# A new approximation method for multi-level programming and application to optimal pricing of communication networks

M. Bergounioux & M. Haddou
MAPMO-UMR 6628
Université d'orléans - BP 6759
45067 Orléans cedex 2
maitine, haddou@labomath.univ-orleans.fr

June 15, 2004

**Abstract**

We present a penalized method to approach multi-level problems occurring in optimal pricing of communication networks. This allows to overcome the difficulty arising from the non uniqueness of different level problems solutions. We prove existence of approximated solutions, give convergence result with Hoffman-like assumptions. We end with cost value error estimates.

## 1 Introduction

Price and revenue optimization in the areas of transport and telecommunications networks is a very active domain of research and has received a considerable attention in the recent literature. See, for example, [5, 6, 8, 12, 14]

The principal motivation of this work, is the theoretical study of a new regularization technique for solving the short time optimal pricing of links in packet switched communication networks.

This problem belongs to a large class of important and strategic economic problems that can be viewed as some asymmetric games. The leader (a telecommunication company) have to price the arcs of its subnetwork in order to maximize some revenue function. This revenue function is in general a part of the cost function that minimizes the strategy of the follower (the user). The leader must take into account the reaction (or any possible reaction when non unique) of the follower. This problem, in its best case version (i.e. if multiple optimal reaction of the follower are possible, the follower is supposed to choose in favor of the leader) can be modelled using a bi-level program. We consider that such program is ill posed when the lower level ( user strategies) can have multiple solutions for each (or some) fixed leader's variables.

These programs are very difficult nonconvex optimization problems. Several heuristics or approximation techniques can be found in the recent literature [5, 6, 8, 12, 14] . In almost all these works, strong assumptions are made to simplify the model. The solutions of the lower level are supposed unique or at least to correspond to the same upper level revenue.

The lower level is replaced by the equivalent optimality conditions and considered as an equilibrium problem. The complementarity part of these optimality condition is then smoothed or penalized using different techniques.
In our approach we consider the realistic situation with different possible reactions of the follower and multiple revenues. We first present the problem of optimal pricing of links communication networks (O.P.L.C.N) and then consider general ill posed bi-level programs.

## 1.1 The O.P.L.C.N model

Let $G = (V, E)$ be a directed graph, where $V$ is a set of $p$ nodes , $E$ is a set of $n$ arcs ($n = n_1 + n_2$, the first $n_1$ arcs are those to be priced ) and $\mathcal{K}$ a set of $K$ commodities to be routed through the network represented by $G$.
Each commodity $k$ has a unique source $o^k$ and a unique sink $p^k$. We denote by $A$ the node-arc incidence matrix of $G$. $A$ is an $p \times n$ matrix such that each column is related to an arc $e \in E$ and has only two nonzero $(-1, 1)$ coefficients in those rows associated with (respectively ) the origin and the destination node of $e$. We will decompose $A := [A_1 A_2]$ where the columns of $A_1$ correspond to the first $n_1$ arcs.
The arcs of the network have a maximal capacity $c \in \mathbb{R}^n$ for all commodities. W denote by $b^k \in \mathbb{R}^p$, for $k \in \mathcal{K}$, a vector of supplies/demands for commodity $k$ at the nodes of the network

$$b_i^k = \begin{cases} +r^k & \text{if } i = o^k, \\ -r^k & \text{if } i = p^k, \\ 0 & \text{elsewhere}, \end{cases}$$

where $r^k$ is the value of the flow of the $k^{th}$ commodity.
Each commodity will be satisfied if there exists a vector $(x_1^k, x_2^k)^T \in \mathbb{R}^{n_1+n_2}$ such that

$$A_1 x_1^k + A_2 x_2^k = b^k.$$

The capacity constraints have to be satisfied by the contributions of the all $K$ commodities

$$\sum_{k \in \mathcal{K}} (x_1^k, x_2^k)^T \ \leq C.$$

We will denote $x_1 := (x_1^1, x_1^2, ..., x_1^K)^T$ and $x_2 := (x_2^1, x_2^2, ..., x_2^K)^T$.
The revenue function of the leader (to be maximized) is of the form

$$f_u(y, x_1).$$

( In the "linear" case, this function is $y^T \sum_{k \in \mathcal{K}} x_1^k$)
and the follower's cost function (to be minimized) is of the form

$$f_l(y, x_1, t, x_2).$$

( In the "linear" case, this function is $y^T \sum_{k\in\mathcal{K}} x_1^k + t^T \sum_{k\in\mathcal{K}} x_2^k$.)
$y$ are the leader's optimization variables ( unit costs on the first $n_1$ arcs in the linear case) and $t$ are some fixed parameters corresponding to the unit costs of the rest of the arcs in the linear case.
We remark that, due to the partial or total linearity on $x$, the follower can have multiple optimal strategies corresponding to different revenues for fixed values of $y$.
The best case bi-level model is then

$$(OPLCN) \left\{ \begin{array}{l} \max f_u(y, x_1) \\ y \in C_y \\ (x_1, x_2) \in \text{argmin} f_l(y, x_1, t, x_2) \\ \qquad \left\{ \begin{array}{ll} A_1 x_1^k + A_2 x_2^k = b^k & \forall k \in \mathcal{K} \\ \sum_{k\in\mathcal{K}} (x_1^k, x_2^k)^T \le C & \\ (x_1, x_2) \ge 0. & \end{array} \right. \end{array} \right. ,$$

Depending on the nature of the demands and the function $f_l$, the lower level can be of different types. To get an extensive description of these different multicommodity flow problems and numerical methods used for their resolution we refer to [7, 11].

## 1.2 General ill posed bi-level programs

Let us consider the general bi-level problem :

$$(\mathcal{P}) \left\{ \begin{array}{l} \max f(y, x) \\ y \in K \ , \ x \in \mathcal{S}(y) \ , \end{array} \right.$$

where $K$ and $C$ are non empty convex, closed, bounded subsets of $\mathbb{R}^n$ and

$$\mathcal{S}(y) = \text{argmin} \{ \, f(y, z) + g(y, z) \ |z \in C \ \}, \tag{1.1}$$

$f$ and $g$ are smooth functions from $\mathbb{R}^n \times \mathbb{R}^n$ to $\mathbb{R}$. Moreover, for each $y \in K$, $f(y, .)$ and $g(y, .)$ are convex (we shall precise assumptions later). The main difficulty comes from the fact that the cost "function" $f(y, x)$ , $x \in \mathcal{S}(y)$ is multivalued application since $\mathcal{S}(y)$ is not reduced to a singleton. In addition, it is not clear that $f(y, x) = f(y, \tilde{x})$ for any $x, \tilde{x} \in \mathcal{S}(y)$. Therefore, it is difficult to compute the solutions (if there are any).

The paper is organized as follows. We present the penalized problem and give an existence result in next section. Section 3 is devoted to an asymptotic analysis and we prove that the cluster points of solution to the penalized problems are solutions of a three-level limit problem . In section 4 we give error estimates.

## 2 The penalized problem

We would like to let the upper level revenue function single valued. So we are going to use a penalization process that allow to compute approximate solutions more easily. More precisely, $\varepsilon > 0$ being given, we consider the following penalized problem

$$(\mathcal{P}_\varepsilon) \begin{cases} \max f(y,x) \\ y \in K \ , \ x \in \mathcal{S}_\varepsilon(y) \ , \end{cases}$$

where

$$\mathcal{S}_\varepsilon(y) = \text{argmin}\,\{\, h_\varepsilon(y,z) \ | z \in C \ \}, \tag{2.1}$$

where

$$h_\varepsilon(y) = f(y,z) + g(y,z) + \varepsilon(f(y,z))^2 \ . \tag{2.2}$$

In what follows, we shall note $h = f + g (= h_o)$.
For each nonnegative $\varepsilon$, the bi-level problem $(\mathcal{P}_\varepsilon)$ is well posed . Furthermore, under some general and non restrictive assumptions on $f$ and $g$ we will prove that the upper level function is single valued and continuous with respect to the leader variables $y$.

This regularization technique makes some selection property on the solutions of the lower level problem which is easy to characterize and have an explicit and simple economic interpretation. In almost all other regularization methods, the lower level is replaced by its optimality conditions. The bi-level problem is then considered as a mathematical program with equilibrium constraints. The "hard" part of these constraints ( namely the complementarity conditions ) is then smoothed or penalized. For ill posed problems, these methods make also some selection on the solution set of the lower level but these selections do not have any economic interpretation and convergence results need more restrictive assumptions.

From now we assume that

$$f \text{ and } g \text{ are continous with respect to both variables } x \text{ and } y \ . \tag{2.3}$$

**Lemma 2.1** *For any $\varepsilon > 0$, the low-level problem*

$$\mathcal{Q}_{\varepsilon,y} = \begin{cases} \min h_\varepsilon(y,z) \\ z \in C \ , \end{cases}$$

*admits (at least) a solution so that $\mathcal{S}_\varepsilon(y) \neq \emptyset$. Moreover, there exists a constant $\kappa_y \in \mathbb{R}$ such that*

$$\forall x \in \mathcal{S}_\varepsilon(y) \qquad f(y,x) = \kappa_y \ .$$

*Proof* - $\mathcal{Q}_{\varepsilon,y}$ may written as follows

$$\begin{cases} \min h(y,z) + \varepsilon t^2 \\ t - f(y,z) = 0 \ , \\ z \in C \ , \ t \in \mathbb{R} \end{cases}$$

The existence of such a (convex, smooth) problem is classical. As the problem is strictly convex with respect to $t$ the solution is unique with respect to $t$. Therefore $f(y,\cdot)$ is constant on $\mathcal{S}_\varepsilon(y)$. $\square$

**Lemma 2.2** *Let be $\varepsilon > 0$ fixed. The multi-application $\mathcal{S}_\varepsilon$ is lower semi-continuous in the following sense : if $y_k \to y$ and $x_k \in \mathcal{S}_\varepsilon(y_k)$ then $x_k \to x \in \mathcal{S}_\varepsilon(y)$ (up to a subsequence).*

*Proof* - Let be $x_k \in \mathcal{S}_\varepsilon(y_k) \subset C$. As $C$ is bounded, then $(x_k)$ is bounded as well and converges to some $x$ (up to a subsequence). As $x_k \in \mathcal{S}_\varepsilon(y_k)$ we get

$$\forall z \in C \qquad h(y_k, x_k) + \varepsilon \left(f(y_k, x_k)\right)^2 \leq h(y_k, z) + \varepsilon \left(f(y_k, z)\right)^2 \ .$$

As $f$ and $h$ are continuous with respect to $y$and $x$ we obtain

$$\forall z \in C \qquad h(y, x) + \varepsilon \left(f(y, x)\right)^2 \leq h(y, z) + \varepsilon \left(f(y, z)\right)^2 \ ,$$

that is $x \in \mathcal{S}_\varepsilon(y)$. $\square$

**Lemma 2.3** *Let be $\varepsilon > 0$ fixed. The cost function*

$$v_\varepsilon : y \mapsto \{ f(y,x) \mid x \in \mathcal{S}_\varepsilon(y) \}$$

*is single-valued and continuous.*

*Proof* - We see that the function $v_\varepsilon$ is single valued, with Lemma 2.1. Let us prove the continuity: let be $(y_k)$ a sequence that converges to some $y$. Then $v_\varepsilon(y_k) = f(y_k, x_k)$ where $x_k \in \mathcal{S}_\varepsilon(y_k)$. Lemma 2.2 yields that $x_k$ converges (up to a subsequence) to $x \in \mathcal{S}_\varepsilon(y)$. As $f$ is continuous with respect to $y$ and $x$ we get

$$v_\varepsilon(y_k) = f(y_k, x_k) \to f(y,x) = v_\varepsilon(y) \ .$$

$\square$

We may now give an existence result :

**Theorem 2.1** *For any $\varepsilon > 0$, problem $(\mathcal{P}_\varepsilon)$ admits at least an optimal solution $y_\varepsilon$.*

*Proof* - As $v_\varepsilon$ is continuous and $K$ is bounded, the result follows. $\square$

# 3 Asymptotic results

## 3.1 An convergence result for problem $(\mathcal{P})_\varepsilon$

Now we study the behavior of solutions of $(\mathcal{P}_\varepsilon)$ as $\varepsilon$ to 0. First, we introduce some notations:

$$\widetilde{\mathcal{S}}(y) = \text{argmin}\{ f^2(y,z) \ | z \in \mathcal{S}(y) \ \}, \tag{3.1}$$

where $\mathcal{S}(y)$ is given by (1.1) and

$$(\widetilde{\mathcal{P}}) \begin{cases} \max f(x,y) \\ y \in K \ , \ x \in \widetilde{\mathcal{S}}(y) \ , \end{cases} \tag{3.2}$$

Note that problem $(\widetilde{\mathcal{P}})$ is a three-level problem that can be written in an extended way as follows:

$$(\widetilde{\mathcal{P}}) \begin{cases} \max f(x,y) \\ y \in K \\ x \in \operatorname{argmin} \left\{ f^2(y,z) \ |z \in \operatorname{argmin} \{ f(y,w) + g(y,w) \ |w \in C \ \} \right\} , \end{cases}$$

**Lemma 3.1** *$x \in \widetilde{\mathcal{S}}(y)$ is equivalent to*

$$x \in \mathcal{S}(y) \text{ and } \forall z \in C \text{ such that } h(y,z) = h(y,x), \qquad f^2(y,z) \geq f^2(y,x) \ .$$

*Proof* - Assume that $z$ satisfies $h(y,z) = h(y,x)$ with $x \in \operatorname{argmin} \{ h(y,t) \mid t \in C \}$. Then $z \in \operatorname{argmin} \{ h(y,t) \mid t \in C \}$. □

**Lemma 3.2** *Let $y$ be fixed. If $x_\varepsilon \in \mathcal{S}_\varepsilon$ converges to some $\bar{x}$, then $\bar{x} \in \widetilde{\mathcal{S}}(y)$*

*Proof* - Assume $x_\varepsilon \in \mathcal{S}_\varepsilon$ and $x_\varepsilon \to \bar{x}$ as $\varepsilon \to 0$. For every $z \in C$ we get

$$h(y,x_\varepsilon) + \varepsilon f^2(y,x_\varepsilon) \leq h(y,z) + \varepsilon f^2(y,z) \ .$$

When $\varepsilon \to 0$, as the functions are continuous we obtain

$$\forall z \in C \qquad h(y,\bar{x}) \leq h(y,z) \ ,$$

that is $\bar{x} \in \mathcal{S}(y)$.
Let be $\tilde{x} \in C$ such that $h(y,\tilde{x}) = h(y,\bar{x})$. Then

$$\begin{aligned} h(y,x_\varepsilon) + \varepsilon f^2(y,x_\varepsilon) &\leq h(y,\tilde{x}) + \varepsilon f^2(y,\tilde{x}) && \text{since } \tilde{x} \in C \\ &\leq h(y,\bar{x}) + \varepsilon f^2(y,\tilde{x}) && \text{since } h(y,\tilde{x}) = h(y,\bar{x}) \\ &\leq h(y,x_\varepsilon) + \varepsilon f^2(y,\tilde{x}) && \text{since } x_\varepsilon \in C \text{ and } \bar{x} \in \mathcal{S}(y) \ . \end{aligned}$$

Therefore

$$\forall \tilde{x} \in C \text{ such that } h(y,\tilde{x}) = h(y,\bar{x}), \ f^2(y,x_\varepsilon) \leq f^2(y,\tilde{x}) \ .$$

Passing to the limit with the continuity of $f$ gives

$$\forall \tilde{x} \in C \text{ such that } h(y,\tilde{x}) = h(y,\bar{x}), \ f^2(y,\bar{x}) \leq f^2(y,\tilde{x}) \ .$$

With Lemma 3.1 we conclude that $x \in \widetilde{\mathcal{S}}(y)$. □

Next Lemma is the most important in the sequel: it deals with a continuity property of the multi-application $\mathcal{S}_\varepsilon$.

**Lemma 3.3** *Let* $(y_\varepsilon, x_\varepsilon \in \mathcal{S}_\varepsilon(y_\varepsilon))$ *converging to* $(\bar{y}, \bar{x})$ *. Then* $\bar{x} \in \widetilde{\mathcal{S}}(\bar{y})$*.*

Assume that we have proved this lemma for the moment. We may give an asymptotic result for problem $(\mathcal{P})_\varepsilon$.

**Theorem 3.1** *Let* $y_\varepsilon$ *an optimal solution to* $(\mathcal{P})_\varepsilon$*. Then* $y_\varepsilon$ *converges to some* $\bar{y}$ *(up to a subsequence) and* $\bar{y}$ *is an optimal solution to* $(\widetilde{\mathcal{P}})$*.*

*Proof* - Let $y_\varepsilon$ an optimal solution to $(\mathcal{P})_\varepsilon$. Then $y_\varepsilon \in K$ which is bounded. So (extracting a subsequence) we may assert that $y_\varepsilon$ converges to $\bar{y}$. As $K$ is closed then $\bar{y} \in K$. As $y_\varepsilon$ is an optimal solution to $(\mathcal{P})_\varepsilon$ we have

$$\forall \tilde{y} \in K \ , \forall \tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(\tilde{y}) \qquad f(y_\varepsilon, x_\varepsilon) \geq f(\tilde{y}, \tilde{x}_\varepsilon) \tag{3.3}$$

where $x_\varepsilon \in \mathcal{S}_\varepsilon(y_\varepsilon)$. Note that $\tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(\tilde{y})$ implies that $\tilde{x}_\varepsilon \in C$. So $\tilde{x}_\varepsilon$ is bounded and converges to $\tilde{x}$ (up to a subsequence) with $\tilde{x} \in \widetilde{\mathcal{S}}(\tilde{y})$ (Lemma 3.2).
Passing to the limit in (3.3) gives

$$\forall \tilde{y} \in K \ , \exists \tilde{x} \in \widetilde{\mathcal{S}}(\tilde{y}) \text{ such that } f(\bar{y}, \bar{x}) \geq f(\tilde{y}, \tilde{x}) \ ,$$

where $\bar{x}$ is the limit (of a subsequence) of $x_\varepsilon$. Lemma 3.3 shows that $\bar{x} \in \widetilde{\mathcal{S}}(\bar{y})$.
Thanks to the definition of $\widetilde{\mathcal{S}}(\tilde{y})$ we note that $f(\tilde{y}, \cdot)$ is constant on $\widetilde{\mathcal{S}}(\tilde{y})$, namely

$$\forall z \in \widetilde{\mathcal{S}}(\tilde{y}) \qquad f(\tilde{y}, z) = f(\tilde{y}, \tilde{x}) \ .$$

Finally

$$\forall \tilde{y} \in K, \ \forall \tilde{x} \in \widetilde{\mathcal{S}}(\tilde{y}) \qquad f(\bar{y}, \bar{x}) \geq f(\tilde{y}, \tilde{x}) \ ,$$

with $\bar{x} \in \widetilde{\mathcal{S}}(\bar{y})$. This means that $\bar{y}$ is an optimal solution to $(\widetilde{\mathcal{P}})$. □

## 3.2 Proof of Lemma 3.3

Let $y_\varepsilon$ converging to $\bar{y}$ and $x_\varepsilon \in \mathcal{S}_\varepsilon(y_\varepsilon)$. As $x_\varepsilon \in C$ (bounded) one may extract a subsequence converging to $\bar{x}$. We are going to prove that $\bar{x} \in \widetilde{\mathcal{S}}(\bar{y})$.
Let us set

$$\alpha_\varepsilon = h(y_\varepsilon, x_\varepsilon) + o(\varepsilon) \ , \tag{3.4}$$

and

$$\Lambda_\varepsilon = \{ \, x \in C \,|\, h(y_\varepsilon, x) \leq \alpha_\varepsilon \, \} \ . \tag{3.5}$$

We first prove that $\bar{x} \in \mathcal{S}(\bar{y})$. As $x_\varepsilon \in \mathcal{S}_\varepsilon(y_\varepsilon)$ we have

$$\forall z \in C \qquad h(y_\varepsilon, x_\varepsilon) + \varepsilon\, f^2(y_\varepsilon, x_\varepsilon) \leq h(y_\varepsilon, z) + \varepsilon\, f^2(y_\varepsilon, z) \ ; \tag{3.6}$$

as $f$ and $g$ are continuous, passing to the limit gives

$$\forall z \in C \qquad h(\bar{y}, \bar{x}) \leq h(\bar{y}, z) \ ,$$

that is $\bar{x} \in \mathcal{S}(\bar{y})$.
Let $\tilde{x} \in \mathcal{S}(\bar{y})$. Suppose for a while that $\exists \tilde{\varepsilon}$ such that

$$\forall \varepsilon \leq \tilde{\varepsilon} \qquad \tilde{x} \in \Lambda_\varepsilon \ . \tag{3.7}$$

We get

$$h(y_\varepsilon, \tilde{x}) \leq h(y_\varepsilon, x_\varepsilon) + o(\varepsilon) \ ;$$

with relation (3.6) this gives

$$\forall z \in C \qquad h(y_\varepsilon, \tilde{x}) + \varepsilon\, f^2(y_\varepsilon, x_\varepsilon) \leq h(y_\varepsilon, z) + \varepsilon\, f^2(y_\varepsilon, z) + o(\varepsilon) \ . \tag{3.8}$$

As $\tilde{x} \in C$ relation (3.6) yields as well

$$h(y_\varepsilon, x_\varepsilon) + \varepsilon\, f^2(y_\varepsilon, x_\varepsilon) \leq h(y_\varepsilon, \tilde{x}) + \varepsilon\, f^2(y_\varepsilon, \tilde{x}) \ .$$

Adding these two relations gives

$$\forall z \in C \qquad h(y_\varepsilon, x_\varepsilon) + 2\varepsilon\, f^2(y_\varepsilon, x_\varepsilon) \leq h(y_\varepsilon, z) + \varepsilon\, f^2(y_\varepsilon, z) + \varepsilon\, f^2(y_\varepsilon, \tilde{x}) + o(\varepsilon) \ ; \tag{3.9}$$

the choice of $z = x_\varepsilon$ implies

$$\varepsilon\, f^2(y_\varepsilon, x_\varepsilon) \leq \varepsilon\, f^2(y_\varepsilon, \tilde{x}) + o(\varepsilon) \ ,$$

that is

$$f^2(y_\varepsilon, x_\varepsilon) \leq f^2(y_\varepsilon, \tilde{x}) + \frac{o(\varepsilon)}{\varepsilon} \ .$$

Passing to the limit gives finally

$$\forall \tilde{x} \in \mathcal{S}(\bar{y}) \qquad f^2(\bar{y}, \bar{x}) \leq f^2(\bar{y}, \tilde{x}) \ .$$

This means that $\bar{x} \in \widetilde{\mathcal{S}}(\bar{y})$.

Unfortunately, there is no reason for "assumption " (3.7) to be satisfied and we must get rid of it. We are going to adapt the previous proof (we gave the main ideas). If $\tilde{x} \notin \Lambda_\varepsilon$ then we perform a projection: we call $\tilde{x}_\varepsilon$ the projection of $\tilde{x}$ on $\Lambda_\varepsilon$. We are going to show that $\tilde{x}_\varepsilon$converges to $\tilde{x}$.
As $\tilde{x} \notin \Lambda_\varepsilon$ we get $\alpha_\varepsilon < h(y_\varepsilon, \tilde{x})$ . Let us call $\sigma_{\alpha_\varepsilon}(h)$ the following real number

$$\sigma_{\alpha_\varepsilon}(h) = \inf_{x \in [\alpha_\varepsilon < h(y_\varepsilon, \cdot)]} \frac{h(y_\varepsilon, x) - \alpha_\varepsilon}{d(x, \Lambda_\varepsilon)} \ , \tag{3.10}$$

where $d(x, \Lambda_\varepsilon)$ is the distance between $x$ and $\Lambda_\varepsilon$) and

$$[\alpha_\varepsilon < h(y_\varepsilon, \cdot)] = \{\ x \in \mathbb{R}^n \mid \alpha_\varepsilon < h(y_\varepsilon, x)\} \ .$$

This so called Hoffman constant can be defined following for instance Azé and Corvellec [2]. Therefore

$$h(y_\varepsilon, \tilde{x}) - \alpha_\varepsilon \geq d(\tilde{x}, \Lambda_\varepsilon)\, \sigma_{\alpha_\varepsilon}(h) \ .$$

As $d(\tilde{x},\Lambda_\varepsilon)=d(\tilde{x},\tilde{x}_\varepsilon)$ we obtain

$$d(\tilde{x},\tilde{x}_\varepsilon)\le \frac{h(y_\varepsilon,\tilde{x})-\alpha_\varepsilon}{\sigma_{\alpha_\varepsilon}(h)} \ .$$

We have to estimate $\sigma_{\alpha_\varepsilon}(h)$. In particular we look for $\sigma_o>0$ such that

$$\forall\varepsilon \qquad \sigma_{\alpha_\varepsilon}(h)\ge\sigma_o \ .$$

In [2], it is shown that

$$\sigma_{\alpha_\varepsilon}(h)\ge \inf_{h(y_\varepsilon,x)=\alpha_\varepsilon} |\nabla_x h(y_\varepsilon,x)| \ ,$$

where $|\nabla_x h(y_\varepsilon,x)|$ stands for the strong slope of $h$ at $(y_\varepsilon,x)$ with respect to $x$ ([2]); the strong-slope of a function $\varphi$ at $x$ is defined as

$$|\nabla\varphi|(x):=\begin{cases}0 & \text{if } x \text{ is a local minimum of } \varphi \ ,\\ \limsup\limits_{y\to x}\dfrac{\varphi(x)-\varphi(y)}{d(x,y)} & \text{otherwise}\end{cases}$$

Assume we can find $\sigma_o>0$ and $\varepsilon_o>0$ such that

$$\forall\varepsilon\le\varepsilon_o \quad \inf_{h(y_\varepsilon,x)=\alpha_\varepsilon}|\nabla_x h(y_\varepsilon,x)|\ge\sigma_o \ , \tag{3.11}$$

then

$$d(\tilde{x},\tilde{x}_\varepsilon)\le\frac{h(y_\varepsilon,\tilde{x})-\alpha_\varepsilon}{\sigma_o}=\frac{h(y_\varepsilon,\tilde{x})-h(y_\varepsilon,x_\varepsilon)+o(\varepsilon)}{\sigma_o}\to 0 \ .$$

Indeed $y_\varepsilon\to\bar{y}$, $x_\varepsilon\to\bar{x}$, $h$ is continuous and $h(\bar{y},\bar{x})=h(\bar{y},\tilde{x})$.
We may now end the proof. We can use relation (3.9) with $\tilde{x}_\varepsilon$ instead of $\tilde{x}$ so that

$$\forall z\in C \qquad h(y_\varepsilon,x_\varepsilon)+2\varepsilon\, f^2(y_\varepsilon,x_\varepsilon)\le h(y_\varepsilon,z)+\varepsilon\, f^2(y_\varepsilon,z)+\varepsilon\, f^2(y_\varepsilon,\tilde{x}_\varepsilon)+o(\varepsilon) \ ;$$

we choose $z=x_\varepsilon$ once again to get

$$f^2(y_\varepsilon,x_\varepsilon)\le f^2(y_\varepsilon,\tilde{x}_\varepsilon)+\frac{o(\varepsilon)}{\varepsilon} \ .$$

Passing to the limit as $\varepsilon\to 0$ gives (for every $\tilde{x}\in\mathcal{S}(\bar{y})$

$$f^2(\bar{y},\bar{x})\le f^2(\bar{y},\tilde{x}) \ .$$

This means that $\bar{x}\in\widetilde{\mathcal{S}}(\bar{y})$. □

**Remark 3.1** *It is clear that assumption* (3.11) *is satisfied if $h$ is linear ("linear" case). Next problem is to find simple conditions for $(\bar{y},\bar{x})$ to get* (3.11) *when $h$ is not linear. One hint is to assume that $h$ is $\mathcal{C}^1$ and that $\|\nabla_x h(\bar{y},\bar{x}\|\neq 0$; then the strong slope $|\nabla_x h(y_\varepsilon,x)|$ coincides with the norm $\|\nabla_x h(y_\varepsilon,x)\|$ of the gradient of $h$ with respect to $x$. With the convergence of $(y_\varepsilon,x_\varepsilon)$ to $(\bar{y},\bar{x})$ (up to a subsequence), there exist $\varepsilon_o$ and $\eta>0$ such that*

$$\forall\varepsilon\le\varepsilon_o \qquad \|\nabla_x h(y_\varepsilon,x_\varepsilon)\|\ge\eta>0 \ ;$$

*next we have to prove that $\|\nabla_x h(y_\varepsilon,x)\|\ge\eta$ for any $x$ such that $h(y_\varepsilon,x)=\alpha_\varepsilon$. A good tool could be an "local inversion theorem" for the multivalued case but it is not obvious. The problem is still open. We have the same challenge in next section.*

### 3.3 Comparison of $(\mathcal{P})$ and $(\widetilde{\mathcal{P}})$

Now, it is clear that a solution of the penalized problem $(\mathcal{P}_\varepsilon)$ is a good approximation of a solution of $(\widetilde{\mathcal{P}})$. Anyway, it is not a solution (a priori) of the problem in consideration $(\mathcal{P})$. So we have to compare $(\mathcal{P})$ and $(\widetilde{\mathcal{P}})$.

The second level of $(\widetilde{\mathcal{P}})$ clearly disappears when the solutions set of the lower level of the initial problem corresponds to the same revenue for each value of $y$ (or are unique). In this case $(\mathcal{P})$ and $(\widetilde{\mathcal{P}})$ are equivalent. In other cases, the solution of $(\widetilde{\mathcal{P}})$ corresponds to some "optimal worst" case solution.

This solution is still important for the decision makers of the upper level problem.

**Remark 3.2** *Using the same regularization technique, if we replace $\varepsilon(f(y,z))^2$ by $\varepsilon(g(y,z))^2$ in the definition of $h_\varepsilon$, we will obtain (at the limit) an optimal solution of $(\mathcal{P})$ which corresponds to an optimal best case solution of our asymmetric game .*

## 4 Error estimates

**In this section we assume that the function $f$ is nonnegative.**

### 4.1 Preliminary results

**Lemma 4.1** *Let be $\varepsilon > \varepsilon' > 0$ and $y \in K$. Let be $x_\varepsilon \in \mathcal{S}_\varepsilon(y)$ and $\tilde{x} \in \mathcal{S}_{\varepsilon'}(y)$ Then we get*

$$f^2(y, x_\varepsilon) \le f^2(y, \tilde{x}) \ .$$

*Proof* - Let us fix $\varepsilon > \varepsilon' > 0$ and choose some $y \in K$. Let be $x_\varepsilon \in \mathcal{S}_\varepsilon(y)$ and $\tilde{x} \in \mathcal{S}_{\varepsilon'}(y)$. Assume that

$$f^2(y, \tilde{x}) < f^2(y, x_\varepsilon) \ . \tag{4.1}$$

As $\tilde{x} \in \mathcal{S}_{\varepsilon'}(y)$ and $x_\varepsilon \in C$, we have

$$h(y, \tilde{x}) + \varepsilon' f^2(y, \tilde{x}) \le h(y, x_\varepsilon) + \varepsilon' f^2(y, x_\varepsilon) \ ,$$

$$h(y, \tilde{x}) + \varepsilon' f^2(y, \tilde{x}) + (\varepsilon - \varepsilon') f^2(y, \tilde{x}) \le h(y, x_\varepsilon) + \varepsilon' f^2(y, x_\varepsilon) + (\varepsilon - \varepsilon') f^2(y, \tilde{x})$$

With (4.1) and $\varepsilon > \varepsilon' > 0$, we obtain

$$h(y, \tilde{x}) + \varepsilon f^2(y, \tilde{x}) \le h(y, x_\varepsilon) + \varepsilon' f^2(y, x_\varepsilon) + (\varepsilon - \varepsilon') f^2(y, x_\varepsilon) < h(y, x_\varepsilon) + \varepsilon f^2(y, x_\varepsilon)$$

So

$$h(y, \tilde{x}) + \varepsilon f^2(y, \tilde{x}) < \text{argmin } \{ \ h(y, x) + \varepsilon f^2(y, x), \ x \in C\}$$

and we get a contradiction. $\square$

**Lemma 4.2** *Let be $\varepsilon > \varepsilon' > 0$ and $y_\varepsilon$ (respectively $y_{\varepsilon'}$) a solution to $(\mathcal{P}_\varepsilon)$ (respectively $(\mathcal{P}_{\varepsilon'})$). Let be $x_\varepsilon \in \mathcal{S}_\varepsilon(y_\varepsilon)$ and $x_{\varepsilon'} \in \mathcal{S}_{\varepsilon'}(y_{\varepsilon'})$ Then*

$$f^2(y_\varepsilon, x_\varepsilon) \leq f^2(y_{\varepsilon'}, x_{\varepsilon'}) \leq f^2(y^*, x^*) \ ,$$

*where $y^*$ is a solution to $(\widetilde{\mathcal{P}})$ with $x^* \in \mathcal{S}(y^*)$.*

*Proof* - Using Lemma 4.1 with $y = y_\varepsilon$ and $x_\varepsilon \in \mathcal{S}_\varepsilon(y_\varepsilon)$ gives

$$\forall \tilde{x} \in \mathcal{S}_{\varepsilon'}(y_\varepsilon) \qquad f^2(y_\varepsilon, x_\varepsilon) \leq f^2(y_\varepsilon, \tilde{x}) \ . \tag{4.2}$$

As $y_{\varepsilon'}$ is a solution of $(\mathcal{P}_{\varepsilon'})$ we get

$$\forall y \in K,\ \forall x \in \mathcal{S}_{\varepsilon'}(y) \qquad f(y_{\varepsilon'}, x_{\varepsilon'}) \geq f(y, x) \ .$$

We may choose in particular $y = y_\varepsilon$ and $x = \tilde{x} \in \mathcal{S}_{\varepsilon'}(y_\varepsilon)$ to get

$$\forall \tilde{x} \in \mathcal{S}_{\varepsilon'}(y_\varepsilon) \qquad f(y_{\varepsilon'}, x_{\varepsilon'}) \geq f(y_\varepsilon, \tilde{x}) \ . \tag{4.3}$$

As $f$ is assumed to be nonnegative we finally obtain

$$f(y_\varepsilon, x_\varepsilon) \leq f(y_\varepsilon, \tilde{x}) \leq f(y_{\varepsilon'}, x_{\varepsilon'}) \ .$$

Therefore the family $(f(y_\varepsilon, x_\varepsilon)$ is increasing. The convergence of $f(y_\varepsilon, x_\varepsilon)$ to $f(y^*, x^*)$ ($f$ is continuous) achieves the proof since $f(y^*, x^*)$ is the limit and the upper bound of the family $(f(y_\varepsilon, x_\varepsilon))$ . □

**Lemma 4.3** *Let be $\varepsilon > 0$ and $\tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(y^*)$ where $y^*$ is a solution to $(\widetilde{\mathcal{P}})$. Then*

$$\forall x_\varepsilon \in \mathcal{S}_\varepsilon(y_\varepsilon) \qquad f(y^*, \tilde{x}_\varepsilon) \leq f(y_\varepsilon, x_\varepsilon) \leq f(y^*, x^*) \ . \tag{4.4}$$

*Proof* - This is a direct consequence of Lemma 4.2 : the relation $f(y_\varepsilon, x_\varepsilon) \leq f(y^*, x^*)$ is obvious. and the relation $f(y^*, \tilde{x}_\varepsilon) \leq f(y_\varepsilon, x_\varepsilon)$ comes from the fact that $y_\varepsilon$ is a solution to $(\mathcal{P}_\varepsilon)$. □

The purpose of this subsection is to study the behavior of $f(y^*, x^*) - f(y_\varepsilon, x_\varepsilon)$ as $\varepsilon \to 0$ and provide (if possible) an error estimate. The previous lemmas show that it is sufficient to study $f(y^*, x^*) - f(y^*, \tilde{x}_\varepsilon)$ for some $\tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(y^*)$.

**We assume from now that $C$ is polyhedral :**

$$C = \{\ x \in \mathbb{R}^n \ \mid\ Ax = b,\ x \geq 0\ \} \ ,$$

where $A$ is a $m \times n$ real matrix and $b \in \mathbb{R}^m$.

In the sequel $y^*$ is a solution to $(\widetilde{\mathcal{P}})$ (which existence is given by Theorem 3.3 ) and $x^* \in \widetilde{\mathcal{S}}(y^*)$ ( see (3.1)) so that

$$x^* \in \ \operatorname{argmin} \ \{\ f^2(y^*, z) \mid z \in \ \operatorname{argmin} \ \{h(y^*, \zeta)\ , \zeta \in C\}\ \} \ .$$

Let us denote

$$\alpha^* = h(y^*, x^*) = f(y^*, x^*) + g(y^*, x^*) \text{ and } \beta^* = f(y^*, x^*) \ . \tag{4.5}$$

Note that $\beta^*$ is the optimal value for $(\widetilde{\mathcal{P}})$ (the upper level) so that we may assume that $\beta^* \neq 0$ (otherwise the problem is trivial). We set

$$C^* = \{ \ x \in C \ \mid \ h(y^*, x) \leq \alpha^* \text{ and } f(y^*, x) \leq \beta^* \ \} \ . \tag{4.6}$$

Let us give an important property of $C^*$ :

**Proposition 4.1** *Assume $y^*$ is a solution to $(\widetilde{\mathcal{P}})$ and $C^*$ is defined with (4.6), then*

$$C^* = \{ \ x \in C \ \mid \ h(y^*, x) = \alpha^* \text{ and } f(y^*, x) = \beta^* \ \}$$

*and*

$$C^* = \{ \ x \in C \ \mid \ h(y^*, x) + f(y^*, x) \leq \sigma^* \overset{def}{:=} \alpha^* + \beta^* \ \} \ .$$

*Proof* - Note that it impossible to get $h(y^*, x) \leq \alpha^*$, if $x \in C^*$. Indeed, as $x^* \in \widetilde{\mathcal{S}}(y^*)$ then $x^* \in S(y^*) = \text{ argmin } \{h(y^*, \zeta) \ , \zeta \in C\}$ . Therefore :

$$\forall \zeta \in C \qquad h(y^*, x^*) \leq h(y^*, \zeta) \ . \tag{4.7}$$

Setting $\zeta = x \in C^*$ gives

$$\alpha^* = h(y^*, x^*) \leq h(y^*, x) \leq \alpha^* \ .$$

So

$$\forall x \in C^* \qquad h(y^*, x) = \alpha^* \ .$$

The same remark holds for $\beta^*$ so that

$$C^* = \{ \ x \in C \ \mid \ h(y^*, x) = \alpha^* \text{ and } f(y^*, x) = \beta^* \ \} \ . \tag{4.8}$$

Let us call

$$C' = \{ \ x \in C \ \mid \ h(y^*, x) + f(y^*, x) \leq \sigma^* \ \} \ .$$

It is obvious that $C^* \subset C'$. Conversely, let be $x \in C'$. Relation (4.7) yields $\alpha^* \leq h(y^*, x)$ so that

$$\alpha^* + f(y^*, x) \leq \alpha^* + \beta^* \ .$$

This gives $f(y^*, x) \leq \beta^*$. Similarly, we get $h(y^*, x) \leq \alpha^*$ and $x \in C^*$. □

The main point is now to estimate (roughly speaking) the distance between the solution $x^*$ and $\mathcal{S}_\varepsilon(y^*)$. As $x^* \in C^*$ and $C^*$ is defined with inequalities, we first assume a Hoffman-type condition.

### 4.2 Error estimates under an Hoffman hypothesis

Following Azé and Corvellec [2] we know that

$$\inf_{[\sigma^*<f(y^*,\cdot)+h(y^*,\cdot)]} |\nabla_x \left(f(y^*,\cdot)+h(y^*,\cdot)\right)| \leq \inf_{x\in[\sigma^*<f(y^*,\cdot)+h(y^*,\cdot)]} \frac{f(y^*,x)+h(y^*,x)-\sigma^*}{d(x,[f(y^*,\cdot)+h(y^*,\cdot)\leq\sigma^*]} .$$

The notation $[\sigma^* < f(y^*,\cdot)+h(y^*,\cdot)]$ stands for the set

$$\{x\in\mathbb{R}^n \mid \sigma^* < f(y^*,x)+h(y^*,x)\ \} .$$

We note that $[f(y^*,\cdot)+h(y^*,\cdot)\leq\sigma^*]=C^*$. In this subsection, we assume the following

$$(\mathcal{H}_1) \qquad \gamma^* := \inf_{[\sigma^*<f(y^*,\cdot)+h(y^*,\cdot)]} |\nabla_x \left(f(y^*,\cdot)+h(y^*,\cdot)\right)| > 0.$$

Let us call $\gamma = \dfrac{1}{\gamma^*}$ : assumption $(\mathcal{H}_1)$ implies that

$$\forall\varepsilon>0,\ \forall\tilde{x}_\varepsilon\in\mathcal{S}_\varepsilon(y^*) \qquad \exists x_\varepsilon^*\in C^* \text{ s.t. } \|\tilde{x}_\varepsilon - x_\varepsilon^*\| \leq \gamma\left[f(y^*,\tilde{x}_\varepsilon)+h(y^*,\tilde{x}_\varepsilon)-\alpha^*-\beta^*\right] . \tag{4.9}$$

Note also that relation (4.4) of Lemma 4.3 yields that

$$\forall\tilde{x}_\varepsilon\in\mathcal{S}_\varepsilon(y^*) \qquad f(y^*,\tilde{x}_\varepsilon)\leq\beta^*$$

and

$$h(y^*,\tilde{x}_\varepsilon)\leq\alpha^*+\varepsilon\beta^*$$

because of the definition of $\mathcal{S}_\varepsilon(y^*)$.Therefore

$$\forall\tilde{x}_\varepsilon\in\mathcal{S}_\varepsilon(y^*) \qquad f(y^*,\tilde{x}_\varepsilon)+h(y^*,\tilde{x}_\varepsilon)-\alpha^*-\beta^*\leq\varepsilon$$

and

$$\forall\varepsilon>0,\ \forall\tilde{x}_\varepsilon\in\mathcal{S}_\varepsilon(y^*) \qquad \exists x_\varepsilon^*\in C^* \text{ s.t. } \|\tilde{x}_\varepsilon - x_\varepsilon^*\| \leq \gamma\varepsilon . \tag{4.10}$$

The existence of such Lipschitzian error bound for convex or general inequalities is, itself, an interesting domain of research. It is strongly related to metric regularity properties. A large number of conditions and characterizations can be found in [2, 3, 9, 10, 13, 15, 16]. This list of references constitutes a small but significant part of the existing literature.

**Remark 4.1** *1. Assumption $(\mathcal{H}_1)$ is fulfilled if the functions $f$ and $g$ are linear with respect to $x$. Indeed they cannot be identically equal to 0 and the strong slope coincides with the norm of gradient which is a positive constant.*
*2. $x_\varepsilon^*$ is the projection of $\tilde{x}_\varepsilon$ on $C^*$.*

**Lemma 4.4** *Both $\tilde{x}_\varepsilon\in\mathcal{S}_\varepsilon(y^*)$ and $x_\varepsilon^*$ given by (4.10) converge to $x^*\in\mathcal{S}(y^*)$ as $\varepsilon\to 0$.*

*Proof* - We know that $\tilde{x}_\varepsilon \to x^*$ (with the previous results). Let us set $d_\varepsilon = \dfrac{\tilde{x}_\varepsilon - x^*_\varepsilon}{\varepsilon}$ . As $d_\varepsilon$ is bounded (by $\gamma$) it clear that $x^*_\varepsilon$ and $\tilde{x}_\varepsilon$ have the same limit point (namely $x^*$). □

In what follows $\tilde{x}_\varepsilon$ is an element of $\mathcal{S}_\varepsilon(y^*)$ and $x^*_\varepsilon$ is the associated element given by (4.10) .

Let us define

$$I(x^*) = \{i \in \{1, \cdots, n\} \ \mid \ x^*_i = 0 \ \} \ , \text{ and } \tilde{C} = \{ \ d \in \mathbb{R}^n \ \mid \ Ad = 0 \ , \ d_{|I(x^*)} \geq 0 \ \} \ .$$

**Let $d$ be in $\tilde{C}$.**

Then, there exists $\varepsilon_d > 0$ such that $\forall \varepsilon < \varepsilon_o, \ x^*_\varepsilon + \varepsilon d \in C$ . Indeed

- $A(x^*_\varepsilon + \varepsilon d) = A(x^*_\varepsilon) + \varepsilon Ad = A(x^*_\varepsilon) = b$ .
- If $i \in I(x^*)$, then $(x^*_\varepsilon + \varepsilon d)_i \geq x^*_{\varepsilon,i} \geq 0$.
- If $i \notin I(x^*)$, then $x^*_i > 0$. As $x^*_\varepsilon \to x^*, \exists \varepsilon_i$ such that $x^*_{\varepsilon,i} > 0$ forall $\varepsilon \leq \varepsilon_i$. Then we choose $\eta = \inf_{i \notin I(x^*)} \{\varepsilon_i\}$ so that

$$\forall \varepsilon \leq \eta \qquad x^*_{\varepsilon,i} > 0 \ .$$

Now choosing $\varepsilon_d \leq \eta$ small enough we get $(x^*_\varepsilon + \varepsilon d)_i \geq 0$ for any $\varepsilon \leq \varepsilon_o$.

As $\tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(y^*)$ and $x^*_\varepsilon + \varepsilon d \in C$ we have

$$h(y^*, \tilde{x}_\varepsilon) + \varepsilon f^2(y^*, \tilde{x}_\varepsilon) \leq h(y^*, x^*_\varepsilon + \varepsilon d) + \varepsilon f^2(y^*, x^*_\varepsilon + \varepsilon d) \ ,$$

$$h(y^*, \tilde{x}_\varepsilon) - h(y^*, x^*_\varepsilon + \varepsilon d) + \varepsilon \left[ f^2(y^*, \tilde{x}_\varepsilon) - f^2(y^*, x^*_\varepsilon + \varepsilon d) \right] \leq 0 \ .$$

As the functions are $\mathcal{C}^1$, we have

$$h(y^*, \tilde{x}_\varepsilon) = h(y^*, x^*_\varepsilon) + \nabla_x h(y^*, x^*_\varepsilon) \cdot (\tilde{x}_\varepsilon - x^*_\varepsilon) + (\tilde{x}_\varepsilon - x^*_\varepsilon) o(\tilde{x}_\varepsilon - x^*_\varepsilon)$$

$$h(y^*, \tilde{x}_\varepsilon) = h(y^*, x^*_\varepsilon) + \varepsilon \nabla_x h(y^*, x^*_\varepsilon) \cdot d_\varepsilon + \varepsilon d_\varepsilon \, o(\varepsilon d_\varepsilon) \ , \tag{4.11}$$

and

$$h(y^*, x^*_\varepsilon + \varepsilon d) = h(y^*, x^*_\varepsilon) + \varepsilon \nabla_x h(y^*, x^*_\varepsilon) \cdot d + \varepsilon d \, o(\varepsilon d) \ , \tag{4.12}$$

where $\nabla_x h$ stands for the derivative of $h$ with respect to $x$. As $x^*_\varepsilon \in C^*$ and $\tilde{x}_\varepsilon \in C$ then

$$h(y^*, x^*_\varepsilon) = \alpha^* = h(y^*, x^*) \leq h(y^*, \tilde{x}_\varepsilon) \ .$$

With relation (4.11) this gives

$$\nabla_x h(y^*, x^*_\varepsilon) \cdot d_\varepsilon + d_\varepsilon \, o(\varepsilon d_\varepsilon) = \frac{h(y^*, \tilde{x}_\varepsilon) - h(y^*, x^*_\varepsilon)}{\varepsilon} \geq 0 \ .$$

As $d_\varepsilon$ is bounded (by $\gamma$), there exist cluster points; passing to the limit gives

$$\nabla_x h(y^*, x^*) \cdot \tilde{d} = \lim_{\varepsilon \to 0} \nabla_x h(y^*, x^*_\varepsilon) \cdot d_\varepsilon \geq 0 \ , \tag{4.13}$$

for any cluster point $\tilde{d}$ of the family $d_\varepsilon$.

In addition, we obtain with (4.11) and (4.12)

$$\varepsilon \nabla_x h(y^*, x^*_\varepsilon) \cdot d_\varepsilon + \varepsilon d_\varepsilon \, o(\varepsilon d_\varepsilon) - \varepsilon \nabla_x h(y^*, x^*_\varepsilon) \cdot d - \varepsilon d \, o(\varepsilon d) + \varepsilon \left[ f^2(y^*, \tilde{x}_\varepsilon) - f^2(y^*, x^*_\varepsilon + \varepsilon d) \right] \leq 0 \ ,$$

that is

$$\nabla_x h(y^*, x^*_\varepsilon) \cdot (d_\varepsilon - d) + d_\varepsilon \, o(\varepsilon d_\varepsilon) - d \, o(\varepsilon d) + \left[ f^2(y^*, \tilde{x}_\varepsilon) - f^2(y^*, x^*_\varepsilon + \varepsilon d) \right] \leq 0 \ .$$

Passing to the limit (with Lemma 4.4) we obtain

$$\nabla_x h(y^*, x^*) \cdot (\tilde{d} - d) \leq 0 \ , \tag{4.14}$$

where $\tilde{d}$ is a cluster point of the sequence $d_\varepsilon$ and any $d \in \tilde{C}$. As $d = 0$ belongs to $\tilde{C}$, we get

$$\nabla_x h(y^*, x^*) \cdot \tilde{d} \leq 0 \ .$$

Finally, we obtain with (4.13)

$$\nabla_x h(y^*, x^*) \cdot \tilde{d} = \lim_{\varepsilon \to 0} \nabla_x h(y^*, x^*_\varepsilon) \cdot \frac{\tilde{x}_\varepsilon - x^*_\varepsilon}{\varepsilon} = 0 \ . \tag{4.15}$$

This means that

$$\nabla_x h(y^*, x^*_\varepsilon) \cdot (\tilde{x}_\varepsilon - x^*_\varepsilon) = o(\varepsilon).$$

As

$$h(y^*, \tilde{x}_\varepsilon) = h(y^*, x^*_\varepsilon) - \nabla_x h(y^*, x^*_\varepsilon) \cdot (x^*_\varepsilon - \tilde{x}_\varepsilon) + (x^*_\varepsilon - \tilde{x}_\varepsilon) \, o(x^*_\varepsilon - \tilde{x}_\varepsilon)$$

we get

$$h(y^*, \tilde{x}_\varepsilon) - h(y^*, x^*_\varepsilon) = o(\varepsilon) - \varepsilon d_\varepsilon \, o(\varepsilon d_\varepsilon) = o(\varepsilon) \ .$$

As $x^*_\varepsilon \in C^*$ then $h(y^*, x^*_\varepsilon) = \alpha^*$ and

$$\forall \tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(y^*) \qquad h(y^*, \tilde{x}_\varepsilon) = h(y^*, x^*) + o(\varepsilon) \ . \tag{4.16}$$

As $h$ and $f^2$ play similar roles we have the same result for $f^2$. More precisely

$$\forall \tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(y^*) \qquad f^2(y^*, \tilde{x}_\varepsilon) - f^2(y^*, x^*) = o(\varepsilon) \ . \tag{4.17}$$

We just proved the following result

**Theorem 4.1** *Assume that* $(\mathcal{H}_1)$ *is satisfied ; let* $y_\varepsilon$ *be a solution to* $(\mathcal{P}_\varepsilon)$ *and* $\tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(y^*)$. *Then*

$$h(y^*, \tilde{x}_\varepsilon) - h(y^*, x^*) = o(\varepsilon) \text{ and } f^2(y^*, \tilde{x}_\varepsilon) - f^2(y^*, x^*) = o(\varepsilon) \text{ as } \varepsilon \to 0 \ .$$

*Moreover*

$$\forall x_\varepsilon \in \mathcal{S}_\varepsilon(y_\varepsilon) \qquad f(y^*, x^*) - f(y_\varepsilon, x_\varepsilon) = o(\varepsilon) \text{ as } \varepsilon \to 0 \ .$$

*Proof* - The first assertion has been proved : relations (4.16) and (4.17) . We use relation (4.4) and the previous result to claim that

$$f^2(y^*, x^*) - f^2(y_\varepsilon, x_\varepsilon) = o(\varepsilon) \ .$$

As $f^2(y^*, x^*) - f^2(y_\varepsilon, x_\varepsilon) = [f(y^*, x^*) + f(y_\varepsilon, x_\varepsilon)]\,[f(y^*, x^*) - f(y_\varepsilon, x_\varepsilon)]$ and $f(y^*, x^*) + f(y_\varepsilon, x_\varepsilon) \to 2f(y^*, x^*) = 2\beta^*$ we get the result since $\beta^* \neq 0$. □

With a bootstrapping technique we obtain the following corollary ;

**Corollary 4.1** *Under the assumptions and notations of the previous theorem, we get* $\forall x_\varepsilon \in \mathcal{S}_\varepsilon(y_\varepsilon)$

$$\forall n \in \mathbb{N} \qquad f(y^*, x^*) - f(y_\varepsilon, x_\varepsilon) = o(\varepsilon^n)$$

*and* $\forall \tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(y^*)$

$$h(y^*, \tilde{x}_\varepsilon) - h(y^*, x^*) = o(\varepsilon^n) \ .$$

*Proof* - Using relations (4.16) and (4.17) in assumption $(\mathcal{H}_1)$ we see that relation (4.10) becomes

$$\forall \varepsilon > 0,\ \forall \tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(y^*) \qquad \exists x^*_\varepsilon \in C^* \text{ s.t. } \|\tilde{x}_\varepsilon - x^*_\varepsilon\| \leq \gamma o(\varepsilon) \ . \tag{4.18}$$

Using the same technique leads to relations (4.16) and (4.17) with $\varepsilon^2$ instead of $\varepsilon$ and so on. □

### 4.3 Error estimates under a "second-order" assumption

If assumption $(\mathcal{H}_1)$ is not ensured, one may, however, give estimates if the following hypothesis is satisfied

$$(\mathcal{H}_2) \qquad \left\{ \begin{array}{c} \exists \varepsilon_o > 0 \ , \exists \delta > 0, \ \text{ such that } \forall x \in C^* + \mathcal{B}(0, \varepsilon_o) \\ \exists \tilde{x} \in C^* \text{ such that } \|x - \tilde{x}\|^2 \leq \delta \left[ (h(y^*, x) - \alpha^*)^+ + (f(y^*, x) - \beta^*)^+ \right] \end{array} \right.$$

**Remark 4.2** $(\mathcal{H}_2)$ *means that* $C^*$ *is H-metrically regular (of the second order). (See the definition of this regularity property for example in [1] Def. 4.3.2).* $(\mathcal{H}_2)$ *also corresponds to a quadratic growth condition [4] Def.3.1 .*
*This assumption is significantly weaker than* $(\mathcal{H}_1)$ *and covers a large class of problems since it is satisfied when* $h(y^*, .) + f(y^*, .)$ *is linear or quadratic.*

We have a rather similar result which proof is the the same as in the previous subsection (so that we do not detail it) :

**Theorem 4.2** *Assume that* $(\mathcal{H}_2)$ *is satisfied ; let* $y_\varepsilon$ *be a solution to* $(\mathcal{P}_\varepsilon)$ *and* $x_\varepsilon \in \mathcal{S}_\varepsilon(y_\varepsilon)$. *Then*

$$f(y^*, x^*) - f(y_\varepsilon, x_\varepsilon) = o(\sqrt{\varepsilon}) \text{ as } \varepsilon \to 0 \text{ , so that}$$

$$\forall \tau > 0 \qquad f(y^*, x^*) - f(y_\varepsilon, x_\varepsilon) = o(\varepsilon^{1-\tau}) \text{ .}$$

*Moreover,* $\forall \tilde{x}_\varepsilon \in \mathcal{S}_\varepsilon(y^*)$

$$\forall \tau > 0 \qquad h(y^*, \tilde{x}_\varepsilon) - h(y^*, x^*) = o(\varepsilon^{1-\tau}) \text{ .}$$

# 5 Conclusion

With this new penalization approach we are able to "solve" more general bi-level problems as usual ones. In fact we do not solve the limit problem of course but the approximated one. We have given error estimates that proved that the approximation is reasonable. Next step is the numerical realization and the comparison with current methods

# References

[1] A. Auslender and M. Teboulle. Asymptotic cones and functions in optimization and variational inequalities. *Springer Monographs in Mathematics*, Springer-Verlag, New York, 2003.

[2] D. Azé and J.N Corvellec. Characterizations of error bounds for lower semicontinuous functions on metric spaces. *submitted*, 2003.

[3] D. Azé and A. Rahmouni. On primal-dual stability in convex optimization. *submitted*, 2003.

[4] J.F. Bonnans and A. Shapiro. Perturbation analysis of optimization problems. *Springer Series in Operations Research*, Springer-Verlag, New York, 2000.

[5] M.Bouhtou, M. Diallo, and L. Wynter. Fair network resource allocation and link pricing: a numerical study. *Optimization and optimal control (Ulaanbaatar, 2002)*, 37–58, Ser. Comput. Oper. Res., 1, World Sci. Publishing, River Edge, NJ, 2003.

[6] L. Brotcorne, M. Labbé, P. Marcotte and G. Savard. A bilevel model for toll optimization on a multicommodity transportation network, *Transportation Science*, 35 (2001) 1-14

[7] P. Chardaire and A. Lisser. Simplex and interior point specialized algorithms for solving non-oriented multicommodity flow problems. *Operation Res*, 50 (2002), no. 2, 260–276.

[8] M. Labbé, P. Marcotte and G. Savard. On a class of bilevel programs. *Nonlinear optimization and related topics*, (Erice, 1998), 183–206, Appl. Optim., 36, Kluwer Acad. Publ., Dordrecht, 2000.

[9] A.S. Lewis and J-S. Pang. Error bounds for convex inequality systems. *Generalized convexity, generalized monotonicity: recent results (Luminy, 1996)*, 75–110, Nonconvex Optim. Appl., 27, Kluwer Acad. Publ., Dordrecht, 1998.

[10] W. Li. Abadie's constraint qualification, metric regularity, and error bounds for differentiable convex inequalities. *SIAM Journal of Optimization*, 7(4):966–978, 1997.

[11] A.Ouorou, P. Mahey and J.P. Vial. A survey of algorithms for convex multicommodity flow problems *Management Science*, vol.46, 1 (2000), pp.126-147.

[12] P. Marcotte et G. Savard. A bilevel programming approach to Price Setting in: Decision and Control in Management Science. *Essays in Honor of Alain Haurie* Georges Zaccour ed., Kluwer Academic Publishers (2002) 97 - 117

[13] K.F. Ng and X.Y. Zheng. Global error bounds with fractional exponents. *Mathematical programming*, Ser B(88):357–370, 2000.

[14] T. Q. Nguyen, M. Bouhtou and J.-L. Lutton. D.C. approach to bilevel bilinear programming problem: application in telecommunication pricing. *Optimization and optimal control (Ulaanbaatar, 2002)*, 211–231, Ser. Comput. Oper. Res., 1, World Sci. Publishing, River Edge, NJ, 2003.

[15] Z. Wu and J. Ye. On error bounds for lower semicontinuous functions. *Mathematical Programming*, Ser. A, 2002.

[16] J. Zhao. The lower semicontinuity of optimal solution sets. *Journal of Mathematical Analysis and Applications*, 207:240–254, 1997.