France
kieffer@lss.supelec.fr.com

ABSTRACT

This paper addresses the problem of locating base stations in a certain area which is highly populated by mobile stations; each mobile station is assumed to select the closest base station. Base stations are modeled by players who choose their best location for maximizing their uplink throughput. The approach of this paper is to make some simplifying assumptions in order to get interpretable analytical results and insights to the problem under study. Specifically, a relatively complete Nash equilibrium (NE) analysis is conducted (existence, uniqueness, determination, and efficiency). Then, assuming that the base station location can be adjusted dynamically, the best-response dynamics and reinforcement learning algorithm are applied, discussed, and illustrated through numerical results.

Categories and Subject Descriptors

H.4 [**Information Systems Applications**]: Miscellaneous; D.2.8 [**Software Engineering**]: Metrics—*complexity measures, performance measures*

Keywords

Game Theory, Location games, Wireless communications

1. INTRODUCTION

Mobile terminals (MTs) are currently gaining increased autonomy of decision to allow a better use of the available wireless resources. For example, MTs may choose their wireless access technology or the base station (BS) or access point to which they want to connect. We could imagine that this may be done in the future independently of the network operator owner of the BS. A mobile operator deploying BSs for a wireless network will have to deal with these new characteristics. If his goal is to maximize the traffic gathered by his own BSs, he will have to take into account the presence of competitor network operators when deciding on the location of BSs. If every operator involved has the same reasoning, this problem of BSs placement may be cast in the framework

of game theory and more precisely in the context of location games.

The history of location games starts with the work of Hotelling [8] in which the notion of spatial competition in a duopoly situation is introduced. More precisely, two firms compete for benefits over a finite segment crowded with customers. This results in the partition of the segment into a convex area of influence for each firm. Plastria [12] gives an overview of optimization approaches to place new facilities in an environment with pre-existing facilities. A large overview on location games is also presented in [7]. Location games are extended to the context of wireless networks with works such as [1] and [2]. The main difference arising in this new context is the interaction between MTs due to the mutual interference they generate. This point makes the association problem between MTs and BSs complex. As an association between a MT and a BS depends on SINR, the association relies on the respective locations of the MT and the BS, but also on the MTs already connected to the BS. Whereas [2] focuses on the downlink case, in [1] the location of BSs and the association choice of the MTs is treated as a Stackelberg game [15] in the uplink case. The context of our work is similar to the one in [1] but several interesting results are obtained in the present paper. The main *contributions* of this paper can be summarized as follows:

- As in [1] MTs are assumed to operate in the uplink and to be distributed along a one-dimensional region. However, each MT is assumed to select the closest BS (e.g., based on measure given by a GPS -global positioning system- receiver). This leads to a convenient form for the BSs utility functions (Sec. 2.1). As a consequence, the existence of a pure Nash equilibrium can be made rigorously (Sec. 3.1).

- Due to the symmetry of the problem, multiple Nash equilibria generally exist. However, if the locations can be ordered (which is easy for one-dimension regions), the Nash equilibrium can be determined and checked to be unique (Sec. 3.3).

- By making the reasonable assumption that the BS heights are much less than the typical distance between the BSs, the game can be further simplified and shown to be a form of Cournot oligopoly [4].

- In the two-player case, the efficiency of the Nash equilibrium is studied by evaluating the price of anarchy

- Assuming that the BS locations can be adjusted dynamically (which would be relevant in scenario like the one of small cells where only some of the small BS have to be active), the best response dynamics and reinforcement learning algorithm [3][14][11] are performed (Sec. 5).

- The made assumptions lead to several interpretations which could be further analyzed in the light of a more general framework (e.g., in two-dimensional regions).

The remainder of the paper is organized as follows. Section 2 introduces the physical model and the parameters of the K-player game. Section 3 describes the Nash equilibrium of the game in the one-dimensional case. In Section 4, the Nash equilibrium, the Stackelberg equilibrium, and the social optimum are compared. Section 5 presents a way to reach equilibrium using best-response dynamics and reinforcement learning. Finally, Section 6 concludes this work.

2. MODEL

Consider a plane to which a frame $\mathbb{R}$ is attached. A MT X located in a position $x \in \mathbb{R}^2$ in this plane is linked with a BS X_1 of height ε situated in $x_1 \in \mathbb{R}^2$, see Figure 1. We define the Signal to Noise Ratio (SNR) and the Signal to Interference plus Noise Ration (SINR) of this MT

$$SNR_X = \frac{P_X . h_{X_1}(x)}{\sigma_{X_1}^2},\qquad(1)$$

$$SINR_X = \frac{P_X . h_{X_1}(x)}{\sigma_{X_1}^2 + I_{X_1(x)}},\qquad(2)$$

where P_X is the transmission power of the MT X, i.e. the level of power chosen by the MT to transmit its signal. $\sigma_{X_1}^2$ is the power of the channel noise, $I_{X_1}(x)$ is some interference term, $h_{X_1}(x)$ is the attenuation introduced by the uplink channel from X to X_1. Here, it is assumed that

$$h_{X_1}(x) = \left(|x - x_1|^2 + \varepsilon^2\right)^{-\frac{\alpha}{2}},\qquad(3)$$

where $|x|$ is the ℓ_2-norm of x and $\alpha \geq 2$ is the path-loss exponent, $\alpha = 2$ corresponding to the free-space path-loss case. A higher value of α suits to worse channel conditions.

With Single User Decoding (SUD) at the BS X_1, there is no hierarchy for decoding the incoming signals at the BS. Hence, the signal from MT X is decoded by taking into account the full interference and the uplink capacity between X and X_1 may be written as

$$C_X = \log\left(1 + SINR_X\right),\qquad(4)$$

Without loss of generality, when several MTs are considered, it is assumed that P_X does not depend on the MT and is normalized, i.e., $P_X = 1$. Moreover, the channel conditions, described by (3), are the same for every MT and the noise power is constant $\sigma_{X_1} = \sigma$. With these assumptions, (4) becomes

$$C_X = \log\left(1 + \frac{\left(|x - x_1|^2 + \varepsilon^2\right)^{-\frac{\alpha}{2}}}{\sigma^2 + I_{X_1}(x)}\right).\qquad(5)$$
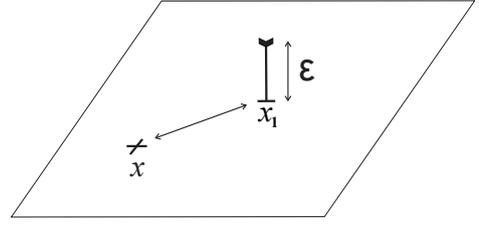

Figure 1: Base station located at a and mobile station located at x

When several BSs are located on the plane, the MTs are assumed to be able to choose the BS they want to be linked with. In this paper, we consider that this association is made based on SNR. Given (3), choosing the BS with the highest SNR is equivalent to choosing the closest BS. Thus, it is assumed that a MT always chooses the closest BS to its location, leading to convex cells for BSs. For example, considering two BSs X_1 and X_2 at positions x_1 and x_2 and a MT X at x, one has

$$C_X = \begin{cases} \log\left(1 + \frac{(|x-x_1|^2+\varepsilon^2)^{-\frac{\alpha}{2}}}{\sigma^2 + I_1(x)}\right) & \text{if } |x - x_1|^2 \leqslant |x - x_2|^2 \\ \log\left(1 + \frac{(|x-x_2|^2+\varepsilon^2)^{-\frac{\alpha}{2}}}{\sigma^2 + I_2(x)}\right) & \text{if } |x - x_2|^2 \leqslant |x - x_1|^2. \end{cases}$$

2.1 Base station utility

The utility of a BS is taken as the sum of uplink capacities it offers to its connected users. We assume that the number of MT is large enough to be represented by a continuous distribution $\rho(x)$. This assumption allows to get the utility for the k-th BS as a continuous sum of the MTs uplink capacities

$$U_k(\underline{x}) = \int_{S_k(\underline{x})} \rho(z) \log\left(1 + \frac{(|z - x_k|^2 + \varepsilon^2)^{-\frac{\alpha}{2}}}{\sigma^2 + I_k(\underline{x})}\right) \mathrm{d}z,\quad(6)$$

where $S_k(\underline{x})$ is the subset of the plane where MTs are linked with the k-th BS and $\underline{x} = {}^t(x_1, x_2, \ldots, x_K)$ is the vector of locations for the set of BSs. This paper will consider only uniform MT distribution with $\rho(x) = 1$.

When considering interferences, a worst-case scenario is considered: there is no mechanism such as beamforming [10] to lower their effects. We consider that there is no interference between MTs connected with different BSs because of frequency reuse. Then only interference between MTs of a same BS has to be considered. This framework is quite similar to the one of [1], where two competing BSs are assumed to use different frequency bands. In our case, we consider K BSs (with $K \geqslant 1$) and each of them uses its own frequency band.

Performing SUD at the i-th BS, one gets

$$I_k(\underline{x}) = \int_{S_k(\underline{x})} \left(|z - x_k|^2 + \varepsilon^2\right)^{-\frac{\alpha}{2}} \mathrm{d}z.\qquad(7)$$

Utility of k-th BS (6) then becomes

$$U_k(\underline{x}) = \int_{\mathcal{S}_k(\underline{x})} \log\left(1 + \frac{(|z - x_k|^2 + \varepsilon^2)^{-\frac{\alpha}{2}}}{\sigma^2 + \int_{\mathcal{S}_k(\underline{x})}(\varepsilon^2 + |z' - x_k|^2)^{-\frac{\alpha}{2}}\mathrm{d}z'}\right) \mathrm{d}z.$$
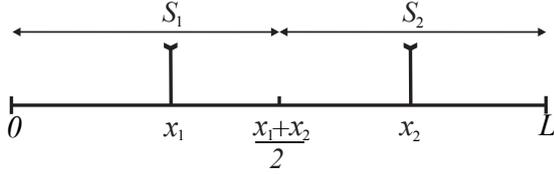
$$(8)$$

Figure 2: Two base stations competing in a one-dimension space

2.2 Utility approximation

When considering the low SINR regime, the useful power of each MT is small compared to the interference term in (8). This is especially true for high MT density. With this assumption, (8) may be approximated as

$$U_k(\underline{x}) \approx \frac{\int_{\mathcal{S}_k(\underline{x})}(|z - x_k|^2 + \varepsilon^2)^{-\frac{\alpha}{2}}\mathrm{d}z}{\sigma^2 + \int_{\mathcal{S}_k(\underline{x})}(|z' - x_k|^2 + \varepsilon^2)^{-\frac{\alpha}{2}}\mathrm{d}z'}$$
$$= \frac{I_k(\underline{x})}{\sigma^2 + I_k(\underline{x})}. \qquad (9)$$

Note that this simplification makes the considered utility based on capacity equivalent to a utility based on SINR such as the ones in [1]. At low-SINR regime, it is equivalent to work with a capacity-based utility or a SINR-based utility.

Also note that one has $f(t) = t/(\sigma^2 + t)$ strictly increasing over $[0, L]$, since its derivative is $f'(t) = \sigma^2/(\sigma^2 + t)^2$. Then maximizing the approximation of $U_k(\underline{x})$ or $I_k(\underline{x})$ is equivalent. Thus the utility we define for the game is $\widehat{U_k}(\underline{x}) = I_k(\underline{x})$, $k \in \mathcal{K}$.

2.3 Definition of the game

In this section, we study the case of K BSs competing on a segment of length L. Each of the BSs uses different carrier frequencies so the MTs of different BSs do not interfere together. As assumed in Section 2, the MT distribution is uniform over the segment, and the set of possible locations for the BS is $[0, L]$. Figure 2 illustrates the context for the two-player case.

Definition 1. The strategic form of the game is given by

$$\mathcal{G} = (\mathcal{K}, \{\mathcal{A}_k\}_{k \in \mathcal{K}}, \{\widehat{U_k}\}_{k \in \mathcal{K}})$$

where

- $\mathcal{K} = \{1, \dots, K\}$ is the set of players, which are here BS.

- $\{\mathcal{A}_k\}_{k \in \mathcal{K}}$ is the set of actions players can consider, here

$$\mathcal{A}_k = \{x_k \in [0, L] \mid 0 < x_1 < \dots < x_K < L\}. \quad (10)$$

Denote $\mathcal{A} = \{\underline{x} \in [0, L]^K \mid 0 < x_1 < \dots < x_K < L\}$.

- $\{\widehat{U_k}\}_{k \in \mathcal{K}}$ is the set of utilities players use.

Note that we are interested in location equilibria that do not superimpose several BSs. Thus, if there exists an equilibrium, there exists a spatial order for BSs at this equilibrium, that is why we introduce this order in the action spaces $\{\mathcal{A}_k\}_{k \in \mathcal{K}}$.

3. NASH EQUILIBRIUM ANALYSIS

The aim of this section is to show the existence of a Nash equilibrium in the location game described in Section 2.3 and to characterize this equilibrium.

3.1 Existence

In this section, we focus of the existence of a Nash equilibrium in the defined game. To prove the existence of a Nash equilibrium, the concavity of $\widehat{U_k}(\underline{x})$ with respect to x_k over $\mathcal{A}_k$, $\forall k \in \mathcal{K}$, has to be established.

LEMMA 1. $\widehat{U_k}(\underline{x})$ *is concave with respect to x_k over $\mathcal{A}_k$, $\forall k \in \mathcal{K}$.*

The proof of this lemma is in Appendix A.1. Then one has the following theorem.

THEOREM 1. *In the game defined by Definition 1, there exists at least one Nash equilibrium.*

PROOF.
- Using Lemma 1, we know that $\widehat{U_k}(\underline{x})$ is concave with respect to x_k over $\mathcal{A}_k$, $\forall k \in \mathcal{K}$,

- $\widehat{U_k}(\underline{x})$ is continuous with respect to $\underline{x}$ over $\mathcal{A}$, $\forall k \in \mathcal{K}$,

- the set of feasible actions is compact and convex for all players in the game.

The Rosen [13] conditions for the existence of a Nash equilibrium are met and Theorem 1 is thus proved. □

3.2 Multiplicity of NE

Regarding to the uniqueness of the Nash equilibrium, as the characteristics of the BSs (height and noise) are assumed to be identical, it is interesting to note that permuting the order of BSs leads to a symmetric system of equation. Thus, without condition on the order of BSs as in (10), there are $K!$ Nash equilibria for the game and all these equilibria are symmetric, meaning that the set of locations at equilibrium is unique. However, if one imposes the condition order (10), the NE can be shown to be unique by using the Diagonally Strict Concavity (DSC) condition [13].

THEOREM 2. *In the game defined by Definition 1, there exists one single Nash equilibrium.*

The proof of this theorem is in Appendix A.2. Having uniqueness under this order condition might seem to be a weak result in comparison with a general uniqueness result. However, in the framework of a dynamic process, the initial locations of the base station might suffice to determine the effectively observed NE (after convergence).
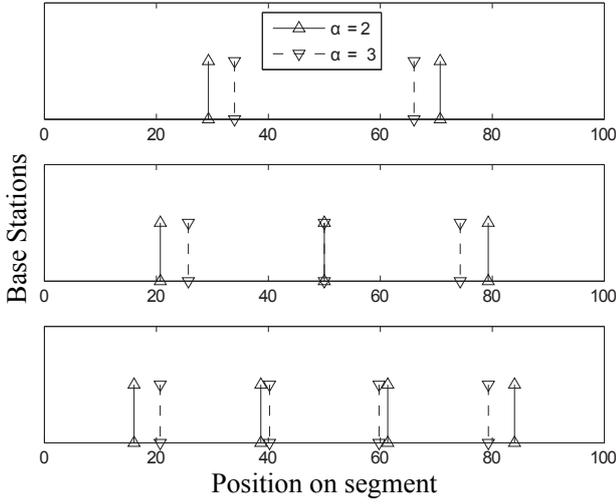
Figure 3: Base Stations at Nash equilibrium for $\alpha = 3$ (dotted line) and $\alpha = 2$ (continuous line). From the top to the bottom, configuration with 2 BSs, 3 BSs and 4 BSs. Parameters are $\varepsilon = 0.1$ and $L = 100$.

3.3 Determination of the NE

A characterization of the equilibrium is provided in this section with examples for small values of K. A real solution $\underline{x}$ for $\alpha > 2$ has to satisfy

$$
\begin{cases}
x_1 = \frac{2\sqrt{(5.2^{\frac{2}{\alpha}} - 2^{\frac{4}{\alpha}} - 4)\varepsilon^2 + 2^{\frac{2}{\alpha}}x_2^2} - 2^{\frac{2}{\alpha}}x_2}{4 - 2^{\frac{2}{\alpha}}}, \\
x_k = \frac{x_{k+1} + x_{k-1}}{2}, \ \forall k \in \{2, \ldots, K-1\}, \\
x_K = \frac{-2\sqrt{(5.2^{\frac{2}{\alpha}} - 2^{\frac{4}{\alpha}} - 4)\varepsilon^2 + 2^{\frac{2}{\alpha}}(L - x_{K-1})^2} - 2^{\frac{2}{\alpha}}x_{K-1} + 4L}{4 - 2^{\frac{2}{\alpha}}}.
\end{cases}
\tag{11}
$$

See Appendix A.3 for detailed derivations. In the case $\alpha = 2$, one obtains

$$
\begin{cases}
x_1 = \sqrt{2\varepsilon^2 + 2x_2^2} - x_2, \\
x_k = \frac{x_{k+1} + x_{k-1}}{2}, \ \forall i \in \{2, \ldots, K-1\}, \\
x_K = -\sqrt{10\varepsilon^2 + 2(L - x_{K-1})^2} - x_{K-1} + 2L.
\end{cases}
\tag{12}
$$

In the 2-player game with $\alpha = 2$, one gets

$$
\begin{cases}
x_1^{\mathrm{ne}} = L - \frac{1}{2}\sqrt{2L^2 - 4\varepsilon^2} \\
x_2^{\mathrm{ne}} = \frac{1}{2}\sqrt{2L^2 - 4\varepsilon^2}
\end{cases}
\tag{13}
$$

Figure 3 illustrates equilibria for $\alpha = 2$ and $\alpha = 3$ for respectively two, three, and four BSs.

4. COMPARING EQUILIBRIA

This section compares the Nash equilibrium described in Section 3, the Stackelberg equilibrium, and the social optimum.

4.1 Stackelberg equilibrium

As written in Section 3, the pure Nash equilibrium needs the players to know their spatial order to be reached. If players do not know this order but play in a chronological order, the problem changes. If the first player plays alone at the first stage of the game knowing that other players will
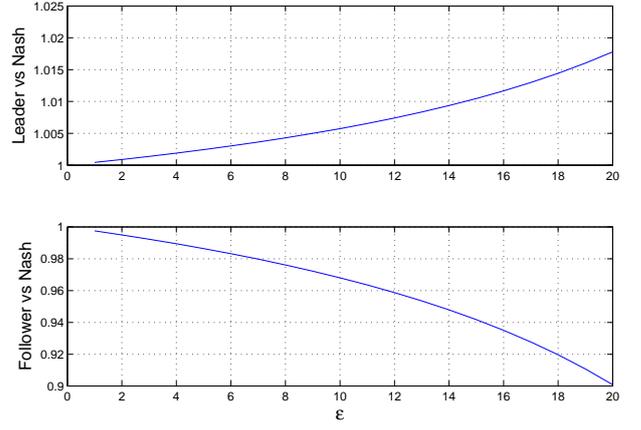

Figure 4: $\dfrac{\widehat{U_1}^{\mathbf{SE}}}{\widehat{U_1}^{\mathbf{NE}}}$ and $\dfrac{\widehat{U_2}^{\mathbf{SE}}}{\widehat{U_2}^{\mathbf{NE}}}$ with regard to ε.

place their BSs after, it turns into a Stackelberg game [15] with the first player being the leader of the game. This idea is illustrated with a two-player game with one leader (BS 1) and one follower (BS 2). The leader chooses its position knowing that the follower will place itself after. Both BSs still want to maximize their utilities and BS 1 knows this point. Then, BS 1 knows how BS 2 is going to be placed regarding to its own position, it is simply the best response of BS 2

$$
x_2(x_1) = \frac{-2\sqrt{(5.2^{\frac{2}{\alpha}} - 2^{\frac{4}{\alpha}} - 4)\varepsilon^2 + 2^{\frac{2}{\alpha}}(L - x_1)^2} - 2^{\frac{2}{\alpha}}x_1 + 4L}{4 - 2^{\frac{2}{\alpha}}}.
$$

Then BS 1 places itself at a location x_1 solution of

$$
\frac{\partial \widehat{U_1}}{\partial x_1}(x_1, x_2(x_1)) = 0.
\tag{14}
$$

With $\alpha = 2$ and neglecting ε compared to every other lengths of the problem, one gets

$$
x_1^{\mathrm{se}} = \left(1 - \sqrt{2} + \sqrt{2 - \sqrt{2}}\right)L,
\tag{15}
$$

and

$$
x_2^{\mathrm{se}} = \left((\sqrt{2} - 1)(1 + \sqrt{2 - \sqrt{2}})\right)L.
\tag{16}
$$

On Figure 4, we compare the utilities of the leader and the follower of the Stackelberg game versus the utilities of the Nash equilibrium. As we see, the leader of the Stackelberg game has a better utility than what he would have get at the Nash equilibrium. On the contrary, the follower has a worst utility. Hence, in a mobile operator point of view, it is more interesting to deploy its BS first.

4.2 Social optimum

Sections 3 and 4.1 provide equilibria corresponding to situations where the BSs only consider what is best for themselves. But these equilibria are not the best for the MTs. In
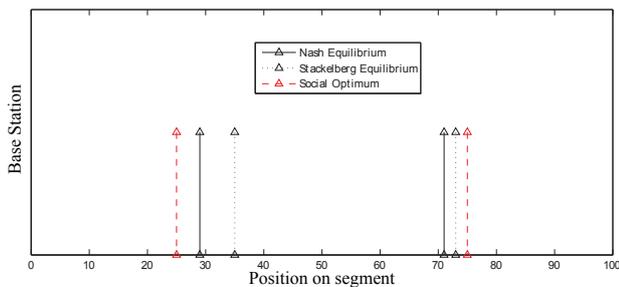
Figure 5: Nash Equilibrium, Stackelberg Equilibrium, and Social Optimum for $\alpha = 2$.

the MTs point-of-view, the utility to consider is the complete utility sum or *social utility*. In a 2-BS case

$$\widehat{U}^{\mathrm{so}}(x_1, x_2) = \widehat{U_1}(x_1, x_2) + \widehat{U_2}(x_1, x_2). \qquad (17)$$

For this utility, we know that there exists an optimum since the strategy set $\mathcal{A}$ is compact and $\widehat{U_1}(x_1, x_2) + \widehat{U_2}(x_1, x_2)$ is continuous with respect to (x_1, x_2) over $\mathcal{A}$. At this optimum, it is proven [1] that

- (i) the BSs place themselves at the middle of their associated subset of segment (respectively $\mathcal{S}_1$ and $\mathcal{S}_2$),

- (ii) the frontier between the two BSs is the middle of the segment.

Thus, we have the optimum

$$x_1^{\mathrm{so}} = \frac{L}{4}, \; x_2^{\mathrm{so}} = \frac{3L}{4}. \qquad (18)$$

Figure 5 shows the Nash equilibrium, the Stackelberg equilibrium, and the social optimum for $\alpha = 2$. The locations of BSs for Nash equilibrium and social optimum are symmetric with respect to the middle of the segment $[0, L]$ whereas this is not the case for the Stackelberg equilibrium.

4.3 Price of anarchy

To compare the Nash and the Stackelberg equilibrium in the two-player case, we look at the utilities at equilibrium. Without any pricing mechanism, the Nash equilibrium and the social optimum are very close in terms of locations. As a result, they are also close in terms of utility sum. The *Price of Anarchy* (PoA), introduced in [9], is an adequate metric to compare these sums

$$PoA(eq) = \frac{\max_{\underline{x} \in \mathcal{A}} \sum_{k \in \mathcal{K}} \widehat{U}_k(\underline{x})}{\sum_{k \in \mathcal{K}} \widehat{U}_k(\underline{x}^{\mathrm{eq}})}. \qquad (19)$$

Note that the PoA is always stronger than 1 and if the PoA is high, it means that the corresponding equilibrium is not that efficient in terms of utility sum. On the contrary, if the PoI is close to one, the corresponding equilibrium is satisfying. Figure 6 illustrates the behavior of the PoA for the Nash equilibrium and the PoA for the Stackelberg equilibrium as a function of ε for 2-player case. It appears that
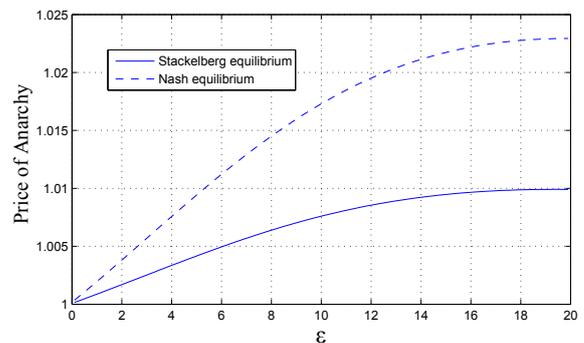

Figure 6: Prices of Anarchy regarding to ε. Parameters are $L = 100$, $\sigma^2 = 10^4$, $\alpha = 2$.

the Stackelberg equilibrium is less efficient than the Nash equilibrium.

5. CONVERGENCE TO THE NASH EQUILIBRIUM

Getting to the equilibrium given in Section 3.3 is not simple: it would require that every player knows the utilities of other players. In practice, this is hardly the case. Thus, we present techniques that enable to reach the Nash equilibrium in a decentralized way: best-response dynamics and reinforcement learning. However the BSs need to be movable to perform these two techniques, hence it is more accurate to talk about Mobile Stations (MSs) in the present section. Note that for these two techniques, the only assumption about MSs location is that they cannot superimpose.

5.1 Best response dynamics

The context of this section remains the same as defined in Section 2.3.

The principle of best-response dynamics is that given a realization of actions for its opponents, every player of the game is able to compare its own possible actions and choose which one is best for itself. Precisely the best-response algorithm is the following

1. At every time step t, each MS k chooses its location $x_k^{\mathrm{br}}(t)$ according to

$$x_k^{\mathrm{br}}(t) = \arg\max_{x_k} \hat{U}_k(x_k, \underline{x}_{-k}(t)).$$

2. Algorithm stops when $\left| \underline{x}^{\mathrm{br}}(t+1) - \underline{x}^{\mathrm{br}}(t) \right| < \beta$, with β fixed.

Regarding to the convergence of this algorithm, we study the system of best-responses (11). If we neglect ε regarding to x_1 and $L - x_{K-1}$, one gets a linear system of equations

$$\begin{cases} x_1 = \frac{2^{\frac{1+\alpha}{\alpha}} - 2^{\frac{2}{\alpha}}}{4 - 2^{\frac{2}{\alpha}}} x_2, \\ x_k = \frac{x_{k+1} + x_{k-1}}{2}, \; \forall k \in \{2, \dots, K-1\}, \qquad (20) \\ x_K = \frac{(4 - 2^{\frac{1+\alpha}{\alpha}})L}{4 - 2^{\frac{2}{\alpha}}} + \frac{2^{\frac{1+\alpha}{\alpha}} - 2^{\frac{2}{\alpha}}}{4 - 2^{\frac{2}{\alpha}}} x_{K-1}. \end{cases}$$

It is a linear Cournot oligopoly [4].

One distinguishes two ways this algorithm can run.

Simultaneous best response dynamics. At every step, every MS adapt their locations simultaneously. In this case, the evolution of the locations with the algorithm can be expressed by

$$\underline{x}^{\mathrm{br}}(t+1) = \underline{a} + \mathrm{A}_K \underline{x}^{\mathrm{br}}(t) \tag{21}$$

with $\underline{a} = {}^t\left(0,\dots,0,\frac{(4-2^{\frac{1+\alpha}{\alpha}})L}{4-2^{\frac{2}{\alpha}}}\right)$ and

$$\mathrm{A}_K = \begin{pmatrix} 0 & g(\alpha) & 0 & \dots\dots\dots\dots & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} & \ddots & & & \vdots \\ 0 & \ddots & \ddots & \ddots & \ddots & & \vdots \\ \vdots & \ddots & \ddots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & \ddots & \ddots & 0 \\ \vdots & & & \ddots & \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & \dots\dots\dots\dots & 0 & g(\alpha) & 0 \end{pmatrix}, \tag{22}$$

with $g(\alpha) = \frac{2^{\frac{1+\alpha}{\alpha}}-2^{\frac{2}{\alpha}}}{4-2^{\frac{2}{\alpha}}}$.

As recalled in Lemma 2.8 of [16], for an irreducible positive square matrix $\mathrm{A} = (a_{ij})_{1\le i\le n}$, then either

$$\sum_{j=1}^{n} a_{ij} = \rho(\mathrm{A}) \; \forall i \in [1,n], \tag{23}$$

or

$$\inf_{1\le i\le n} \sum_{j=1}^{n} a_{ij} < \rho(\mathrm{A}) < \sup_{1\le i\le K} \sum_{j=1}^{n} a_{ij}, \tag{24}$$

with $\rho(\mathrm{A})$ being the radius of A. In our case, A_K can be verified to be positive and irreducible, then one has

$$g(\alpha) < \rho(\mathrm{A}_K) < 1. \tag{25}$$

Hence, the convergence of the algorithm is ensured.

Sequential best response dynamics. The other case is the sequential algorithm: at every step, only one MS adapts its location. Depending on the MS k adapting its location, the algorithm evolves according to

$$\underline{x}^{\mathrm{br}}(t+1) = \underline{a} + \mathrm{A}_K^k \underline{x}^{\mathrm{br}}(t), \tag{26}$$

with A_K^k corresponding to the adaptation of the k-th MS. A_K^k may have several forms.

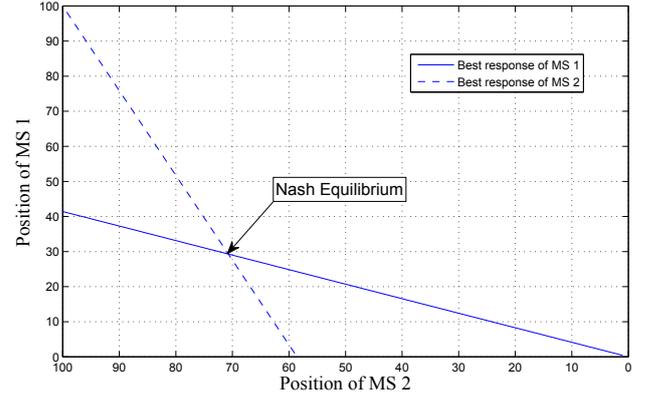- If the k-th MS, $k \in \{2,\dots,K-1\}$ adapts its location,

Figure 7: Best-responses of MS 1 and MS 2 for $\alpha = 2$

then

$$\mathrm{A}_K^k = \begin{pmatrix} 1 & 0 & \dots\dots\dots\dots & 0 \\ \ddots & \ddots & \ddots & & & \vdots \\ \vdots & & 0 & 1 & 0 & & \vdots \\ \vdots & & & \frac{1}{2} & 0 & \frac{1}{2} & & \vdots \\ \vdots & & & & 0 & 1 & 0 & \vdots \\ \vdots & & & & & \ddots & \ddots & \ddots \\ 0 & \dots\dots\dots\dots & & & 0 & 1 \end{pmatrix}$$

In this case, A_K^k is irreducible and positive and

$$\rho(\mathrm{A}_K^k) = 1. \tag{27}$$

- If the first MS adapts its location, the matrix has the form

$$\mathrm{A}_K^1 = \begin{pmatrix} 0 & g(\alpha) & 0 & \dots \\ 0 & 1 & 0 & \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \dots & 0 & 1 \end{pmatrix}$$

Again, A_K^1 is irreducible and positive, but this time

$$g(\alpha) < \rho(\mathrm{A}_K^1) < 1. \tag{28}$$

Note that if the K-th MS adapts its location, the same reasoning can be done.

$$g(\alpha) < \rho(\mathrm{A}_K^K) < 1. \tag{29}$$

Hence, $\rho(\prod_{k=1}^{K} \mathrm{A}_K^k) = \prod_{k=1}^{K} \rho(\mathrm{A}_K^k) < 1$ and the algorithm converges.

In Figure 7, we compare MSs 1 and 2 sequential best-responses for $\alpha = 2$.

5.2 Reinforcement Learning

The following notations are used in this section. For a set of MSs $\mathcal{K} = \{1,\dots,K\}$, let $\tilde{\mathcal{A}}_k = \{y_{k1},\dots,y_{km_k}\}$ be the set of possible locations for MS k (m_k being the cardinality of $\tilde{\mathcal{A}}_k$), which corresponds to a discretization of the set $[0,L]$.

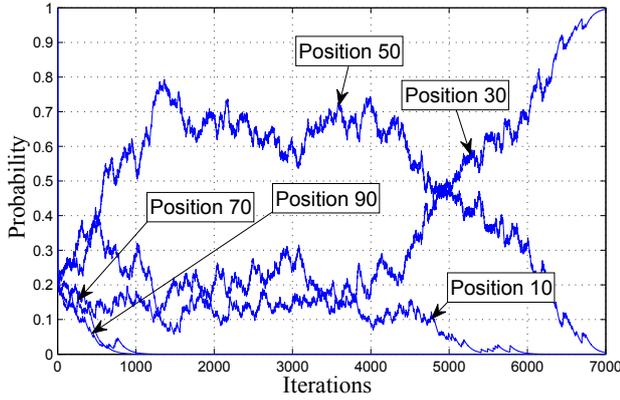Figure 8: Evolution of action probabilities for BS 1 with discrete stochastic learning.

Figure 9: Evolution of action probabilities for BS 2 with discrete stochastic learning.

We implement a discrete stochastic learning algorithm in the sense that the action space is discrete and at every step, actions are chosen in a stochastic way. The only information available to a MS is the value of its utility after each iteration (note that the MS does not necessarily knows its utility expression).

We define $\underline{p}_k(t) = (p_{k1}(t), \ldots, p_{km_k}(t))$ the probability distribution vector of MS k at time t.

$$\mathrm{P}\left[x_k(t) = y_{ki}\right] = p_{ki}(t),\ i \in \{1, \ldots, m_k\}. \qquad (30)$$

The algorithm used by each MS is then the following.

1. Initialize the distribution probability vector.
   $\forall k \in \mathcal{K},\ \forall i \in \{1, \ldots, m_k\},\ p_{ki}(0) = \frac{1}{m_k}$.

2. At every step t, each MS k chooses a location $x_k(t)$ according to its probability vector $\underline{p}_k(t)$.

3. Each MS gets $U_k(t)$.

4. Each MS updates its probability distribution vector $\underline{p}_k(t)$

   $p_{ki}(t+1) = p_{ki}(t) - bU_k(t)p_{ki}(t),$ if $x_t(t) \neq y_{ki}$
   $p_{ki}(t+1) = p_{ki}(t) + bU_k(t)\sum_{s \neq i} p_k s(t),$ if $x_k(t) = y_{ki}$
   (31)

5. Algorithm stops when $\underline{p}_k(t+1) = \underline{p}_k(t)$, else go to step 2.

Figures 8 and 9 illustrate the evolution of the probability distribution vectors of two MSs. The parameters of the simulation are the following: each MS has the same set of possible positions $\{10, 30, 50, 70, 90\}$ and the step of the learning algorithm is $b = 0.01$.

Depending on the choice of b, two phenomena occur.

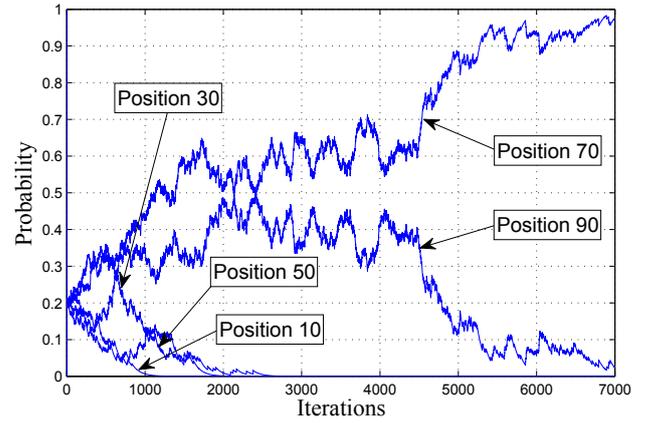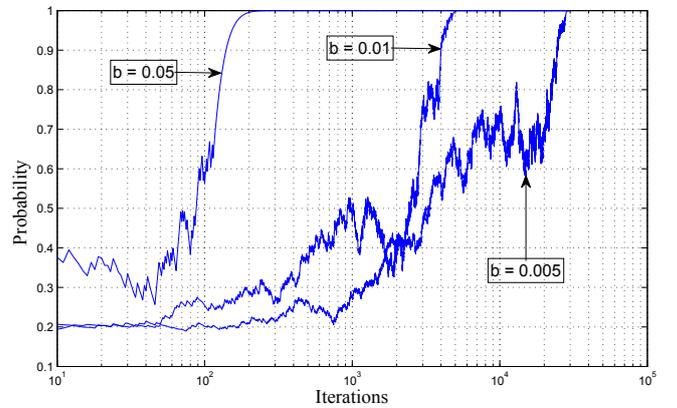- The higher the value of b, the lower the convergence time of the algorithm,

Figure 10: Convergence of probability with respect to step b

- However, if b is chosen too high, the algorithm may converge to locations that do no correspond to an equilibrium of the game defined in Section 2.3.

Figure 10 illustrates the convergence time as a function of b.

The choice of b is hence a trade-off between convergence time and accuracy of the convergence.

6. CONCLUSION AND PERSPECTIVES

Obviously the proposed model is relatively simple and should be improved to obtain refined results. However, this led us to several interesting results such as a full characterization of Nash equilibria and interesting behavior in terms of convergence. It would be very relevant to extend this work to two-dimensional scenario, define a suitable order for which uniqueness would be ensured. More connections with the famous multi-source Weber problem [6] should be established to better understand the general problem of *deployment games*. Indeed, if an operator has to locate a set of

base stations, the problem becomes more complicated. The problem becomes even more interesting if the set of possible constellations is discrete, which would lead us to make connections with Voronoi games [5]. The authors believe there is large avenue for contributions to the general problem under study, especially in finding relevant assumptions to simplify it without too loss in terms of understanding. Random matrix theory and stochastic geometry might be for great help to achieve this challenging objective.

7. REFERENCES

[1] Eitan Altman, Anurag Kumar, Chandramani Kishore Singh, and Rajesh Sundaresan. Spatial sinr games combining base station placement and mobile association. In *INFOCOM*, pages 1629–1637. IEEE, 2009.

[2] Eitan Altman, Alonso Silva, Amidou Tembine, and Merouane Debbah. Spatial games and global optimization for the mobile association problem: the downlink case. In *IEEE/SIAM CDC'*, 2010.

[3] Robert R. Bush and Frederick Mosteller. *Stochastic Models for Learning.* Wiley, 1st edition, 1955.

[4] A. Cournot. *Researches into the principles of wealth.* Irwin Paperback Classics in Economics (original: Recherches sur les principes mathématiques de la théorie des richesses, 1838), 1963 (English Translation).

[5] Christoph Dürr and Nguyen Kim Thang. Nash equilibria in Voronoi games on graphs. *CoRR*, abs/cs/0702054, 2007.

[6] Sándor P. Fekete, Joseph S. B. Mitchell, and Karin Beurer. On the continuous fermat-weber problem. *CoRR*, cs.CG/0310027, 2003.

[7] Jean J. Gabszewicz and Jacques-Francois Thisse. Location. In R.J. Aumann and S. Hart, editors, *Handbook of Game Theory with Economic Applications*, volume 1 of *Handbook of Game Theory with Economic Applications*, chapter 9, pages 281–304. Elsevier, 1992.

[8] Harold Hotelling. Stability in competition. *The Economic Journal*, 39(153):41–57, 1929.

[9] Elias Koutsoupias and Christos Papadimitriou. Worst-case equilibria. In *the 16th Annual Symposium on Theoretical Aspects of Computer Science*, pages 404–413, 1999.

[10] John Litva and Titus K. Lo. *Digital Beamforming in Wireless Communications.* Artech House, Inc., Norwood, MA, USA, 1996.

[11] Kumpati S. Narendra and Mandayam A. L. Thathachar. *Learning automata: an introduction.* Prentice-Hall, Inc., Upper Saddle River, NJ, USA, 1989.

[12] Frank Plastria. Static competitive facility location: An overview of optimisation approaches. *European Journal of Operational Research*, 129(3):461–470, March 2001.

[13] J. B. Rosen. Existence and Uniqueness of Equilibrium Points for Concave N-Person Games. *Econometrica*, 33(3):520–534, 1965.

[14] P.S. Sastry, V.V. Phansalkar, and M.A.L. Thathachar. Decentralized learning of nash equilibria in multi-person stochastic games with incomplete information. *Systems, Man and Cybernetics, IEEE Transactions on*, 24(5):769 –777, May 1994.

[15] H. Stackelberg. *Marketform und Gleichgewicht.* Oxford, U.K., 1934.

[16] Richard S. Varga. *Matrix iterative analysis.* Springer series in computational mathematics, 27. Springer, 2nd rev. and expanded ed. edition, 2000.

APPENDIX

A. PROOFS

A.1 Proof of Lemma 1

We need to prove that $\widehat{U_k}(\underline{x})$ is concave with respect to x_k over $\mathcal{A}_k$.

We know that the index of players verify (10). In this case, the regions associated to the BSs are

$$
\begin{cases}
[0, \frac{x_1+x_2}{2}] \text{ for BS 1,} \\
[\frac{x_{k-1}+x_k}{2}, \frac{x_k+x_{k+1}}{2}] \text{ for BS } k, \ k \in \{2,\ldots,K-1\} \\
[\frac{x_{K-1}+x_K}{2}, L] \text{ for BS } K.
\end{cases} \quad (32)
$$

Then

$$
\begin{cases}
\widehat{U_1}(\underline{x}) = \int_0^{\frac{x_1+x_2}{2}} (\varepsilon^2 + (x-x_1)^2)^{-\frac{\alpha}{2}} \, dx \\
\widehat{U_k}(\underline{x}) = \int_{\frac{x_{k-1}+x_k}{2}}^{\frac{x_k+x_{k+1}}{2}} (\varepsilon^2 + (x-x_k)^2)^{-\frac{\alpha}{2}} \, dx, \ k \in \{2,\ldots,K-1\} \\
\widehat{U_K}(\underline{x}) = \int_{\frac{x_{K-1}+x_K}{2}}^{L} (\varepsilon^2 + (x-x_K)^2)^{-\frac{\alpha}{2}} \, dx
\end{cases} \quad (33)
$$

which may be rewritten as

$$
\begin{cases}
\widehat{U_1}(\underline{x}) = \int_{-x_1}^{\frac{x_2-x_1}{2}} (\varepsilon^2 + x^2)^{-\frac{\alpha}{2}} \, dx \\
\widehat{U_k}(\underline{x}) = \int_{\frac{x_{k-1}-x_k}{2}}^{\frac{x_{k+1}-x_k}{2}} (\varepsilon^2 + x^2)^{-\frac{\alpha}{2}} \, dx, \ k \in \{2,\ldots,K-1\} \\
\widehat{U_K}(\underline{x}) = \int_{\frac{x_{K-1}-x_K}{2}}^{L-x_K} (\varepsilon^2 + x^2)^{-\frac{\alpha}{2}} \, dx
\end{cases} \quad (34)
$$

To prove the existence of a Nash equilibrium, the concavity of $\widehat{U_k}(\underline{x})$ with respect to x_k $\forall k \in \mathcal{K}$ has now to be established. One has the first-order partial derivatives

$$
\begin{cases}
\frac{\partial \widehat{U_1}}{\partial x_1}(\underline{x}) = -\frac{1}{2}(\varepsilon^2 + (\frac{x_2-x_1}{2})^2)^{-\frac{\alpha}{2}} + (\varepsilon^2 + x_1^2)^{-\frac{\alpha}{2}} \\
\frac{\partial \widehat{U_k}}{\partial x_k}(\underline{x}) = -\frac{1}{2}(\varepsilon^2 + (\frac{x_{k+1}-x_k}{2})^2)^{-\frac{\alpha}{2}} + \frac{1}{2}(\varepsilon^2 + (\frac{x_{k-1}-x_k}{2})^2)^{-\frac{\alpha}{2}} \\
\frac{\partial \widehat{U_K}}{\partial x_K}(\underline{x}) = -(\varepsilon^2 + (L-x_K)^2)^{-\frac{\alpha}{2}} + \frac{1}{2}(\varepsilon^2 + (\frac{x_{K-1}-x_K}{2})^2)^{-\frac{\alpha}{2}}
\end{cases} \quad (35)
$$

and the second-order partial derivatives

$$
\begin{cases}
\frac{\partial^2 \widehat{U_1}}{\partial x_1^2}(\underline{x}) = -\frac{\alpha}{8} \frac{x_2-x_1}{(\varepsilon^2+(\frac{x_2-x_1}{2})^2)^{\frac{\alpha}{2}+1}} - \alpha \frac{x_1}{(\varepsilon^2+x_1^2)^{\frac{\alpha}{2}+1}} \\
\frac{\partial^2 \widehat{U_k}}{\partial x_k^2}(\underline{x}) = \frac{\alpha}{8}\left(-\frac{x_{k+1}-x_k}{(\varepsilon^2+(\frac{x_{k+1}-x_k}{2})^2)^{\frac{\alpha}{2}+1}} + \frac{x_{k-1}-x_k}{(\varepsilon^2+(\frac{x_{k-1}-x_k}{2})^2)^{\frac{\alpha}{2}+1}}\right) \\
\frac{\partial^2 \widehat{U_K}}{\partial x_K^2}(\underline{x}) = -\alpha \frac{L-x_K}{(\varepsilon^2+(L-x_K)^2)^{\frac{\alpha}{2}+1}} + \frac{\alpha}{8} \frac{x_{K-1}-x_K}{(\varepsilon^2+(\frac{x_{K-1}-x_K}{2})^2)^{\frac{\alpha}{2}+1}}
\end{cases} \quad (36)
$$

Given (10), we have $\frac{\partial^2 \widehat{U_k}}{\partial x_k^2}(\underline{x}) < 0$ $\forall k \in \mathcal{K}$. Thus $\widehat{U_k}(\underline{x})$ is concave with respect to x_k over $\mathcal{A}_k$, $\forall k \in \mathcal{K}$.

A.2 Proof of Theorem 2

In the context of K-player game, the DSC [13] condition writes $\forall(\underline{a}, \underline{a}') \in \mathcal{A}^2$ such that $\underline{a} \neq \underline{a}'$

$$
\sum_{k=1}^{K} (a'_k - a_k) \left(\frac{\partial \widehat{U_k}}{\partial x_k}(\underline{a}) - \frac{\partial \widehat{U_k}}{\partial x_k}(\underline{a}') \right) > 0 \quad (37)
$$

For clarity reasons, we denote

$$
g(a,b) = \left((\epsilon^2 + a^2)^{-\frac{\alpha}{2}} - (\epsilon^2 + b^2)^{-\frac{\alpha}{2}} \right), \ (a,b) \in \mathbb{R}^2. \quad (38)
$$

By (35), it turns

$$
\begin{cases}
\frac{\partial \widehat{U_1}}{\partial x_1}(\underline{a}) - \frac{\partial \widehat{U_1}}{\partial x_1}(\underline{a}') = g(a_1, a'_1) - \frac{1}{2}\left(g(\frac{a_1-a_2}{2}, \frac{a'_1-a'_2}{2}) \right) \\
\frac{\partial \widehat{U_k}}{\partial x_k}(\underline{a}) - \frac{\partial \widehat{U_k}}{\partial x_k}(\underline{a}') = \frac{1}{2}\left(g(\frac{a_{k-1}-a_k}{2}, \frac{a'_{k-1}-a'_k}{2}) - g(\frac{a_k-a_{k+1}}{2}, \frac{a'_k-a'_{k+1}}{2}) \right) \\
\frac{\partial \widehat{U_K}}{\partial x_K}(\underline{a}) - \frac{\partial \widehat{U_K}}{\partial x_K}(\underline{a}') = \frac{1}{2}\left(g(\frac{a_{K-1}-a_K}{2}, \frac{a'_{K-1}-a'_K}{2}) \right) - g(L-a_K, L-a'_K)
\end{cases} \quad (39)
$$

Equation (37) becomes

$$
\begin{aligned}
&(a'_1-a_1)\left(g(a_1,a'_1) - \frac{1}{2}g(\frac{a_1-a_2}{2}, \frac{a'_1-a'_2}{2}) \right) \\
&+ \sum_{k=2}^{K-1} \frac{a'_k - a_k}{2}\left(g(\frac{a_{k-1}-a_i}{2}, \frac{a'_{k-1}-a'_k}{2}) - g(\frac{a_k-a_{k+1}}{2}, \frac{a'_k-a'_{k+1}}{2}) \right) \\
&+ (a'_K - a_K)\left(\frac{1}{2}g(\frac{a_{K-1}-a_K}{2}, \frac{a'_{K-1}-a'_K}{2}) - g(L-a_K, L-a'_K) \right) > 0
\end{aligned} \quad (40)
$$

which can also be written

$$
\begin{aligned}
&(a'_1 - a_1)g(a_1, a'_1) \\
&+ \sum_{k=2}^{K} \frac{a'_k - a'_{k-1} - (a_k - a_{k-1})}{2} g\left(\frac{a_{k-1}-a_k}{2}, \frac{a'_{k-1}-a'_k}{2} \right) \\
&+ (L - a'_K - (L - a_K))g(L-a_K, L-a'_K) > 0
\end{aligned} \quad (41)
$$

However $\forall(a,b) \in \mathbb{R}^{*2}$,

$$
(b-a)g(a,b) > 0 \quad (42)
$$

and by the order condition (10)

$$
\begin{cases}
(a'_1, a_1) \in \mathbb{R}^{*2} \\
(\frac{a'_k - a'_{k-1}}{2}, \frac{a_k - a_{k-1}}{2}) \in \mathbb{R}^{*2} \ \forall k \in \{2,\ldots,K\} \\
(L - a'_K, L - a_K) \in \mathbb{R}^{*2}
\end{cases} \quad (43)
$$

Then the DSC condition is verified and the equilibrium is unique.

A.3 Derivation of (11)

To obtain a formal expression of a Nash equilibrium, the intersection of the best-responses has to be considered

$$
\frac{\partial \hat{U}_k}{\partial x_k}(\underline{x}) = 0, \ \forall k \in \mathcal{K}, \ \forall \alpha \geqslant 2. \quad (44)
$$

(44) leads to

$$
\begin{cases}
(2^{\frac{2}{\alpha}}(\varepsilon^2 + (\frac{x_2-x_1}{2})^2))^{\frac{\alpha}{2}} = (\varepsilon^2 + x_1^2)^{\frac{\alpha}{2}} \\
(\varepsilon^2 + (\frac{x_{k+1}-x_k}{2})^2)^{\frac{\alpha}{2}} = (\varepsilon^2 + (\frac{x_{k-1}-x_k}{2})^2)^{\frac{\alpha}{2}}, \ \forall k \in \{2,\ldots,K-1\} \\
(\varepsilon^2 + (L - x_K)^2)^{\frac{\alpha}{2}} = (2^{\frac{2}{\alpha}}(\varepsilon^2 + (\frac{x_{K-1}-x_K}{2})^2))^{\frac{\alpha}{2}}.
\end{cases} \quad (45)
$$

Since all terms elevated at power $\alpha/2$ are positive, one gets

$$
\begin{cases}
2^{\frac{2}{\alpha}}(\varepsilon^2 + (\frac{x_2-x_1}{2})^2) = \varepsilon^2 + x_1^2 \\
(\frac{x_{i+k}-x_k}{2})^2 = (\frac{x_{k-1}-x_k}{2})^2, \ \forall i \in \{2,\ldots,K-1\} \\
2^{\frac{2}{\alpha}}(\varepsilon^2 + (\frac{x_{K-1}-x_K}{2})^2) = \varepsilon^2 + (L-x_K)^2
\end{cases} \quad (46)
$$

One real solution $\underline{x}$ verifying (10) has thus to satisfy (11) when $\alpha > 2$ and (12) when $\alpha = 2$.