DIPARTIMENTO DI INFORMATICA
E SISTEMISTICA ANTONIO RUBERTI

SAPIENZA
UNIVERSITÀ DI ROMA

Exact Penalty Functions for Nonlinear Integer Programming Problems

Stefano Lucidi
Francesco Rinaldi

Technical Report n. 10, 2009

Exact Penalty Functions for Nonlinear Integer Programming Problems

S. Lucidi, F. Rinaldi

Dipartimento di Informatica e Sistemistica
Sapienza Università di Roma
Via Ariosto, 25 - 00185 Roma - Italy

e-mail: stefano.lucidi@dis.uniroma1.it
e-mail: rinaldi@dis.uniroma1.it

Abstract

In this work, we study exact continuous reformulations of nonlinear integer programming problems. To this aim, we preliminarily state conditions to guarantee the equivalence between pairs of general nonlinear problems. Then, we prove that optimal solutions of a nonlinear integer programming problem can be obtained by using various exact penalty formulations of the original problem in a continuous space.

Keywords nonlinear integer programming, continuous programming, exact penalty functions.

1 Introduction

Many real world problems can be formulated as a nonlinear minimization problem where some (or all) of the variables only assume integer values. When the dimensions of the problem get large, finding an optimal solution becomes a tough task. A reasonable approach can be that of transforming the original problem into an equivalent continuous problem. A number of different transformations have been proposed in the literature (see e.g. [1, 2, 5, 8, 9, 10, 11, 12]).
In this work, we consider a particular continuous reformulation which comes out by relaxing the integer constraints on the variables and by adding a penalty term to the objective function. This approach was first described by Ragavachari in [12] to solve zero-one linear programming problems. There are many other papers closely related to the one by Ragavachari (see e.g. [3, 4, 6, 7, 13, 14]). In [4], the exact penalty approach has been extended to general nonlinear integer programming problems. In [13], various penalty terms have been proposed for solving zero-one concave programming problems. We generalize the results described in [4], and we show that a general class of penalty functions, including the ones proposed in [13], can be used for solving general nonlinear integer problems.
In Section 2, we state a general result concerning the equivalence between an unspecified optimization problem and a parameterized family of problems. In Section 3, by using the general results described in Section 2, we prove that a specific class of penalty terms can be used to define exact equivalent continuous reformulations of a general zero-one programming problem. In Section 4, following the idea of Section 3, we show a general nonlinear integer programming problem is equivalent to a continuous penalty problem. The results proposed in Section 3 and 4 can be easily extended to mixed integer programming problems.

2 A General Equivalence Result using Penalization

We start from the general nonlinear constrained problem:

$$\min_{x\in W} f(x) \tag{1}$$

where $W \subset R^n$ and $f(x) : R^n \to R$.
For any $\varepsilon \in R_+$, we consider the following problem:

$$\min_{x\in X} f(x) + \varphi(x,\varepsilon). \tag{2}$$

where $W \subseteq X \subset R^n$, and $\varphi(\cdot,\varepsilon) : R^n \to R$.
In the following Theorem we show that, under suitable assumptions on f and φ, Problem (1) and (2) are equivalent.

Theorem 1 *Let W and X be compact sets. Let $\|\cdot\|$ be a suitably chosen norm. We assume that*

a) f is bounded on X, and there exists an open set $A \supset W$ and real numbers $\alpha, L > 0$ such that, for any $x, y \in A$, f satisfies the following condition:

$$|f(x) - f(y)| \le L\|x - y\|^{\alpha}. \tag{3}$$

b) the function φ satisfies the following:

(i) For every $x, y \in W$, and every $\varepsilon \in R_+$

$$\varphi(x, \varepsilon) = \varphi(y, \varepsilon).$$

(ii) There exist a value $\hat{\varepsilon}$ and an open set $S \supset W$ such that, for every $z \in W$, $x \in S \cap (X \setminus W)$ and $\varepsilon \in (0, \hat{\varepsilon}]$, we have

$$|\varphi(x, \varepsilon) - \varphi(z, \varepsilon)| \geq \hat{L}\|x - z\|^{\alpha} \tag{4}$$

where $\hat{L} > L$.
Furthermore, there exists a point $\bar{x} \notin S$ such that

$$\lim_{\varepsilon \to 0} [\varphi(\bar{x}, \varepsilon) - \varphi(z, \varepsilon)] = \infty \tag{5}$$

for every $z \in W$, and

$$\varphi(x, \varepsilon) \geq \varphi(\bar{x}, \varepsilon). \tag{6}$$

for every $x \in X \setminus S$, and for every $\varepsilon > 0$;

Then, a real value $\tilde{\varepsilon}$ exists such that, for any $\varepsilon \in (0, \tilde{\varepsilon}]$, problem (2) and problem (1) have the same minimum points.

Proof. First we prove that every optimal solution of Problem (2) is also an optimal solution of Problem (1).

For all $\varepsilon > 0$ if $x^\star$ is an optimal solution of problem (2) we have

$$f(x^\star) + \varphi(x^\star, \varepsilon) \leq f(x) + \varphi(x, \varepsilon) \quad \forall\, x \in X. \tag{7}$$

Since $W \subseteq X$ it follows that

$$f(x^\star) + \varphi(x^\star, \varepsilon) \leq f(z) + \varphi(z, \varepsilon) \quad \forall\, z \in W. \tag{8}$$

If $x^\star \in W$, assumption (i) ensures that

$$f(x^\star) \leq f(z) \quad \forall\, z \in W,$$

which shows that $x^\star$ is a global minimum of Problem (1).

Now we prove that there exists a value $\tilde{\varepsilon}$ such that, for all $\varepsilon \in (0, \tilde{\varepsilon}]$, every global minimum point of Problem (2) belongs to the set W.

Let $\bar{x}$ and S be respectively the point and the open set defined in Assumption (ii). Hence, by (5), there exists a value $\bar{\varepsilon}$ such that for all $\varepsilon \in (0, \bar{\varepsilon}]$ the following inequality holds:

$$\varphi(\bar{x}, \varepsilon) - \varphi(z, \varepsilon) > \sup_{x \in W} f(x) - \inf_{x \in X \setminus S} f(x). \tag{9}$$

Then we can introduce the value $\tilde{\varepsilon}$ as follows

$$\tilde{\varepsilon} = \min\{\bar{\varepsilon}, \hat{\varepsilon}\} \tag{10}$$

where $\hat{\varepsilon}$ is defined as in (ii).

Now, suppose, by contradiction, that for a value $\varepsilon \in (0, \tilde{\varepsilon}]$ there exists a global minimum of Problem (2) $x^\star$ which does not belong to W, namely $x^\star \notin W$.
We consider two different cases:

1) $x^\star \in S$:
without any loss of generality, consider $S \subseteq A$. In this case for any $z \in W$, using the definition of $\hat{\varepsilon}$, assumption $a)$ and (ii), we obtain

$$f(z) - f(x^\star) \leq |f(z) - f(x^\star)| \leq L\|x^\star - z\|^\alpha < \hat{L}\|x^\star - z\|^\alpha \leq \varphi(x^\star, \varepsilon) - \varphi(z, \varepsilon) \qquad (11)$$

and we get the contradiction

$$f(z) + \varphi(z, \varepsilon) < f(x^\star) + \varphi(x^\star, \varepsilon). \qquad (12)$$

2) $x^\star \notin S$:
in this case we have that $x^\star \in X \setminus S$ and, recalling Assumption (ii), by using (6) we can write for any $z \in W$:

$$\begin{aligned} f(x^\star) + \varphi(x^\star, \varepsilon) \quad & \geq \inf_{x \in X \setminus S} f(x) + \varphi(x^\star, \varepsilon) \\ & \geq f(z) - \sup_{x \in W} f(x) + \inf_{x \in X \setminus S} f(x) + \varphi(x^\star, \varepsilon) \\ & \geq f(z) - \sup_{x \in W} f(x) + \inf_{x \in X \setminus S} f(x) + \varphi(\bar{x}, \varepsilon), \end{aligned} \qquad (13)$$

adding and subtracting $\varphi(z, \varepsilon)$ we write

$$f(x^\star) + \varphi(x^\star, \varepsilon) \geq f(z) + \varphi(z, \varepsilon) + \varphi(\bar{x}, \varepsilon) - \varphi(z, \varepsilon) - \sup_{x \in W} f(x) + \inf_{x \in X \setminus S} f(x).$$

Recalling definition of $\tilde{\varepsilon}$ and (9), for all $\epsilon \in (0, \tilde{\varepsilon}]$ we obtain the contradiction:

$$f(x^\star) + \varphi(x^\star, \varepsilon) > f(z) + \varphi(z, \varepsilon). \qquad (14)$$

Now we prove that, for all $\varepsilon \in (0, \tilde{\varepsilon}]$ (where $\tilde{\varepsilon}$ is defined as in (10)), every optimal solution of Problem (1) is also an optimal solution of Problem (2).
Suppose, by contradiction, that there exists an $\varepsilon \in (0, \tilde{\varepsilon}]$ such that

$$f(x^\star) + \varphi(x^\star, \varepsilon) < f(z^\star) + \varphi(z^\star, \varepsilon), \qquad (15)$$

where $z^\star$ is an optimal solution of Problem (1) and $x^\star$ is an optimal solution of Problem (2).
Recalling the first part of the proof, we have that, for all $\varepsilon \in (0, \tilde{\varepsilon}]$, the point $x^\star$ is also a optimal point of Problem (1) and, hence, using assumption (i), we have

$$f(x^\star) < f(z^\star) \qquad (16)$$

and this contradicts the fact that $z^\star$ is an optimal solution of problem (1).□

3 Smooth Penalty Functions for Solving Zero-one Programming Problems

We consider the following problem

$$\begin{aligned} & \min f(x) \\ & x \in T \\ & x \in \{0,1\}^n \end{aligned} \qquad (17)$$

where $T \subseteq R^n$, and f is a function satisfying assumption $a)$ of Theorem 1.
Our aim consists in showing that the zero-one problem (17) is equivalent to the following continuous formulation:

$$\begin{array}{l} \min f(x) + \varphi(x, \varepsilon) \\ x \in T \\ 0 \leq x \leq e \end{array} \tag{18}$$

where $\varepsilon > 0$, and $\varphi(x, \varepsilon)$ is a suitably chosen penalty term.
In [4], the equivalence between (17) and (18) has been proved for

$$\varphi(x, \varepsilon) = \frac{1}{\varepsilon} \sum_{i=1}^{n} x_i(1 - x_i). \tag{19}$$

In this section, by using Theorem 1, we can prove the equivalence between (17) and (18) for a more general class of penalty terms including (19).
In particular, the penalty terms we consider are:

$$\varphi(x, \varepsilon) = \sum_{i=1}^{n} \{\log(x_i + \varepsilon) + \log[(1 - x_i) + \varepsilon]\} \tag{20}$$

$$\varphi(x, \varepsilon) = \sum_{i=1}^{n} \{-(x_i + \varepsilon)^{-p} - [(1 - x_i) + \varepsilon]^{-p}\} \tag{21}$$

$$\varphi(x, \varepsilon) = \frac{1}{\varepsilon} \sum_{i=1}^{n} \Big\{ \Big[1 - \exp(-\alpha \cdot x_i)\Big] + \Big[1 - \exp(-\alpha \cdot (1 - x_i))\Big] \Big\} \tag{22}$$

$$\varphi(x, \varepsilon) = \frac{1}{\varepsilon} \sum_{i=1}^{n} \{(x_i + \varepsilon)^q + [(1 - x_i) + \varepsilon]^q\} \tag{23}$$

$$\varphi(x, \varepsilon) = \frac{1}{\varepsilon} \sum_{i=1}^{n} \Big\{ \Big[1 + \exp(-\alpha \cdot x_i)\Big]^{-1} + \Big[1 + \exp(-\alpha \cdot (1 - x_i))\Big]^{-1} \Big\} \tag{24}$$

where ε, α, $p > 0$ and $0 < q < 1$. Functions (20)-(23) have been proposed in [13], where the equivalence between (17) and (18) has been proved in the case when f is a concave objective function and T is a polyhedral set. The use of penalty term (24) in formulation (18) has never been proposed before.
We set

$$W = \Big\{ x \in T \;:\; x \in \{0, 1\}^n \Big\}$$

and

$$X = \Big\{ x \in T \;:\; 0 \leq x \leq e \Big\}.$$

Proposition 1 *For every penalty term* (20)-(24)*, there exists a value* $\bar{\varepsilon} > 0$ *such that, for any* $\varepsilon \in (0, \bar{\varepsilon}]$*, problem* (18) *and problem* (17) *have the same minimum points.*

Proof. As we assumed that function f satisfies assumption $a)$ of Theorem 1, the proof can be derived by showing that every penalty term (20)-(24) satisfies assumption $b)$ of Theorem 1.

Consider the penalty term (20).
For any $x \in \{0,1\}^n$ we have

$$\varphi(x,\varepsilon) = n \cdot \log[\varepsilon \cdot (1+\varepsilon)]$$

and (i) is satisfied.
We study the behavior of the i-th function $\varphi_i(x_i,\varepsilon)$ in a neighborhood of a feasible point z_i. We can consider three different cases:

1. $z_i = 0$ and $0 < x_i < \rho$: Using the mean theorem we obtain

$$\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) = \left(\frac{1}{\tilde{x}_i+\varepsilon} - \frac{1}{1-\tilde{x}_i+\varepsilon}\right)|x_i - z_i| \tag{25}$$

where $\tilde{x}_i \in (0, x_i)$. Choosing $\rho < \frac{1}{2}$, we have

$$\begin{aligned} \varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) &\geq \left(\frac{1}{\rho+\varepsilon} - \frac{1}{1-\rho+\varepsilon}\right)|x_i - z_i| \quad (26)\\ &\geq \left(\frac{1}{\rho+\varepsilon} - 2\right)|x_i - z_i| \quad (27)\end{aligned}$$

Choosing ρ and ε such that

$$\rho + \varepsilon \leq \frac{1}{\tilde{L}+2}, \tag{28}$$

we obtain

$$\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) \geq \tilde{L}|x_i - z_i|. \tag{29}$$

2. $z_i = 1$ and $1-\rho < x_i < 1$: Using the mean theorem we obtain

$$\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) = \left(\frac{1}{1-\tilde{x}_i+\varepsilon} - \frac{1}{\tilde{x}_i+\varepsilon}\right)|x_i - z_i|. \tag{30}$$

Then, repeating the same reasoning as in case 1, we have again that (29) holds when ρ and ε satisfy (28).

3. $z_i = x_i = 0$, or $z_i = x_i = 1$: We have $\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) = 0$.

We can conclude that, when ρ and ε satisfy (28),

$$\varphi(x,\varepsilon) - \varphi(z,\varepsilon) \geq \tilde{L}\sum_{i=1}^{n}|x_i - z_i| = \tilde{L}\|x-z\|_1 \geq \tilde{L}\|x-z\|_\infty \tag{31}$$

for all $z \in \{0,1\}^n \cap T$ and for all x such that $\|x-z\|_\infty < \rho$.
Now we define $S(z) = \{x \in R^n : \|x-z\|_\infty < \rho\}$ and $S = \bigcup_{i=1}^{N} S(z_i)$, where N is the number of points $z \in \{0,1\}^n \cap T$, and (4) holds.
Let $\bar{x}$ be a point such that $\bar{x}_j = \rho$ ($\bar{x}_j = 1-\rho$), and $\bar{x}_i \in \{0,1\}$ for all $i \neq j$. If $\{\varepsilon^k\}$ is an infinite sequence such that $\varepsilon^k \to 0$ for $k \to \infty$, we can write for each $z \in \{0,1\}^n$:

$$\lim_{k\to\infty}[\varphi(\bar{x},\varepsilon^k) - \varphi(z,\varepsilon^k)] = \lim_{k\to\infty}\left\{\log[(\rho+\varepsilon^k)\cdot(1-\rho+\varepsilon^k)] - \log[\varepsilon^k \cdot (1+\varepsilon^k)]\right\} = +\infty,$$

and (5) holds.
Then for every $x \in X \setminus S$, and for every $\varepsilon > 0$ we have

$$\begin{aligned}\varphi(x,\varepsilon) - \varphi(\bar{x},\varepsilon) = \quad & \sum_{i \neq \tilde{\jmath}} \Big\{ \log[(x_i+\varepsilon)\cdot(1-x_i+\varepsilon)] \Big\} - (n-1)\cdot \log[\varepsilon \cdot (1+\varepsilon)] \\ & + \log[(x_{\tilde{\jmath}}+\varepsilon)\cdot(1-x_{\tilde{\jmath}}+\varepsilon)] - \log[(\rho+\varepsilon)\cdot(1-\rho+\varepsilon)] \geq 0,\end{aligned}$$

where $\rho \leq x_{\tilde{\jmath}} \leq 1-\rho$. Then (6) holds, and Assumption (ii) is satisfied.

Consider the penalty term (21).
For any $x \in \{0,1\}^n$ we have

$$\varphi(x,\varepsilon) = -n \cdot [(\varepsilon)^{-p} + (1+\varepsilon)^{-p}]$$

and (i) is satisfied.
We study the behavior of the i-th function $\varphi_i(x_i,\varepsilon)$ in a neighborhood of a feasible point z_i. We can consider three different cases:

1. $z_i = 0$ and $0 < x_i < \rho$: Using the mean theorem we obtain

$$\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) = \left[\frac{p}{(\tilde{x}_i+\varepsilon)^{p+1}} - \frac{p}{(1-\tilde{x}_i+\varepsilon)^{p+1}} \right] |x_i - z_i| \tag{32}$$

   where $\tilde{x}_i \in (0, x_i)$. Choosing $\rho < \frac{1}{2}$, we have

$$\begin{aligned}\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) \quad & \geq \left[\frac{p}{(\rho+\varepsilon)^{p+1}} - \frac{p}{(1-\rho+\varepsilon)^{p+1}} \right] |x_i - z_i| & (33) \\ & \geq \left[\frac{p}{(\rho+\varepsilon)^{p+1}} - p \cdot 2^{p+1} \right] |x_i - z_i| & (34)\end{aligned}$$

   Choosing ρ and ε such that

$$(\rho+\varepsilon)^{p+1} \leq \frac{p}{\tilde{L} + p \cdot 2^{p+1}}, \tag{35}$$

   we obtain

$$\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) \geq \tilde{L}|x_i - z_i|. \tag{36}$$

2. $z_i = 1$ and $1-\rho < x_i < 1$: Using the mean theorem we obtain

$$\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) = \left[\frac{p}{(1-\tilde{x}_i+\varepsilon)^{p+1}} - \frac{p}{(\tilde{x}_i+\varepsilon)^{p+1}} \right] |x_i - z_i| \tag{37}$$

   Then, repeating the same reasoning as in case 1, we have again that (36) holds when ρ and ε satisfy (35).

3. $z_i = x_i = 0$, or $z_i = x_i = 1$: We have $\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) = 0$.

We can conclude that, when ρ and ε satisfy (35),

$$\varphi(x,\varepsilon)-\varphi(z,\varepsilon)\geq \tilde{L}\sum_{i=1}^{n}|x_i-z_i|=\tilde{L}\|x-z\|_1\geq \tilde{L}\|x-z\|_\infty \tag{38}$$

for all $z\in\{0,1\}^n\cap T$ and for all x such that $\|x-z\|_\infty<\rho$.
Now we define $S(z)=\{x\in R^n:\|x-z\|_\infty<\rho\}$ and $S=\bigcup_{i=1}^{N}S(z_i)$, where N is the number of points $z\in\{0,1\}^n\cap T$, and (4) holds.
Let $\bar{x}$ be a point such that $\bar{x}_j=\rho$ ($\bar{x}_j=1-\rho$), and $\bar{x}_i\in\{0,1\}$ for all $i\neq j$. If $\{\varepsilon^k\}$ is an infinite sequence such that $\varepsilon^k\to 0$ for $k\to\infty$, we can write for each $z\in\{0,1\}^n$:

$$\lim_{k\to\infty}[\varphi(\bar{x},\varepsilon^k)-\varphi(z,\varepsilon^k)]=\lim_{k\to\infty}\Big\{-(\rho+\varepsilon^k)^{-p}-[(1-\rho)+\varepsilon^k]^{-p}+[(\varepsilon^k)^{-p}+(1+\varepsilon^k)^{-p}]\Big\}=+\infty,$$

and (5) holds.
Then for every $x\in X\setminus S$, and for every $\varepsilon>0$ we have

$$\begin{aligned}\varphi(x,\varepsilon)-\varphi(\bar{x},\varepsilon)= \quad & \sum_{i\neq\tilde{\jmath}}\Big\{-(x_i+\varepsilon)^{-p}-[(1-x_i)+\varepsilon]^{-p}\Big\}+(n-1)\cdot[(\varepsilon)^{-p}+(1+\varepsilon)^{-p}]\\ & -(x_{\tilde{\jmath}}+\varepsilon)^{-p}-[(1-x_{\tilde{\jmath}})+\varepsilon]^{-p}+(\rho+\varepsilon)^{-p}+[(1-\rho)+\varepsilon]^{-p}\geq 0\end{aligned}$$

where $\rho\leq x_{\tilde{\jmath}}\leq 1-\rho$. Then (6) holds, and Assumption (ii) is satisfied.□

Consider the penalty term (22).
For any $x\in\{0,1\}^n$ we have

$$\varphi(x,\varepsilon)=\frac{n}{\varepsilon}\Big[1-\exp(-\alpha)\Big]$$

and (i) is satisfied.
We study the behavior of the i-th function $\varphi_i(x_i,\varepsilon)$ in a neighborhood of a feasible point z_i. We can consider three different cases:

1. $z_i=0$ and $0<x_i<\rho$: Using the mean theorem we obtain

$$\varphi_i(x_i,\varepsilon)-\varphi_i(z_i,\varepsilon)=\frac{\alpha}{\varepsilon}\Big[\exp(-\alpha\cdot\tilde{x}_i)-\exp(-\alpha\cdot(1-\tilde{x}_i))\Big]|x_i-z_i| \tag{39}$$

where $\tilde{x}_i\in(0,x_i)$. Choosing $\rho<\frac{1}{2}$, we have

$$\varphi_i(x_i,\varepsilon)-\varphi_i(z_i,\varepsilon)\geq\frac{\alpha}{\varepsilon}\Big[\exp(-\alpha\cdot\rho)-\exp(-\alpha\cdot(1-\rho))\Big]|x_i-z_i| \tag{40}$$

Choosing ρ and ε such that

$$\tilde{L}\leq\frac{\alpha}{\varepsilon}\Big[\exp(-\alpha\cdot\rho)-\exp(-\alpha\cdot(1-\rho))\Big], \tag{41}$$

we obtain

$$\varphi_i(x_i,\varepsilon)-\varphi_i(z_i,\varepsilon)\geq\tilde{L}|x_i-z_i|. \tag{42}$$

2. $z_i=1$ and $1-\rho<x_i<1$: Using the mean theorem we obtain

$$\varphi_i(x_i,\varepsilon)-\varphi_i(z_i,\varepsilon)=\frac{\alpha}{\varepsilon}\Big[\exp(-\alpha\cdot(1-\tilde{x}_i)-\exp(-\alpha\cdot\tilde{x}_i))\Big]|x_i-z_i| \tag{43}$$

Then, repeating the same reasoning as in case 1, we have again that (42) holds when ρ and ε satisfy (41).

3. $z_i = x_i = 0$, or $z_i = x_i = 1$: We have $\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = 0$.

We can conclude that, when ρ and ε satisfy (41),

$$\varphi(x, \varepsilon) - \varphi(z, \varepsilon) \geq \tilde{L} \sum_{i=1}^{n} |x_i - z_i| = \tilde{L}\|x - z\|_1 \geq \tilde{L}\|x - z\|_\infty \tag{44}$$

for all $z \in \{0,1\}^n \cap T$ and for all x such that $\|x - z\|_\infty < \rho$.
Now we define $S(z) = \{x \in R^n : \|x - z\|_\infty < \rho\}$ and $S = \bigcup_{i=1}^{N} S(z_i)$, where N is the number of points $z \in \{0,1\}^n \cap T$, and (4) holds.
Let $\bar{x}$ be a point such that $\bar{x}_j = \rho$ ($\bar{x}_j = 1 - \rho$), and $\bar{x}_i \in \{0,1\}$ for all $i \neq j$. If $\{\varepsilon^k\}$ is an infinite sequence such that $\varepsilon^k \to 0$ for $k \to \infty$, we can write for each $z \in \{0,1\}^n$:

$$\lim_{k\to\infty} [\varphi(\bar{x}, \varepsilon^k) - \varphi(z, \varepsilon^k)] = \lim_{k\to\infty} \frac{1}{\varepsilon^k} \left\{ \Big[1 - \exp(-\alpha \cdot \rho)\Big] + \Big[1 - \exp(-\alpha \cdot (1-\rho))\Big] - \Big[1 - \exp(-\alpha)\Big] \right\} = \infty,$$

and (5) holds.
Then for every $x \in X \setminus S$, and for every $\varepsilon > 0$ we have

$$\begin{aligned}\varphi(x, \varepsilon) - \varphi(\bar{x}, \varepsilon) = \quad & \frac{1}{\varepsilon} \sum_{i \neq \tilde{j}} \Big\{ \Big[1 - \exp(-\alpha \cdot x_i)\Big] + \Big[1 - \exp(-\alpha \cdot (1 - x_i))\Big] \Big\} - (n-1) \cdot \Big[1 - \exp(-\alpha)\Big] \\ & + \Big\{ \Big[1 - \exp(-\alpha \cdot x_{\tilde{j}})\Big] + \Big[1 - \exp(-\alpha \cdot (1 - x_{\tilde{j}}))\Big] \\ & - \Big[1 - \exp(-\alpha) \cdot \rho\Big] - \Big[1 - \exp(-\alpha \cdot (1 - \rho))\Big] \Big\} \geq 0\end{aligned}$$

where $\rho \leq x_{\tilde{j}} \leq 1 - \rho$. Then (6) holds, and Assumption (ii) is satisfied.$\Box$

Consider the penalty term (23).
For any $x \in \{0,1\}^n$ we have

$$\varphi(x, \varepsilon) = n \cdot [(\varepsilon)^q + (1 + \varepsilon)^q]$$

and (i) is satisfied.
We study the behavior of the i-th function $\varphi_i(x_i, \varepsilon)$ in a neighborhood of a feasible point z_i. We can consider three different cases:

1. $z_i = 0$ and $0 < x_i < \rho$: Using the mean theorem we obtain

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = \left[\frac{q}{(\tilde{x}_i + \varepsilon)^{1-q}} - \frac{q}{(1 - \tilde{x}_i + \varepsilon)^{1-q}} \right] |x_i - z_i| \tag{45}$$

   where $\tilde{x}_i \in (0, x_i)$. Choosing $\rho < \frac{1}{2}$, we have

$$\begin{aligned}\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \quad & \geq \left[\frac{q}{(\rho + \varepsilon)^{1-q}} - \frac{q}{(1 - \rho + \varepsilon)^{1-q}} \right] |x_i - z_i| & (46) \\ & \geq \left[\frac{q}{(\rho + \varepsilon)^{1-q}} - q \cdot 2^{1-q} \right] |x_i - z_i| & (47)\end{aligned}$$

   Choosing ρ and ε such that

$$(\rho + \varepsilon)^{1-q} \leq \frac{q}{\tilde{L} + q \cdot 2^{1-q}}, \tag{48}$$

   we obtain

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq \tilde{L}|x_i - z_i|. \tag{49}$$

2. $z_i = 1$ and $1-\rho < x_i < 1$: Using the mean theorem we obtain

$$\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) = \left[\frac{q}{(1-\tilde{x}_i+\varepsilon)^{1-q}} - \frac{q}{(\tilde{x}_i+\varepsilon)^{1-q}}\right]|x_i - z_i| \tag{50}$$

Then, repeating the same reasoning as in case 1, we have again that (49) holds when ρ and ε satisfy (48).

3. $z_i = x_i = 0$, or $z_i = x_i = 1$: We have $\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) = 0$.

We can conclude that, when ρ and ε satisfy (48),

$$\varphi(x,\varepsilon) - \varphi(z,\varepsilon) \geq \tilde{L}\sum_{i=1}^{n}|x_i - z_i| = \tilde{L}\|x-z\|_1 \geq \tilde{L}\|x-z\|_\infty \tag{51}$$

for all $z \in \{0,1\}^n \cap T$ and for all x such that $\|x-z\|_\infty < \rho$.
Now we define $S(z) = \{x \in R^n : \|x-z\|_\infty < \rho\}$ and $S = \bigcup_{i=1}^{N} S(z_i)$, where N is the number of points $z \in \{0,1\}^n \cap T$, and (4) holds.
Let $\bar{x}$ be a point such that $\bar{x}_j = \rho$ ($\bar{x}_j = 1-\rho$), and $\bar{x}_i \in \{0,1\}$ for all $i \neq j$. If $\{\varepsilon^k\}$ is an infinite sequence such that $\varepsilon^k \to 0$ for $k \to \infty$, we can write for each $z \in \{0,1\}^n$:

$$\lim_{k\to\infty}[\varphi(\bar{x},\varepsilon^k) - \varphi(z,\varepsilon^k)] = \lim_{k\to\infty}\Big\{(\rho+\varepsilon^k)^q + [(1-\rho)+\varepsilon^k]^q - [(\varepsilon^k)^q + (1+\varepsilon^k)^q]\Big\} = +\infty,$$

and (5) holds.
Then for every $x \in X \setminus S$, and for every $\varepsilon > 0$ we have

$$\begin{aligned}\varphi(x,\varepsilon) - \varphi(\bar{x},\varepsilon) = \quad & \sum_{i\neq\bar{\jmath}}\Big\{(x_i+\varepsilon)^q + [(1-x_i)+\varepsilon]^q\Big\} - (n-1)\cdot[(\varepsilon)^q + (1+\varepsilon)^q] \\ & +(x_{\bar{\jmath}}+\varepsilon)^q + [(1-x_{\bar{\jmath}})+\varepsilon]^q - (\rho+\varepsilon)^q - [(1-\rho)+\varepsilon]^q \geq 0\end{aligned}$$

where $\rho \leq x_{\bar{\jmath}} \leq 1-\rho$. Then (6) holds, and Assumption (ii) is satisfied.□

Consider the penalty term (24).
For any $x \in \{0,1\}^n$ we have

$$\varphi(x,\varepsilon) = \frac{n}{\varepsilon}\Big\{0.5 + \Big[1+\exp(-\alpha)\Big]^{-1}\Big\}$$

and (i) is satisfied.
We study the behavior of the i-th function $\varphi_i(x_i,\varepsilon)$ in a neighborhood of a feasible point z_i. We can consider three different cases:

1. $z_i = 0$ and $0 < x_i < \rho$: Using the mean theorem we obtain

$$\begin{aligned}\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) = \quad & \frac{\alpha}{\varepsilon}\Big\{[1+\exp(-\alpha\cdot\tilde{x}_i)]^{-2}\cdot\exp(-\alpha\cdot\tilde{x}_i) \\ & -[1+\exp(-\alpha\cdot(1-\tilde{x}_i))]^{-2}\cdot\exp(-\alpha\cdot(1-\tilde{x}_i))\Big\}|x_i - z_i|\end{aligned} \tag{52}$$

where $\tilde{x}_i \in (0, x_i)$. Choosing $\rho < \frac{1}{2}$, we have

$$\begin{aligned}\varphi_i(x_i,\varepsilon) - \varphi_i(z_i,\varepsilon) \quad & \geq \frac{\alpha}{\varepsilon}\Big\{[1+\exp(-\alpha\cdot\rho)]^{-2}\cdot\exp(-\alpha\cdot\rho) \\ & -[1+\exp(-\alpha\cdot(1-\rho))]^{-2}\cdot\exp(-\alpha\cdot(1-\rho))\Big\}|x_i - z_i| \\ & \geq \frac{\alpha}{\varepsilon}\Big[0.5\cdot\exp(-\alpha\cdot\rho) - \exp(-\alpha\cdot(1-\rho))\Big]|x_i - z_i|\end{aligned} \tag{53}$$

Choosing ρ and ε such that

$$\tilde{L} \leq \frac{\alpha}{\varepsilon}\Big[0.5 \cdot \exp(-\alpha \cdot \rho) - \exp(-\alpha \cdot (1-\rho))\Big], \tag{54}$$

we obtain

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq \tilde{L}|x_i - z_i|. \tag{55}$$

2. $z_i = 1$ and $1 - \rho < x_i < 1$: Using the mean theorem we obtain

$$\begin{aligned} \varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \quad &= \frac{\alpha}{\varepsilon}\Big\{[1 + \exp(-\alpha \cdot (1 - \tilde{x}_i))]^{-2} \cdot \exp(-\alpha \cdot (1 - \tilde{x}_i)) \\ &\quad -[1 + \exp(-\alpha \cdot \tilde{x}_i)]^{-2} \cdot \exp(-\alpha \cdot \tilde{x}_i)\Big\}|x_i - z_i| \end{aligned} \tag{56}$$

   Then, repeating the same reasoning as in case 1, we have again that (55) holds when ρ and ε satisfy (54).

3. $z_i = x_i = 0$, or $z_i = x_i = 1$: We have $\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = 0$.

We can conclude that, when ρ and ε satisfy (54),

$$\varphi(x, \varepsilon) - \varphi(z, \varepsilon) \geq \tilde{L}\sum_{i=1}^{n} |x_i - z_i| = \tilde{L}\|x - z\|_1 \geq \tilde{L}\|x - z\|_\infty \tag{57}$$

for all $z \in \{0,1\}^n \cap T$ and for all x such that $\|x - z\|_\infty < \rho$.
Now we define $S(z) = \{x \in R^n : \|x - z\|_\infty < \rho\}$ and $S = \bigcup_{i=1}^{N} S(z_i)$, where N is the number of points $z \in \{0,1\}^n \cap T$, and (4) holds.
Let $\bar{x}$ be a point such that $\bar{x}_j = \rho$ ($\bar{x}_j = 1 - \rho$), and $\bar{x}_i \in \{0,1\}$ for all $i \neq j$. If $\{\varepsilon^k\}$ is an infinite sequence such that $\varepsilon^k \to 0$ for $k \to \infty$, we can write for each $z \in \{0,1\}^n$:

$$\begin{aligned} \lim_{k\to\infty}[\varphi(\bar{x}, \varepsilon^k) - \varphi(z, \varepsilon^k)] \quad &= \lim_{k\to\infty} \frac{1}{\varepsilon^k}\ \Big\{\Big[1 + \exp(-\alpha \cdot \rho)\Big]^{-1} \\ &+\Big[1 + \exp(-\alpha \cdot (1-\rho))\Big]^{-1} - 0.5 - \Big[1 + \exp(-\alpha)\Big]^{-1}\Big\} = \infty, \end{aligned}$$

and (5) holds.
Then for every $x \in X \setminus S$, and for every $\varepsilon > 0$ we have

$$\begin{aligned} \varphi(x, \varepsilon) - \varphi(\bar{x}, \varepsilon) = \quad & \frac{1}{\varepsilon}\ \sum_{i \neq \tilde{\jmath}}\Big\{\Big[1 + \exp(-\alpha \cdot x_i)\Big]^{-1} + \Big[1 + \exp(-\alpha \cdot (1 - x_i))\Big]^{-1}\Big\} \\ & -(n-1) \cdot \Big\{0.5 + \Big[1 + \exp(-\alpha)\Big]^{-1}\Big\} \\ & +\Big\{\Big[1 + \exp(-\alpha \cdot x_{\tilde{\jmath}})\Big]^{-1} + \Big[1 + \exp(-\alpha \cdot (1 - x_{\tilde{\jmath}}))\Big]^{-1}\Big\} \\ & -\Big\{\Big[1 + \exp(-\alpha \cdot \rho)\Big]^{-1} + \Big[1 + \exp(-\alpha \cdot (1 - \rho))\Big]^{-1}\Big\} \geq 0 \end{aligned}$$

where $\rho \leq x_{\tilde{\jmath}} \leq 1 - \rho$. Then (6) holds, and Assumption (ii) is satisfied.□

4 Smooth Penalty Functions for Solving Integer Programming Problems

In this section we consider the following problem

$$\begin{aligned} &\min f(x) \\ &x \in T \\ &x \in D = D_1 \times \ldots \times D_n \end{aligned} \tag{58}$$

where f is a function satisfying assumption $a)$ of Theorem 1, T is a compact set, and

$$D_i = \{d_j \in Z,\ j = 1, \ldots, m_{D_i}\}. \tag{59}$$

It is well known (see i.e. [4]) that Problem (58) can be reformulated as a zero-one programming problem by using the following representation for the integer variables:

$$x_i = \sum_{k=0}^{M} y_k^{(i)} \cdot 2^k \quad y_k^{(i)} \in \{0, 1\}, \quad i = 1, \ldots, n \tag{60}$$

where M is an upper integer bound for $\log x_i$. This approach can be troublesome, especially when dealing with problems having sets D_i not uniformly distributed in Z. In order to face this type of problems, we propose a different approach that directly penalizes the constraints $x_i \in D_i$. Once again, by using Theorem 1, we prove the equivalence between (58) and the following continuous penalty formulation:

$$\begin{aligned} &\min f(x) + \varphi(x, \varepsilon) \\ &x \in T, \end{aligned} \tag{61}$$

where the penalty term can assume different forms. An example of such penalty terms is the following:

$$\varphi(x, \varepsilon) = \sum_{i=1}^{n} \min_{d_j \in D_i} \Big\{ \log[|x_i - d_j| + \varepsilon] \Big\} \tag{62}$$

Proposition 2 *For the penalty term* (62)*, there exists a value* $\bar{\varepsilon} > 0$ *such that, for any* $\varepsilon \in (0, \bar{\varepsilon}]$*, problem* (61) *and problem* (58) *have the same minimum points.*

Proof. As we assumed that function f satisfies assumption $a)$ of Theorem 1, the proof can be derived by showing that penalty term (62) satisfy assumption $b)$ of Theorem 1.

Consider the penalty term (62).
For any $x \in D$ we have

$$\varphi(x, \varepsilon) = n \cdot \log \varepsilon$$

and (i) is satisfied.
We study the behavior of the i-th function $\varphi_i(x_i, \varepsilon)$ in a neighborhood of a feasible point z_i. We can consider three different cases:

1. $z_i = d_j$ and $d_j < x_i < d_j + \rho$: Choosing ρ sufficiently small, and using the mean theorem we obtain

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = \frac{1}{(\tilde{x}_i - d_j) + \varepsilon}|x_i - z_i| \tag{63}$$

where $\tilde{x}_i \in (d_j, x_i)$. Then, we have

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq \frac{1}{\rho + \varepsilon}|x_i - z_i| \tag{64}$$

Choosing ρ and ε such that

$$\rho + \varepsilon \leq \frac{1}{\tilde{L}}, \tag{65}$$

we obtain

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq \tilde{L}|x_i - z_i|. \tag{66}$$

2. $z_i = d_j$ and $d_j - \rho < x_i < d_j$: Using the mean theorem we obtain

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = \frac{1}{(d_j - \tilde{x}_i) + \varepsilon}|x_i - z_i|. \tag{67}$$

Then, repeating the same reasoning as in case 1, we have again that (66) holds when ρ and ε satisfy (65).

3. $z_i = x_i = d_j$: We have $\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = 0$.

We can conclude that, when ρ and ε satisfy (65),

$$\varphi(x, \varepsilon) - \varphi(z, \varepsilon) \geq \tilde{L}\sum_{i=1}^{n}|x_i - z_i| = \tilde{L}\|x - z\|_1 \geq \tilde{L}\|x - z\|_\infty \tag{68}$$

for all $z \in T$ and for all x such that $\|x - z\|_\infty < \rho$.
Now we define $S(z) = \{x \in R^n : \|x - z\|_\infty < \rho\}$ and $S = \bigcup_{i=1}^{N} S(z_i)$, where N is the number of points $z \in D \cap T$, and (4) holds.
Let $\bar{x}$ be a point such that $\bar{x}_l = d_l \pm \rho$, with $d_l \in D_l$ and $\bar{x}_i \in D_i$ for all $i \neq l$. If $\{\varepsilon^k\}$ is an infinite sequence such that $\varepsilon^k \to 0$ for $k \to \infty$, we can write for each $z \in D$:

$$\lim_{k\to\infty}[\varphi(\bar{x}, \varepsilon^k) - \varphi(z, \varepsilon^k)] = \lim_{k\to\infty}\left\{\log(\rho + \varepsilon^k) - \log\varepsilon^k\right\} = +\infty,$$

and (5) holds.
Then for every $x \in X \setminus S$, and for every $\varepsilon > 0$ we have

$$\begin{aligned}\varphi(x, \varepsilon) - \varphi(\bar{x}, \varepsilon) = \quad & \sum_{i=1}^{n}\min_{d_j\in D_i}\log[|x_i - d_j| + \varepsilon] - \sum_{i=1}^{n}\min_{d_j\in D_i}\log[|\bar{x}_i - d_j| + \varepsilon] = \\ & \sum_{i\neq\tilde{l}}\left\{\min_{d_j\in D_i}\log[|x_i - d_j| + \varepsilon]\right\} - (n-1)\cdot\log\varepsilon \\ & + \log[|x_{\tilde{l}} - \bar{d}| + \varepsilon] - \log(\rho + \varepsilon) \geq 0,\end{aligned}$$

where $|x_{\tilde{l}} - \bar{d}| \geq \rho$ and

$$\bar{d} = \arg\min_{d_j\in D_{\tilde{l}}}\log[|x_{\tilde{l}} - d_j| + \varepsilon].$$

Then (6) holds, and Assumption (ii) is satisfied.□

Remark It is possible to introduce different types of penalty terms for Problem (58) by replacing in (62) the log function with the functions used in Section 3. Taking inspiration from equation (21), we have:

$$\varphi(x,\varepsilon) = \sum_{i=1}^{n} \min_{d_j \in D_i} \Big\{ -[|x_i - d_j| + \varepsilon]^{-p} \Big\} \tag{69}$$

In this case, the proof of the equivalence follows by repeating the same arguments used for proving Propositions 1 and 2.

Remark II Function (69) is equivalent to the following penalty term:

$$\varphi(x,\varepsilon) = \sum_{i=1}^{n} \min_{d_j \in D_i} \Big\{ \min\{-[x_i - d_j + \varepsilon]^{-p},\ -[d_j - x_i + \varepsilon]^{-p}\} \Big\}.$$

This penalty term should be easier to handle from a computational point of view.

References

[1] Abello, J., Butenko, S., Pardalos, P.M., Resende, M., *Finding independent sets in a graph using continuous multivariable polynomial formulations.* J. Glob. Optim. 21, 111137 (2001)

[2] Balasundaram, B., Butenko, S., *Constructing test functions for global optimization using continuous formulations of graph problems.* Optim. Methods Softw. 20, 439452 (2005)

[3] Borchardt M., *An Exact Penalty Approach for Solving a Class of Minimization Problems with Boolean Variables.* Optimization. 19 (6), pp. 829-838, 1988.

[4] Giannessi F., Niccolucci F., *Connections between nonlinear and integer programming problems.* Symposia Mathematica, Academic Press, New York , Vol. 19, pp. 161-176. 1976.

[5] Horst, R., Pardalos, P.M., Thoai, N.V., *Introduction to Global Optimization.* 2nd edn. Kluwer, Dordrecht (2000)

[6] Kalantari B., Rosen J.B., *Penalty Formulation for Zero-One Integer Equivalent Problem.* Mathematical Programming, Vol. 24, pp. 229-232, 1982.

[7] Kalantari B., Rosen J.B., *Penalty Formulation for Zero-One Nonlinear Programming.* Discrete Applied Mathematics, Vol. 16, No. 2, pp. 179-182, 1987.

[8] Mangasarian, O.L., *Knapsack Feasibility as an Absolute Value Equation Solvable by Successive Linear Programming.* Data Mining Institute Technical Report 08-03, September 2008. Optimization Letters, to appear.

[9] Murray W., Ng K. M., *An algorithm for nonlinear optimization problems with binary variables.* Computational Optimization and Applications, to appear.

[10] Pardalos P. M., Prokopyev O. A., Busygin S., *Continuous Approaches for Solving Discrete Optimization Problems.* Handbook on Modelling for Discrete Optimization, Springer US, Vol. 88, pp. 39-60, 2006.

[11] Pardalos, P.M., Wolkowicz, H. *Topics in Semidefinite and Interior-Point Methods.* Am. Math. Soc., Providence (1998)

[12] Ragavachari M., *On Connections Between Zero-One Integer Programming and Concave Programming Under Linear Constraints*, Operation Research Vol. 17, No. 4, pp. 680-684, 1969.

[13] Rinaldi F. *New results on the equivalence between zero-one programming and continuous concave programming*, Optimization Letters, to appear.

[14] Zhu W. X., *Penalty Parameter for Linearly Constrained 0-1 Quadratic Programming*, Journal of Optimization Theory and Applications, Vol. 116, No. 1, pp. 229-239, 2003.