Exact Penalty Functions for Nonlinear Integer Programming Problems

S. Lucidi · F. Rinaldi

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.

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. [1–8]).

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 [8] to solve zero-one linear programming problems. There are many other papers closely related

S. Lucidi (✉) · F. Rinaldi
Dipartimento di Informatica e Sistemistica "Antonio Ruberti", Sapienza University of Rome, Rome, Italy
e-mail: stefano.lucidi@dis.uniroma1.it

F. Rinaldi
e-mail: rinaldi@dis.uniroma1.it

to the one by Ragavachari (see e.g. [9–14]). In [10], 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 [10], 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 Sect. 2, we state a general result concerning the equivalence between an unspecified optimization problem and a parameterized family of problems. In Sect. 3, by using the general results described in Sect. 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 Sect. 4, following the idea of Sect. 3, we show that a general nonlinear integer programming problem is equivalent to a continuous penalty problem. The results proposed in Sects. 3 and 4 can be extended easily to mixed integer programming problems.

2 General Equivalence Result Using Penalization

We start from the general nonlinear constrained problem:

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

where $W \subset \mathbb{R}^n$ and $f(x) : \mathbb{R}^n \to \mathbb{R}$.

For any $\varepsilon \in \mathbb{R}_+$, we consider the following problem:

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

where $W \subseteq X \subset \mathbb{R}^n$, and $\varphi(\cdot, \varepsilon) : \mathbb{R}^n \to \mathbb{R}$. In (1), (2) and in the sequel, min denotes the global minimum.

In the following theorem, we show that, under suitable assumptions on f and φ, Problems (1) and (2) are equivalent.

Theorem 2.1 *Let W and X be compact sets. Let $\| \cdot \|$ be a suitably chosen norm. We make the following assumptions.*

(A1) *The function f is bounded on X and there exists an open set $A \supset W$ and real numbers $\alpha, L > 0$, such that, $\forall x, y \in A$, f satisfies the following condition*:

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

The function φ satisfies the following conditions:

(A2) $\forall x, y \in W$ *and* $\forall \varepsilon \in \mathbb{R}_+$,

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

(A3) *There exist a value $\hat{\varepsilon}$ and, $\forall z \in W$, there exists a neighbourhood $S(z)$ such that, $\forall x \in S(z) \cap (X \setminus W)$, and $\varepsilon \in]0, \hat{\varepsilon}]$, we have*

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

where $\hat{L} > L$ and α is chosen as in (3). *Furthermore, let* $S = \bigcup_{z \in W} S(z)$, $\exists \, \bar{x} \notin S$ *such that*

$$\lim_{\varepsilon \to 0} [\varphi(\bar{x}, \varepsilon) - \varphi(z, \varepsilon)] = +\infty, \quad \forall \, z \in W, \qquad (5)$$

$$\varphi(x, \varepsilon) \geq \varphi(\bar{x}, \varepsilon), \quad \forall \, x \in X \setminus S, \, \forall \, \varepsilon > 0. \qquad (6)$$

Then, $\exists \, \tilde{\varepsilon} \in \mathbb{R}$ *such that*, $\forall \, \varepsilon \in]0, \tilde{\varepsilon}]$, *Problems* (1) *and* (2) *have the same minimum points.*

Proof First, we prove that a minimum point of (2) is also a minimum point of (1). For all $\varepsilon > 0$, if $x^{\star}$ is a minimum point of (2), then we have

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

Since $W \subseteq X$, it follows that

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

If $x^{\star} \in W$, Assumption (A2) ensures that

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

which shows that $x^{\star}$ is a minimum point of (1). Now we prove that there exists a value $\tilde{\varepsilon}$ such that, $\forall \, \varepsilon \in]0, \tilde{\varepsilon}]$, every minimum point of (2) belongs to W. Let $\bar{x}$ and S be respectively the point and the open set defined in Assumption (A3). Hence, by (5), there exists a value $\bar{\varepsilon}$ such that, for all $\varepsilon \in]0, \bar{\varepsilon}]$, the following inequality holds:

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

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

$$\tilde{\varepsilon} := \min\{\bar{\varepsilon}, \hat{\varepsilon}\}, \qquad (10)$$

where $\hat{\varepsilon}$ is defined as in (A3). Ab absurdo, suppose that, for a value $\varepsilon \in]0, \tilde{\varepsilon}]$, there exists a minimum point of (2), say $x^{\star}$, such that $x^{\star} \notin W$. We consider two different cases:

Case (1): $x^{\star} \in S$. Without any loss of generality, consider $S \subseteq A$. In this case, there exists $z_i \in W$ such that $x^{\star} \in S(z_i)$. Using the definition of $\hat{\varepsilon}$, Assumptions (A1) and (A3), 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

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

Case (2): $x^\star \notin S$. In this case, we have that $x^\star \in X \setminus S$. Then,

$$f(x^\star) + \varphi(x^\star, \varepsilon) \geq \inf_{x \in X \setminus S} f(x) + \varphi(x^\star, \varepsilon). \tag{13}$$

$\forall z \in W$, we have $\sup_{x \in W} f(x) - f(z) \geq 0$. Then:

$$f(x^\star) + \varphi(x^\star, \varepsilon) \geq f(z) - \sup_{x \in W} f(x) + \inf_{x \in X \setminus S} f(x) + \varphi(x^\star, \varepsilon). \tag{14}$$

By using (6) of Assumption (A3), we obtain

$$f(x^\star) + \varphi(x^\star, \varepsilon) \geq f(z) - \sup_{x \in W} f(x) + \inf_{x \in X \setminus S} f(x) + \varphi(\bar{x}, \varepsilon). \tag{15}$$

Adding and subtracting $\varphi(z, \varepsilon)$ leads to

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

Finally, recalling the definition of $\tilde{\varepsilon}$ and exploiting (9), $\forall \varepsilon \in {]0, \tilde{\varepsilon}]}$ we obtain the contradiction

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

Now, we prove that, $\forall \varepsilon \in {]0, \tilde{\varepsilon}]}$, where $\tilde{\varepsilon}$ is defined as in (10), every minimum point of (1) is also a minimum point of (2). Ab absurdo, suppose that $\exists \varepsilon \in {]0, \tilde{\varepsilon}]}$ such that

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

where $z^\star$ is a minimum point of (1) and $x^\star$ is a minimum point of (2).

Recalling the first part of the proof, we have that, $\forall \varepsilon \in {]0, \tilde{\varepsilon}]}$, the point $x^\star$ is also a minimum point of (1); hence, using Assumption (A2), we have

$$f(x^\star) < f(z^\star),$$

which contradicts the fact that $z^\star$ is a minimum point of (1).   □

3 Smooth Penalty Functions for Solving Zero-one Programming Problems

We consider the following problem:

$$\min f(x), \quad \text{s.t.} \quad x \in T \cap \{0, 1\}^n, \tag{16}$$

where $T \subseteq \mathbb{R}^n$ and f is a function satisfying Assumption (A1) of Theorem 2.1.

Our aim consists in showing that the zero-one problem (16) is equivalent to the following continuous formulation:

$$\min f(x) + \varphi(x, \varepsilon), \quad x \in T, \ 0 \leq x \leq e, \tag{17}$$

where $\varepsilon > 0$ and $\varphi(x, \varepsilon)$ is a suitably chosen penalty term.

In [10], the equivalence between (16) and (17) has been proved for

$$\varphi(x, \varepsilon) = \frac{1}{\varepsilon} \sum_{i=1}^{n} x_i(1 - x_i). \tag{18}$$

In this section, by using Theorem 2.1, we can prove the equivalence between (16) and (17) for a more general class of penalty terms including (18).

In particular, the penalty terms that we consider are:

$$\varphi(x, \varepsilon) = \sum_{i=1}^{n} \{\log(x_i + \varepsilon) + \log[(1 - x_i) + \varepsilon]\}, \tag{19}$$

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

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

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

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

where $\varepsilon, \alpha, p > 0$ and $0 < q < 1$. The functions (19)–(22) have been proposed in [13], where the equivalence between (16) and (17) has been proved in the case when f is a concave objective function and T is a polyhedral set. The use of the penalty term (23) in the formulation (17) has never been proposed before.

We set

$$W = \{x \in T \ : \ x \in \{0, 1\}^n\}, \quad X = \{x \in T \ : \ 0 \le x \le e\}.$$

Proposition 3.1 *For every penalty term* (19)–(23), *there exists a value* $\bar{\varepsilon} > 0$ *such that, for any* $\varepsilon \in]0, \bar{\varepsilon}]$, *Problems* (16) *and* (17) *have the same minimum points.*

Proof As we assumed that the function f satisfies Assumption (A1) of Theorem 2.1, the proof can be derived by showing that every penalty term (19)–(23) satisfies Assumptions (A2) and (A3) of Theorem 2.1.

Consider the penalty term (19). For any $x \in \{0, 1\}^n$, we have

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

and Assumption (A2) 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

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = \left(\frac{1}{\tilde{x}_i + \varepsilon} - \frac{1}{1 - \tilde{x}_i + \varepsilon} \right) |x_i - z_i| \tag{24}$$

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

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq \left(\frac{1}{\rho + \varepsilon} - \frac{1}{1 - \rho + \varepsilon} \right) |x_i - z_i| \tag{25}$$

$$\geq \left(\frac{1}{\rho + \varepsilon} - 2 \right) |x_i - z_i|. \tag{26}$$

Choosing ρ and ε such that

$$\rho + \varepsilon \leq \frac{1}{\tilde{L} + 2}, \tag{27}$$

we obtain

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq \tilde{L}|x_i - z_i|. \tag{28}$$

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

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = \left(\frac{1}{1 - \tilde{x}_i + \varepsilon} - \frac{1}{\tilde{x}_i + \varepsilon} \right) |x_i - z_i|. \tag{29}$$

Then, repeating the same reasoning as in case 1, we have again that (28) holds when ρ and ε satisfy (27).
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 (27),

$$\varphi(x, \varepsilon) - \varphi(z, \varepsilon) \geq \tilde{L} \sum_{i=1}^{n} |x_i - z_i| = \tilde{L}\|x - z\|_1 \geq \tilde{L}\|x - z\|_\infty \tag{30}$$

for all $z \in \{0, 1\}^n \cap T$ and for all x such that $\|x - z\|_\infty < \rho$.

Now we define $S(z) = \{x \in \mathbb{R}^n : \|x - z\|_\infty < \rho\}$ 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 $z \in \{0, 1\}^n$:

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

$$= +\infty,$$

and (5) holds.

Then $\forall\, x \in X \setminus S$, and $\forall\, \varepsilon > 0$ we have

$$\varphi(x, \varepsilon) - \varphi(\bar{x}, \varepsilon) = \sum_{i \neq \tilde{j}} \{\log[(x_i + \varepsilon) \cdot (1 - x_i + \varepsilon)]\} - (n - 1) \cdot \log[\varepsilon \cdot (1 + \varepsilon)]$$

$$+ \log[(x_{\bar{j}} + \varepsilon) \cdot (1 - x_{\bar{j}} + \varepsilon)] - \log[(\rho + \varepsilon) \cdot (1 - \rho + \varepsilon)] \geq 0,$$

where $\rho \leq x_{\bar{j}} \leq 1 - \rho$. Then (6) holds, and Assumption (A3) is satisfied.

The proofs of the equivalence between (16) and (17) using the penalty terms (20)–(23) follow by repeating the same arguments used for proving the equivalence for the penalty term (19) (see [15] for further details). □

4 Smooth Penalty Functions for Solving Integer Programming Problems

In this section we consider the following problem

$$\min f(x), \quad \text{s.t.} \quad x \in T \cap D \tag{31}$$

where f is a function satisfying assumption (A1) of Theorem 2.1, T is a compact set, $D = D_1 \times \cdots \times D_n$, and

$$D_i = \{d_j \in Z, \ j = 1, \ldots, m_{D_i}\}. \tag{32}$$

It is well known (see i.e. [10]) that Problem (31) can be reformulated as a zero-one programming problem by using the following representation for the integer variables:

$$x_i = \sum_{k=0}^{M} y_k^{(i)} \cdot 2^k \quad y_k^{(i)} \in \{0, 1\}, \ i = 1, \ldots, n, \tag{33}$$

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 2.1, we prove the equivalence between (31) and the following continuous penalty formulation:

$$\min f(x) + \varphi(x, \varepsilon), \quad \text{s.t.} \quad x \in T, \tag{34}$$

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

$$\varphi(x, \varepsilon) = \sum_{i=1}^{n} \min_{d_j \in D_i} \{\log[|x_i - d_j| + \varepsilon]\}. \tag{35}$$

Proposition 4.1 *For the penalty term (35), there exists a value $\bar{\varepsilon} > 0$ such that, for any $\varepsilon \in]0, \bar{\varepsilon}]$, problem (31) and problem (34) have the same minimum points.*

Proof As we assumed that function f satisfies assumption (A1) of Theorem 2.1, the proof can be derived by showing that penalty term (35) satisfy assumptions (A2) and (A3) of Theorem 2.1.

Consider the penalty term (35).

$\forall\, x \in D$ we have

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

and (A2) 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

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = \frac{1}{(\tilde{x}_i - d_j) + \varepsilon} |x_i - z_i| \tag{36}$$

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

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq \frac{1}{\rho + \varepsilon} |x_i - z_i|. \tag{37}$$

Choosing ρ and ε such that

$$\rho + \varepsilon \leq \frac{1}{\tilde{L}}, \tag{38}$$

we obtain

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) \geq \tilde{L} |x_i - z_i|. \tag{39}$$

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

$$\varphi_i(x_i, \varepsilon) - \varphi_i(z_i, \varepsilon) = \frac{1}{(d_j - \tilde{x}_i) + \varepsilon} |x_i - z_i|. \tag{40}$$

Then, repeating the same reasoning as in case 1, we have again that (39) holds when ρ and ε satisfy (38).
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 (38),

$$\varphi(x, \varepsilon) - \varphi(z, \varepsilon) \geq \tilde{L} \sum_{i=1}^{n} |x_i - z_i| = \tilde{L} \|x - z\|_1 \geq \tilde{L} \|x - z\|_\infty \tag{41}$$

for all $z \in T$ and for all x such that $\|x - z\|_\infty < \rho$.

Now we define $S(z) = \{x \in \mathbb{R}^n : \|x - z\|_\infty < \rho\}$ 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 $z \in D$:

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

and (5) holds.

Then $\forall\, x \in X \setminus S$, and $\forall\, \varepsilon > 0$ we have

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

$$= \sum_{i \neq \bar{l}} \Big\{ \min_{d_j \in D_i} \log[|x_i - d_j| + \varepsilon] \Big\} - (n-1)\cdot \log\varepsilon$$

$$+ \log[|x_{\bar{l}} - \bar{d}| + \varepsilon] - \log(\rho + \varepsilon) \geq 0,$$

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

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

Then (6) holds, and Assumption (A3) is satisfied. □

Remark 4.1 It is possible to introduce different types of penalty terms for Problem (31) by replacing in (35) the log function with the functions used in Sect. 3. Taking inspiration from (20), we have:

$$\varphi(x,\varepsilon) = \sum_{i=1}^{n} \min_{d_j \in D_i} \{-[|x_i - d_j| + \varepsilon]^{-p}\}. \tag{42}$$

In this case, the proof of the equivalence follows by repeating the same arguments used for proving Propositions 3.1 and 4.1.

Remark 4.2 Function (42) is equivalent to the following penalty term:

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

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**, 111–137 (2001)
2. Balasundaram, B., Butenko, S.: Constructing test functions for global optimization using continuous formulations of graph problems. Optim. Methods Softw. **20**, 439–452 (2005)
3. Horst, R., Pardalos, P.M., Thoai, N.V.: Introduction to Global Optimization, 2nd edn. Kluwer, Dordrecht (2000)
4. Mangasarian, O.L.: Knapsack feasibility as an absolute value equation solvable by successive linear programming. Optim. Lett. **3**(2) (2009)
5. Murray, W., Ng, K.M.: An algorithm for nonlinear optimization problems with binary variables. Comput. Optim. Appl. (2010, to appear)
6. Pardalos, P.M., Prokopyev, O.A., Busygin, S.: Continuous Approaches for Solving Discrete Optimization Problems. Handbook on Modelling for Discrete Optimization, vol. 88, pp. 39–60. Springer, New York (2006)

7. Pardalos, P.M., Wolkowicz, H.: Topics in Semidefinite and Interior-Point Methods. Am. Math. Soc., Providence (1998)

8. Ragavachari, M.: On connections between zero-one integer programming and concave programming under linear constraints. Oper. Res. **17**(4), 680–684 (1969)

9. Borchardt, M.: An exact penalty approach for solving a class of minimization problems with Boolean variables. Optimization **19**(6), 829–838 (1988)

10. Giannessi, F., Niccolucci, F.: Connections between nonlinear and integer programming problems. In: Symposia Mathematica, vol. 19, pp. 161–176. Academic Press, New York (1976)

11. Kalantari, B., Rosen, J.B.: Penalty formulation for zero-one integer equivalent problem. Math. Program. **24**, 229–232 (1982)

12. Kalantari, B., Rosen, J.B.: Penalty formulation for zero-one nonlinear programming. Discrete Appl. Math. **16**(2), 179–182 (1987)

13. Rinaldi, F.: New results on the equivalence between zero-one programming and continuous concave programming. Optim. Lett. **3**(3), 377–386 (2009)

14. Zhu, W.X.: Penalty parameter for linearly constrained 0-1 quadratic programming. J. Optim. Theory Appl. **116**(1), 229–239 (2003)

15. Lucidi, S., Rinaldi, F.: Exact penalty functions for nonlinear integer programming problems. Technical Report Dipartimento Informatica e Sistemistica, Sapienza Università di Roma, vol. 1, n. 10 (2009)