AN EXACT PENALTY APPROACH FOR MATHEMATICAL PROGRAMS WITH EQUILIBRIUM CONSTRAINTS.

L. Abdallah[1] and M. Haddou[2]

Abstract. We propose an exact penalty approach to solve the mathematical problems with equilibrium constraints (MPECs). This work is based on the smoothing functions introduced in [3] but we do not need any complicate updating rule for the penalty parameter. We present some numerical results to prove the viability of the approach. We consider two generic applications : the binary quadratic programs and simple number partitioning problems.

Keywords: Optimization, nonlinear programming, exact penalty function.

1. Introduction

Mathematical programs with equilibrium constraints (MPECs) represent an optimization problem including a set of parametric variational inequality or complementary constraints. In this paper, we consider optimization problems with complementary constraints, in the following form

$$(1) \qquad (P) \quad \begin{cases} f^* = \min f(x,y) \\ < x.y >= 0 \\ (x,y) \in D \end{cases}$$

where $f : \mathbb{R}^{2n} \to \mathbb{R}$ and $D = [0, v]^{2n}$. $< . >$ denotes the inner product on $\mathbb{R}^n$.
We made this choice for D only to simplify the exposition. We an use only compact set.
MPECs problems are know to be very difficult. They arise in numerous areas and applications. The aim of this paper is to develop some simple heuristic to obtain a local solution. The main difficulty relies in the complementarity constraints so that a direct use of standard non linear programming algorithms generally fails. Many regularization and relaxations techniques have already been proposed in the literature [2, 5]. In this study, we proposed smoothing technique to regularize the complementary constraints based on [3], we replace each constraint

$$x_i y_i = 0$$

by

$$\theta_\varepsilon(x_i) + \theta_\varepsilon(y_i) \leq 1$$

where the functions $\theta_\varepsilon : \mathbb{R}_+ \to [0, 1]$ are at least C^2 and satisfy :

$$\theta_\varepsilon(x) \begin{cases} \simeq 1 & \text{if} \quad x \neq 0 \\ = 0 & \text{if} \quad x = 0 \end{cases}$$

The main difference of our approach is that we do not need any complicate strategy to update the parameter ε since we will consider it as some new optimization variable. This paper is organized as follows. In section 2, we present some preliminaries and assumptions on the smoothing functions. In Section 3, we consider a smooth constrained optimization problem and introduce a penalty function. We

[1]Université Libanaise, Laboratoire de Mathématiques et Applications (LaMA), Tripoli, Liban. E-mail: lina_abdallah@hotmail.fr

[2]MAPMO, UMR 6628, Univ. Orléans, France. E-mail: mounir.haddou@univ-orleans.fr

prove under some mild assumptions an existence result for the approximate problem and an exact penalty property. In section 4, we present numerical experiments concerning academic MPECs of small sizes. The last section presents a large set of numerical experiments considering binary quadratic programs and simple number partitioning problems.

2. Preliminaries

In this section, we present some preliminaries concerning the regularization and approximation process. We consider functions $\theta_\varepsilon(\varepsilon > 0)$ with the following properties:

(1) θ_ε is nondecreasing, strictly concave and continuously differentiable,
(2) $\forall \varepsilon > 0,\ \theta_\varepsilon(0) = 0$,
(3) $\forall x > 0,\ \lim\limits_{\varepsilon \to 0} \theta_\varepsilon(x) = 1$,
(4) $\lim\limits_{\varepsilon \to 0} \theta'_\varepsilon(0) > 0$,
(5) $\exists\, m > 0, \exists \varepsilon_0 > 0\, \forall x \in [0, v], \forall \varepsilon \in]0, \varepsilon_0], |\partial_\varepsilon \theta_\varepsilon(x)| \leq \dfrac{m}{\varepsilon^2}$

For $\varepsilon = 0$, we set $\theta_0(0) = 0$ and $\theta_0(x) = 1,\ \forall\, x \neq 0$.

Examples of such functions are:

$$
\begin{array}{ll}
(\theta^1_\varepsilon): & \theta_\varepsilon(x) = \dfrac{x}{x+\varepsilon} \\
(\theta^{w_1}_\varepsilon): & \theta_\varepsilon(x) = (1 - e^{-(\frac{x}{\varepsilon})})^k,\ \text{for } k \leq 1 \\
(\theta^{log}_\varepsilon): & \theta_\varepsilon(x) = \dfrac{\log(1+x)}{\log(1+x+\varepsilon)}
\end{array}
$$

Using function θ_ε, we obtain the relaxed following problem :

$$
(2) \qquad (P_\varepsilon) \quad \begin{cases} f^*_\varepsilon = \min f(x,y) \\ \theta_\varepsilon(x_i) + \theta_\varepsilon(y_i) \leq 1, \quad i = 1, \ldots, n \\ (x,y) \in D \end{cases}
$$

Remark 2.1. $< x.y >= 0 \Rightarrow \forall \varepsilon > 0,\ \theta_\varepsilon(x_i) + \theta_\varepsilon(y_i) \leq 1$. *Thus any feasible point for* (P_ε) *is also feasible for* (P) *and then* $\forall \varepsilon > 0, f^*_\varepsilon \leq f^*$.

We first transform the inequality constraints into equality constraints, by introducing some slacks variables e_i:

$$
(3) \qquad \theta_\varepsilon(x_i) + \theta_\varepsilon(y_i) + e_i - 1 = 0, \quad e_i \geq 0 \quad i = 1, \ldots, n.
$$

The problem (P_ε) becomes:

$$
(4) \qquad (\widetilde{P}_\varepsilon) \quad \begin{cases} \min f(x,y) \\ \theta_\varepsilon(x_i) + \theta_\varepsilon(y_i) + e_i - 1 = 0 \quad i = 1, \ldots, n \\ (x,y,e) \in D \times [0,1]^n \end{cases}
$$

Indeed each e_i can not exceed 1.
The limit problem $(\widetilde{P}_\varepsilon)$ for $\varepsilon = 0$

$$
(5) \qquad (\widetilde{P}) \quad \begin{cases} \min f(x,y) \\ \theta_0(x_i) + \theta_0(y_i) + e_i = 1, \quad i = 1, \ldots, n \\ e_i \in [0,1], \quad i = 1, \ldots, n \end{cases}
$$

which is equivalent to (P).

Until now, this relaxation process was introduced in [3]. To avoid the updating of parameters problem, we define the penalty functions f_σ on $D \times [0,1] \times [0,\overline{\varepsilon}]$:

$$f_\sigma(x,y,e,\varepsilon) = \begin{cases} f(x,y) & \text{if} \quad \varepsilon = \Delta(x,y,e,\varepsilon) = 0; \\ f(x,y) + \dfrac{1}{2\varepsilon}\Delta(x,y,e,\varepsilon) + \sigma\beta(\varepsilon) & \text{if} \quad \varepsilon > 0, \\ +\infty & \text{if} \quad \varepsilon = 0 \text{ and } \Delta(x,y,e,\varepsilon) \neq 0 \end{cases}$$

where Δ measures the feasibility violation and the function $\beta : [0,\overline{\varepsilon}] \to [0,\infty)$ is continuously differentiable on $(0,\overline{\varepsilon}]$ with $\beta(0) = 0$. $\Delta(z,\varepsilon) = \|G_\varepsilon(z)\|^2$ where $(G_\varepsilon(z))_i = (\theta_\varepsilon(x) + \theta_\varepsilon(y) + e - 1)_i$ and $z = (x,y,e)$.

Remark 2.2. $\forall z \in D^{'}, \Delta(z,0) = 0 \Leftrightarrow z$ *feasible for* $\widetilde{P} \Leftrightarrow (x,y)$ *feasible for* (P).

Then we consider the following problem:

$$(6) \qquad (P_\sigma) \qquad \begin{cases} \min f_\sigma(x,y,e,\varepsilon) \\ (x,y,e,\varepsilon) \in D \times [0,1]^n \times [0,\overline{\varepsilon}] \end{cases}$$

From now on, we will denote

$$(7) \qquad D^{'} = D \times [0,1]^n$$

Definition 2.1. *We say that the Mangasarian-Fromovitz condition* [7] *for* P_σ *holds at* $z \in D^{'}$ *if* $G^{'}_\varepsilon(z)$ *has full rank and there exists a vector* $p \in \mathbb{R}^n$ *such that* $G^{'}_\varepsilon(z)p = 0$ *and*

$$p_i \quad \begin{cases} > 0 & \text{if} \quad z_i = 0 \\ < 0 & \text{if} \quad z_i = w_i \end{cases}$$

with

$$w_i = \begin{cases} v & \text{if} \quad i \in \{1 \dots 2n\} \\ 1 & \text{if} \quad i \in \{2n+1 \dots 3n\} \end{cases}$$

Remark 2.3. *This regularity condition can be replaced by one of those proposed in* [8].

3. The smoothing technique

The following theorem yields a condition to find a solution for (P_σ). It also proves a direct link to (P):

Theorem 3.1. *we suppose we suppose that* $z \in D^{'}$ *satisfies the Mangasarian-Fromovitz condition, and that*

$$\beta^{'}(\varepsilon) \geq \beta_1 > 0 \ \text{ for } \ 0 < \varepsilon < \overline{\varepsilon}.$$

i) If σ *is sufficiently large, there is no KKT point of* P_σ *with* $\varepsilon > 0$.
ii) For σ *sufficiently large, every local minimizer* $(z^*,\varepsilon^*), (z^* = (x^*,y^*,e^*))$ *of the problem* (P_σ) *has the form* $(z^*,0)$*, where* (x^*,y^*) *is a local minimizer of the problem* (P).

Proof:
$i)$ Let (z,ε) a Kuhn Tucker point of P_σ, then there exist λ and $\mu \in \mathbb{R}^{3n+1}$ such that:

$$(8) \qquad \begin{array}{ll} \text{(i)} & \nabla \ell(z,\varepsilon) = \nabla f_\sigma(z,\varepsilon) + \lambda - \mu = 0 \\ \text{(ii)} & \min(\lambda, z_i) = \min(\mu, w_i - z_i) = 0, \ i = 1 \dots 3n \\ \text{(iii)} & \mu_{3n+1} = \min(\lambda_{3n+1}, \overline{\varepsilon} - \varepsilon) = 0, \end{array}$$

where ∇f_σ is the gradient of f_σ with respect to (z,ε).
Assume that there exists a sequence of KKT points (z_k,ε_k) of P_{σ_k} with $\varepsilon_k \neq 0, \forall k$

and $\lim_{k\to+\infty} \sigma_k = +\infty$.
Since $D^{'}$ is bounded and closed, up to a subsequence, we have

$$\lim_{k\to+\infty} \varepsilon_k = \varepsilon^*$$
$$\lim_{k\to+\infty} z_k = z^*$$

(9.i) yields to $\partial_\varepsilon f_{\sigma_k}(z_k,\varepsilon_k) + \lambda_{3n+1} - \mu_{3n+1} = 0$. So that $\partial_\varepsilon f_{\sigma_k}(z_k,\varepsilon_k) \leq 0$.
Then, if we denote $\Delta_k = \Delta(z_k,\varepsilon_k)$, we have

$$\begin{aligned} \partial_\varepsilon f_{\sigma_k} &= -\frac{1}{4\varepsilon_k^2}\Delta_k + \frac{1}{2\varepsilon_k}\partial_\varepsilon\Delta_k + \sigma_k\beta^{'}(\varepsilon_k) \\ &= -\frac{1}{4\varepsilon_k^2}\Delta_k + \frac{1}{\varepsilon_k}(\theta_\varepsilon(x_k)+\theta_\varepsilon(y_k)+e_k+1)(\partial_\varepsilon\theta_\varepsilon(x_k)+\partial_\varepsilon\theta_\varepsilon(y_k)) + \sigma_k\beta^{'}(\varepsilon_k) \leq 0 \end{aligned}$$

Multiplying by $4\varepsilon^3$, we obtain

$$4\varepsilon_k^2(\theta_\varepsilon(x_k)+\theta_\varepsilon(y_k)+e_k-1)(\partial_\varepsilon\theta_\varepsilon(x_k)+\partial_\varepsilon\theta_\varepsilon(y_k)) + 4\varepsilon_k^3\sigma_k\beta^{'}(\varepsilon_k) \leq \varepsilon_k\Delta_k$$

Since $\Delta_k, \theta_\varepsilon$ and $\varepsilon^2\partial_\varepsilon\theta_\varepsilon$ are bounded (by definition (v)), $\sigma_k \to \infty$ when $k \to \infty$. We have $\varepsilon^* = 0$.
(ii) Let σ sufficiently large and (z^*,ε^*) a local minimizer for (P_σ). If (z^*,ε^*) satisfies the Magasarian-Fromovitz condition, then (z^*,ε^*) is a Kuhn-Tucker points for f_σ. By (i), we conclude that $\varepsilon^* = 0$.

Let $\mathcal{V}$ a neighborhood of $(z^*,0)$, for any z feasible for $\widetilde{P}$ such that $(z,0) \in \mathcal{V}$ we have

$$f_\sigma(z^*,0) \leq f_\sigma(z,0) = f(x,y) < +\infty \tag{9}$$

(since $\Delta(z,0)=0$).
We can conclude that $\Delta(z^*,0) = 0$, otherwise $f_\sigma(z^*,0)$ would be $+\infty$. So that $< x^*, y^* >= 0$ and (x^*,y^*) is a feasible point of (P).

Back to (9) $f(x^*,y^*) = f_\sigma(z^*,0) \leq f_\sigma(z,0) = f(x,y)$.
Therefore (x^*,y^*) is a local minimizer for (P). □

Remark 3.1. *The previous theorem is still valid if we consider penalty functions of the form*

$$f_\sigma(x,y,e,\sigma) = f(x,y) + \alpha(\varepsilon)\Delta(x,y,e,\varepsilon) + \sigma\beta(\varepsilon) \tag{10}$$

with $\alpha(\varepsilon) > \frac{1}{2\varepsilon}$.

4. Numerical results

In this section we consider some preliminary results obtained with the approach described in the previous section. We used the SNOPT solver [6] for the solution on the AMPL optimization platform [1]. In all our tests, we take $\beta(\varepsilon) := \sqrt{\varepsilon}$.
We consider various MPECs where the optimal value is know [4]. Tables 1 and 2 summarizes our different informations concerning the computational effort of the SNOPT, by using respectively θ^{w_1} and θ_1 function:

- Obj.value : is the optimal value
- it : correspond to the total number of iterations
- (Obj.) and (grad.) : correspond to the total number of objective function evaluations and objective function gradient evaluation
- (constr.) and (jac.) : give respectively the total number of constraints and constraints gradient evaluation.

Problem	Data	Obj.val.	it	Obj.	grad	constr.	Jac
bard1	no	17	331	193	192		
desilva	$(0,0)$	-1	892	655	656	655	656
	$(2,2)$	-1	448	416	415	416	415
Df1	no	$3.26e-12$	3	28	27	28	27
Bilevel1	$(25,25)$	5	470	214	213		
	$(50,50)$	5	295	168	169		
Bilevel2	$(0,0,0,0)$	-6600	232	55	54		
	$(0,5,0,20)$	-6600	180	56	55		
	$(5,0,15,10)$	-6600	-6599.9	331	97	96	
flp4	flp4-1.dat	$1.9e-29$	66	9	8		
	flp4-2.dat	$3.08e-29$	66	9	8		
	flp4-3.dat	$1.1e-28$	66	9	8		
gauvin	no	0	184	71	70		
jr1	no	0.5	1175	814	813		
scholtes1	1	2	12	10	9	10	9
hs044	no	14.97	375	101	100		
nash1	$(0,0)$	0	0	2	1		
	$(5,5)$	$1.72e-17$	13	13	12		
	$(10,10)$	$2.24e-12$	12	12	11		
	$(10,0)$	$4.29e-12$	12	11	10		
	$(0,10)$	$1.46e-13$	13	12	11		
qpec1	no	80	1249	443	442		
liswet1-inv	liswet1-050	0.0929361	215	126	125		
Stack1	0	-3266.67	27	26			
	100	-3266.67	7	17	16		
	200	-3266.67	7	17	16		
Water-net	Water-net.dat	931.1	2070	886	885	886	885

TABLE 1. using the θ^{w_1} function

Problem	Data	Obj.val.	it	Obj.	grad	constr.	Jac
bard1	no	17	433	248	247		
desilva	$(0,0)$	-1	7	255	254	255	254
Df1	no	0	657	961	960	961	960
gauvin	no	$9.5e-05$	164	82	81	82	81
Bilevel1	$(25,25)$	5	401	190	198		
	$(50,50)$	5	391	183	182		
Bilevel2	$(0,0,0,0)$	-6600	2458	487	486	487	486
	$(0,5,0,20)$	-6600	2391	727	721		
	$(5,0,15,10)$	-6600	2391	727	721		
hs044	no	17.08	617	261	260	261	260
jr1	no	0.5	67	54	53		
nash1	$(0,0)$	$3.35e-13$	203	111	110		
	$(5,5)$	$6.7e-24$	146	71	70		
	$(10,10)$	$2.3e-17$	133	85	84		
	$(10,0)$	$8.1e-16$	379	238	237		
	$(0,10)$	$2.37e-18$	1228	848	847		
qpec1	no	80	1895	518	517		
liswet1-inv	liswet1-050	0.028	3559	462	461		
scholtes1	1	2	51	106	105	106	105
Stack1	0	-3266.67	64	58	57		
	100	-3266.67	30	32	31		
	200	-3266.67	30	32	31		
Water-net	Water-net.dat	931.369	919	282	281	282	281

TABLE 2. using the θ^1 function

We remark that by considering θ^{w_1} or θ^1 we obtain the optimal know value in almost all the considered test problems.

5. Application to simple partitioning problem and binary quadratic problems

In this section, we consider two real applications : the simple number partitioning and binary quadratic problems. These two classes of problems are know to be

NP hard. We propose here a simple heuristic to obtain local solutions.

5.1. **Application to simple partitioning problem.** The number partitioning problem can be stated as a quadratic binary problem. We model this problem as follows.
We consider a set of numbers $S = \{s_1, s_2, s_3, \ldots, s_m\}$. The goal is to partition S into two subsets such that the subset sums are as close to each other as possible. Let $x_j = 1$ if s_j is assigned to subset 1, 0 otherwise. Then sum_1, subset 1's sum, is $\mathrm{sum}_1 = \sum_{j=1}^{m} s_j x_j$ and the sum for subset 2 is $\mathrm{sum}_2 = \sum_{j=1}^{m} s_j - \sum_{j=1}^{m} s_j x_j$. The difference in the sums is then given by

$$\mathrm{diff} = \sum_{j=1}^{m} s_j - 2\sum_{j=1}^{m} s_j x_j = c - 2\sum_{j=1}^{m} s_j x_j. \qquad (c = \sum_{j=1}^{m} s_j)$$

We will minimize the square of this difference

$$\mathrm{diff}^2 := \{c - 2\sum_{j=1}^{m} s_j x_j\}^2,$$

We can rewrite diff^2 as

$$\mathrm{diff}^2 = c^2 + 4x^T Q x,$$

where

$$q_{ii} = s_i(s_i - c), \quad q_{ij} = s_i s_j.$$

Dropping the additive and multiplicative constants, our optimization problem becomes simply

$$UQP \begin{cases} \min \ x^T Q x \\ x \in \{0,1\}^n \end{cases}$$

We rewrite the problem as the follows:

$$UQP \begin{cases} \min \ x^T Q x \\ x.(1-x) = 0 \end{cases}$$

We can now, use the proposed algorithm to get some local solutions for (UQP).

The results reported here on modest-sized random problems of size $m = 25$ and $m = 75$. An instance of each size are considered with the element drawn randomly from the interval $(50, 100)$.

Each of the problems was solved by our approach, using the two functions θ_ε^1 and θ_ε^w. We present in the table $3, 4$ the number of optimal solution obtained with 100 different initial points generated randomly from the interval $[0, 1]$:

- Best sum diff : corresponds to the best value of $|\sum_{i=1}^{100}(Q*round(x[i]) - 0.5*c|$
- Integrality measure : correspond to the $\max_i |round(x_i) - x_i|$
- Nb: correspond to the number of tests such that the best sum is satisfied.
- Nb_{10} : correspond to the number of tests such that the sum : $|\sum_{i=1}^{100}(Q*round(x[i]) - 0.5*c| \leq 10$

Problem	Best sum diff (θ^1, θ^{w_1})	Nb (θ^1, θ^{w_1})	Integrality measure (θ^1, θ^{w_1})	Nb_{10} (θ^1, θ^{w_1})
$NP25.1$	$(1, 0)$	$(1, 2)$	$(0.011, 0)$	$(15, 15)$
$NP25.2$	$(0, 0)$	$(2, 2)$	$(0.0055, 0.005)$	$(16, 14)$
$NP25.3$	$(0, 0)$	$(1, 1)$	$(0, 0)$	$(16, 14)$
$NP25.4$	$(0, 0)$	$(1, 2)$	$(0, 0)$	$(22, 22)$
$NP25.5$	$(0, 0)$	$(1, 4)$	$(0.008, 0.0045)$	$(11, 10)$
$NP75.1$	$(0, 0)$	$(1, 2)$	$(0.003, 0)$	$(14, 14)$
$NP75.2$	$(0, 0)$	$(2, 1)$	$(0, 0)$	$(15, 15)$
$NP75.3$	$(0, 0)$	$(1, 1)$	$(0, 0)$	$(17, 17)$
$NP75.4$	$(0, 0)$	$(2, 2)$	$(0, 0)$	$(18, 18)$
$NP75.5$	$(0, 1)$	$(1, 1)$	$(0, 0)$	$(17, 17)$

TABLE 3. using the θ^1 and θ^{w_1} function

5.2. **Application to binary quadratic problems.** We consider some test problems from the Biq Mac Library [9]. These problems are written in the simple following formulation:

$$\min y^T Q y$$
$$y \in \{0, 1\}^n$$

where Q is a symmetric matrix of order n. For the Q matrix, ten instances have been generated. The parameters are the following:

- diagonal coefficients in the range $[-100, 100]$
- off-diagonal coefficients in the range $[-50, 50]$,
- seeds $1, 2, \ldots, 10$.

We apply the technique described in section 2. We present in the table 5 the number of optimal solution obtained with 100 different initial points (Nbop) generated randomly from the interval $[0, 1]$, and for a size matrix equal 100. The fourth column precise the obtained value when different to know optimal value.

Problem	Know. value	$Nbop\,(\theta^1, \theta^{w_1})$	Found value (θ^1, θ^{w_1})
$be100.1$	-19412	$(17, 14)$	
$be100.2$	-17290	$(14, 12)$	
$be100.3$	-17565	$(9, 13)$	
$be100.4$	-19125	$(9, 14)$	
$be100.5$	-15868	$(2, 2)$	
$be100.6$	-17368	$(31, 31)$	
$be100.7$	-18629	$(0, 0)$	$(-18473, -18475)$
$be100.8$	-18649	$(1, 1)$	
$be100.9$	-13294	$(0, 0)$	$(-13248, -13248)$
$be100.10$	-15352	$(11, 4)$	

TABLE 4. using the θ^1 and θ^{w_1} functions

Using θ^{w_1} or θ^1 we obtain the optimal know value in almost of our tests. We obtain a local solutions for only two examples. For each instance, the algorithm found an optimal solution and needs < 1 s for the resolution.

6. CONCLUSION

In this paper, we have introduced an exact penalty approach to solve the mathematical program with equilibrium constraints.
We have proved a convergence result under suitable constraint qualification conditions. We performed a numerical computation by applying our approach to some tests from the library MacMPEC. Then, we considered some examples from the Biq Mac Library and some randomly generated partitioning problems. We used two different smoothing functions and our limited numerical tests gave almost the same result for each one.

REFERENCES

[1] AMPL.: A Modeling Language for Mathematical Programming. http://www.ampl.com
[2] Birbil, S.I., Fang, S-H., and Han, J.: An entropic approach for mathematical programs with equilibrium constraints. Computer and Operations Research, 31
[3] Haddou, M.: A new class of smoothing methods for mathematical programs with equilibrium constraints, Pacific Journal of Optimization, Vol. 5, pp. 87-95 (2009)
[4] MacMPEC.: Ampl collection of Mathematical Programs with Equilibrium Constraints. http://www.unix.mcs.anl.gov/leyffer/MacMPEC
[5] Facchinei, F., Jiang, H., and Qi, L. A smoothing method for mathematical programs with equilibrium constraints. Mathematical Programming, Vol. 85, pp. 81-106 (1995)
[6] Gill, P., Murray, W., Saunders, M.: SNOPT, A large-scale smooth optimization problems having linear or nonlinear objectives and constraints. http://www-neos.mcs.anl.gov/neos/solvers
[7] Mangasarian, O.L., Fromovitz, S.: The Fritz John necessary optimality conditions in the presence of equality and inequality constraints, J. Math. Anal. Appl., 17(1967), pp.37-47.
[8] Ralph, D. and S. J. Wright: Some properties of regularization and penalization schemee for MPECS, Springer-Verlag, New York, 2000.
[9] Wiegele A.: Biq Mac Library - A collection of Max-Cut and quadratic 0-1 programming instances of medium size (2007)