Journal of Global Optimization **18:** 235–254, 2000.

Success Guarantee of Dual Search in Integer Programming: p-th Power Lagrangian Method ⋆

D. LI⋆⋆ and X. L. SUN
Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, NT, Hong Kong (email: dli@se.cuhk.edu.hk); Department of Mathematics, Shanghai University, Baoshan, Shanghai 200436, P. R. China

(Received 6 July 1999; accepted in revised form 15 April 2000)

Abstract. Although the Lagrangian method is a powerful dual search approach in integer programming, it often fails to identify an optimal solution of the primal problem. The p-th power Lagrangian method developed in this paper offers a success guarantee for the dual search in generating an optimal solution of the primal integer programming problem in an equivalent setting via two key transformations. One other prominent feature of the p-th power Lagrangian method is that the dual search only involves a one-dimensional search within [0,1]. Some potential applications of the method as well as the issue of its implementation are discussed.

Key words: Integer programming, Dual search, Lagrangian method

1. Introduction

The following general class of finite integer programming problems is considered in this paper:

$$(P) \qquad \min \ f(x) \tag{1.1a}$$

$$\text{s.t.} \quad g_i(x) \leqslant b_i, \quad i = 1, 2, \ldots, m, \tag{1.1b}$$

$$x \in X \subseteq \mathbb{R}^n, \tag{1.1c}$$

where X is a finite integer set. Problem (P) is termed the primal problem. Without loss of generality, f and g_i, $i = 1, 2, \ldots, m$, are assumed to be strictly positive for all $x \in X$. Constraints in (1.1b) are called Lagrangian constraints. Define F to be the feasible region of the decision vector x in (P),

$$F = \{x \mid g_i(x) \leqslant b_i, \ i = 1, 2, \ldots, m; \ x \in X\}. \tag{1.2}$$

Denote by $v(Q)$ the optimal value of an optimization problem (Q). Thus the optimal objective value of the primal problem is $v(P)$.

⋆ This research was partially supported by the Research Grants Council of Hong Kong, Grant CUHK 4130/97E, and the National Science Foundation of China under Grant 79970107.

⋆⋆ Corresponding author.

The dual search method plays a significant role in integer optimization. The Lagrangian methods are widely used in linear integer programming in finding an optimal solution, see, e.g., Geoffrion (1974), Fisher and Shapiro (1974), Bell and Shapiro (1977), Shapiro (1979), and Fisher (1981). In most situations, the Lagrangian methods provide a lower bound for $v(P)$. Incorporating the set of Lagrangian constraints into the objective function by introducing a nonnegative Lagrangian multiplier vector, $\lambda = (\lambda_1, \lambda_2, \ldots, \lambda_m) \in \mathbb{R}^m_+$, yields a Lagrangian relaxation:

$$(PR_\lambda) \qquad \min_{x \in X} L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i [g_i(x) - b_i]. \tag{1.3}$$

The Lagrangian dual is an optimization problem in λ,

$$(D) \qquad \max_{\lambda \in \mathbb{R}^m_+} [v(PR_\lambda)]. \tag{1.4}$$

The Lagrangian method searches for an optimal solution of (P) via maximizing the dual function $v(PR_\lambda)$.

If $\hat{x}$ solves both (P) and $(PR_{\hat{\lambda}})$ with $\hat{\lambda} \in \mathbb{R}^m_+$, then $\hat{\lambda}$ is said to be an optimal generating Lagrangian multiplier vector. If $\hat{x}$ solves both (P) and $(PR_{\hat{\lambda}})$ with $\hat{\lambda} \in \mathbb{R}^m_+$, and $\hat{\lambda}$ solves the dual problem (D), then $\{\hat{x}, \hat{\lambda}\}$ is said to be an optimal primal-dual pair of (P).

While the Lagrangian method is a powerful constructive dual search method, it often fails to identify an optimal solution of the primal integer optimization problem. Two critical situations could be present that prevent the Lagrangian method from succeeding in the dual search. Firstly, the optimal solution of (P) may not even be generated by solving (PR_λ) for any $\lambda \geqslant 0$. Secondly, the optimal solution to (PR_{λ^*}), with λ^* being a solution to the dual problem (D), is not necessarily an optimal solution to (P), or even not feasible. The first situation mentioned above is associated with the existence of an optimal generating Lagrangian multiplier vector. The second situation is related to the existence of an optimal primal-dual pair.

As an illustrative example, let us consider Example 5.12 in Parker and Rardin (1988):

$$\begin{aligned} \min \quad & 3x_1 + 2x_2 \\ \text{s.t.} \quad & g_1(x) = 10 - 5x_1 - 2x_2 \leqslant 7, \\ & g_2(x) = 15 - 2x_1 - 5x_2 \leqslant 12, \\ & x \in X = \left\{ \begin{array}{c} \text{integer} \\ 0 \leqslant x_1 \leqslant 1,\ 0 \leqslant x_2 \leqslant 2 \\ 8x_1 + 8x_2 \geqslant 1 \end{array} \right\}. \end{aligned} \tag{1.5}$$

Note that in order to conform with the problem assumption in (1.1) the two Lagrangian constraints in (1.5) take forms equivalent to the original Lagrangian constraints in Ex. 5.12 of Parker and Rardin (1988). The explicit expression of set X is

$X = \{(0, 1), (0, 2), (1, 0), (1, 1), (1, 2)\}$. It is easy to check that only (0, 2), (1, 1), and (1, 2) are feasible for the problem. The optimal solution is $x^* = (0, 2)$ with optimal value $v(P) = 4$. If the conventional Lagrangian method is applied, the solutions, (1, 0) and (0, 1), to problem (PR_{λ^*}) with $\lambda^* = (1/3, 0)$ being the maximizer of (D) are not optimal solutions of (P). They are both infeasible for (P).

The main purpose of this paper is to integrate two equivalent transformations that ensure the existence of an optimal primal-dual pair in an equivalent problem setting, thus offering a success guarantee for the dual search in generating an optimal solution of the primal integer programming problem. Based on the existence of an optimal primal-dual pair, we propose a convergent p-th power Lagrangian method. An optimal solution to a Lagrangian relaxation problem is obtained at each iteration of the method and a new multiplier is generated via dual search. One prominent feature is that the dual search only involves a one-dimensional search of a scalar Lagrangian multiplier within interval [0,1].

The organization of this paper is as follows. In Section 2, a t-norm surrogate constraint method is adopted to construct a single-constraint surrogate model that is exactly equivalent to the primal problem. The surrogate constraint method developed in this paper is based on a similar technique recently developed in Li (1999) for a more general problem setting. In Section 3, a p-th power transformation is investigated. Applying the p-th power transformation to the objective function guarantees the existence of an optimal primal-dual pair, thus ensuring the success of the dual search. In Section 4, the results in Sections 2 and 3 lead to the development of the p-th power Lagrangian method. In Section 5, two classes of nonlinear integer programming problems with real-world background are presented to show the potential applications of the proposed p-th power Lagrangian method. The paper concludes in Section 6 with suggestions for future research.

2. Equivalent t-norm surrogate constraint formulation

The use of the surrogate constraint formulation in integer programming was investigated in Glover (1968), Karwan and Rardin (1979) and Karwan and Rardin (1980). The surrogate constraint method converts a mathematical programming problem with multiple constraints into a one with a single aggregated constraint using a multiplier vector. The multiplier vector is successively adjusted such that a surrogate dual is maximized. The surrogate dual in general, however, does not guarantee the generation of an optimal solution of the primal problem. A surrogate strategy termed p-norm surrogate constraint method was recently developed in Li (1999) for general integer programming problems that yields an exact equivalence between the primal problem and the surrogated one without any assumption of convexity. We will give in this section a revised version of the p-norm surrogate formulation in Li (1999).

Problem (P) can be always converted to an equivalent form with $b_1 = b_2 = \cdots = b_m$. We thus further assume, without loss of generality, that $b_1 = b_2 =$

$\cdots = b_m = b > 0$ in (1.1). Problem (P) is then equivalent to the following single-constraint problem:

$$\begin{aligned} \min \quad & f(x) && \text{(2.6a)}\\ \text{s.t.} \quad & g_M(x) := \max\{g_1(x), \ldots, g_m(x)\} \leqslant b, && \text{(2.6b)}\\ & x \in X. && \text{(2.6c)} \end{aligned}$$

Let $g(x) = (g_1(x), \ldots, g_m(x))$. Note that the nonsmooth function $g_M(x)$ is exactly the infinite norm $\|g(x)\|_\infty$, which can be approximated by the t-th norm

$$\|g(x)\|_t = \sqrt[t]{[g_1(x)]^t + \ldots + [g_m(x)]^t}$$

as t tends to infinity. We further have

$$\frac{g_M(x)}{\sqrt[t]{m}} \leqslant \frac{\|g(x)\|_t}{\sqrt[t]{m}} \leqslant g_M(x). \tag{2.7}$$

A t-norm surrogate constraint formulation of (P) is formed by replacing $g_M(x)$ in (2.6) by $G_t(x) = \|g(x)\|_t/\sqrt[t]{m}$ for $t > 0$,

$$\begin{aligned} (S_t) \quad & \min \quad f(x) && \text{(2.8a)}\\ & \text{s.t.} \quad G_t(x) \leqslant b, && \text{(2.8b)}\\ & \qquad\ x \in X. && \text{(2.8c)} \end{aligned}$$

Let F_t denote the feasible region of (S_t),

$$F_t = \{x \mid G_t(x) \leqslant b;\ x \in X\}.$$

It is clear from (2.7) that $F \subseteq F_t$ for any $t > 0$. The surrogate problem (S_t) is thus a relaxation of (P) when $t \geqslant 1$. The following theorem shows that the sets F_t and F will be identical if t is chosen sufficiently large.

THEOREM 2.1. *Assume that $X \setminus F \neq \emptyset$. Let*

$$U = \min\{\frac{g_M(x)}{b} \mid x \in X \setminus F\}. \tag{2.9}$$

Then $F = F_t$ holds for all $t > t_0$, where $t_0 = \ln(m)/\ln(U)$.

Proof. Since $F \subseteq F_t$, we only need to prove $F_t \subseteq F$. We first note from (2.9) that $U > 1$ since $g_M(x) > b$ for any $x \in X \setminus F$. Hence, we have $t_0 \geqslant 0$. If $t > t_0$, then

$$\min\{\frac{g_M(x)}{b\sqrt[t]{m}} \mid x \in X \setminus F\} > 1. \tag{2.10}$$

For any $\hat{x} \in X \setminus F$, from (2.7) and (2.10), we have

$$\frac{G_t(\hat{x})}{b} \geqslant \frac{g_M(\hat{x})}{b\sqrt[t]{m}} > 1,$$

that is, $\hat{x} \notin F_t$ and hence $F_t \subseteq F$. □

The bound t_0 is smaller in general situations than the one obtained in Li (1999). Under a mild condition, an explicit bound of t in Theorem 2.1 can be specified.

COROLLARY 2.1. *Suppose that all g_i, $i = 1, 2, \ldots, m$, are integer-valued functions, e.g., polynomial functions with integer coefficients, and b is a positive integer. Then $F = F_t$ for all $t > t_1$, where*

$$t_1 = \frac{\ln(m)}{\ln[(1+b)/b]}. \tag{2.11}$$

Proof. Since $g_M(x) \geqslant b + 1$ for all $x \in X \setminus F$, we have $U \geqslant (1 + b)/b$. The conclusion then follows from Theorem 2.1. □

By selecting a sufficiently large t, all infeasible solutions of the primal problem will be excluded from F_t. In other words, the feasible set defined by the t-norm surrogate constraint, F_t, will exactly match the feasible set of the primal problem for a sufficiently large t. For illustration, let us consider the example problem, Ex. 5.12 in Parker and Rardin (1988), which we discussed in Section 1. To make the right-hand sides equal for the two constraints, we multiply $g_1(x)$ by 12/19 and $g_2(x)$ by 7/19. Applying then the t-norm surrogate constraint method yields the following formulation,

$$\begin{aligned}
&\min \ 3x_1 + 2x_2 \qquad\qquad (2.12)\\
&\text{s.t.} \quad \left[(12/19)^t(10 - 5x_1 - 2x_2)^t + (7/19)^t(15 - 2x_1 - 5x_2)^t\right]^{1/t} \leqslant 2^{1/t}\, 84/19,\\
&\qquad x \in X = \left\{ \begin{array}{c} \text{integer} \\ 0 \leqslant x_1 \leqslant 1, \quad 0 \leqslant x_2 \leqslant 2 \\ 8x_1 + 8x_2 \geqslant 1 \end{array} \right\}.
\end{aligned}$$

It can be verified that when $t \geqslant 9$, $F_t = F$ and the t-norm surrogate problem (2.12) is equivalent to the problem (1.5).

An appropriate single surrogate constraint can be always constructed in aggregating multiple Lagrangian constraints of the primal problem such that a surrogate formulation and the primal problem are exactly equivalent. This result offers a basis in developing the p-th power Lagrangian method in the next section.

3. p-th power transformation

We have shown in the last section that the problems (S_t) and (P) are equivalent when $t > t_0$. We will develop in this section a dual search scheme using a p-th power transformation for problem (S_t) with a fixed $t > t_0$. The convexification results derived from the p-th power transformation are based on the analysis on the perturbation function of (S_t). The perturbation function associated with (S_t) is defined by:

$$\phi(y) \;=\; \min\,\{f(x) \mid G_t(x) \leqslant y;\;\; x \in X\}.$$

It can be easily seen that the perturbation function ϕ is a nonincreasing piecewise-constant function of y. The value of ϕ remains at a constant level when no new integer solution with smaller f value becomes feasible. Hence, the perturbation function is continuous from the right. The domain of $\phi(\cdot)$ is

$$Y = \{y \mid \text{there exists } x \in X \text{ with } G_t(x) \leqslant y\}.$$

Based on the problem assumption, it is clear $Y = [\underline{y}, \infty)$ with $\underline{y} = \min_{x\in X} G_t(x)$. By the finiteness of X, there exists a finite $\overline{y} > \overline{0}$ such that $\phi(y)$ remains at a constant level, $\min_{x\in X}\; f(x)$, for any $y \in [\overline{y}, \infty)$. Therefore, the number of the discontinuous points of ϕ is finite. List them as $\{a_1, a_2, \cdots, a_N\}$ with

$$\underline{y} = a_0 < a_1 < a_2 < \cdots < a_N = \overline{y}. \tag{3.13}$$

If $a_N \leqslant b$, then (P) can be reduced to an equivalent unconstrained integer programming problem without considering the constraint. In the following discussion, we assume that $a_N > b$. Let $c_i = \phi(a_i)$, $i = 0, 1 \ldots, N$. By the definition of ϕ, we have

$$c_0 > c_1 > c_2 > \ldots > c_N > 0. \tag{3.14}$$

Now we impose a p-th power on the objective function of (S_t). Problem (S_t) can then be represented by the following equivalent form,

$$(P_p) \qquad \min\; [f(x)]^p \tag{3.15a}$$

$$\text{s.t.}\quad G_t(x) \leqslant b, \tag{3.15b}$$

$$x \in X, \tag{3.15c}$$

where $p > 0$. The Lagrangian relaxation of problem (P_p) is given as follows with a Lagrangian multiplier $\mu \geqslant 0$,

$$(P_pR_\mu) \qquad \min_{x\in X}\; L_p(x, \mu) := [f(x)]^p + \mu[G_t(x) - b]. \tag{3.16}$$

The Lagrangian dual of (P_p) is,

$$(D_p) \qquad \max_{\mu\in\mathbb{R}_+}\; v(P_pR_\mu). \tag{3.17}$$

Denote by $\phi_p(y)$ the perturbation function associated with (P_p). It is clear that $\phi_p(y) = [\phi(y)]^p$. The domain of $\phi_p(y)$ and the set of discontinuous points of $\phi_p(y)$ are still the same as their counterparts in $\phi(y)$. Let

$$\begin{aligned} \Phi_p &= \{(y, y_0) \mid y_0 = \phi_p(y);\;\; y \in Y\}, \\ E_p &= \{(a_i, c_i^p) \mid i = 0, 1, \ldots, N\}. \end{aligned} \tag{3.18}$$

A point in E_p will be called a *noninferior point* of Φ_p or ϕ_p. Obviously, $(y, y_0) \in E_p$ iff $(y, y_0) \in \Phi_p$ and $(z, y_0) \notin \Phi_p$ for any $z < y$.

LEMMA 3.1. *If* $a_k \leqslant b < a_{k+1}$ *for a* $k \in \{0, 1, \ldots, N-1\}$, *then*

$$v(P_p) = c_k^p \tag{3.19}$$

and $\hat{x} = \arg\min\{[f(x)]^p \mid G_t(x) \leqslant a_k\}$ *is an optimal solution of* (P_p) *and of* (P).

Proof. The lemma is obvious from the feasibility requirement and the fact that the perturbation function is nonincreasing. □

LEMMA 3.2. (i) *For any* $y \in Y$, *if* x^* *solves the perturbated problem*

$$\phi_p(y) = \min\{[f(x)]^p \mid G_t(x) \leqslant y;\ x \in X\},$$

then $(G_t(x^*), [f(x^*)]^p) \in \Phi_p$.

(ii) *If* x^* *solves* $(P_p R_\mu)$ *for some* $\mu > 0$, *then* $(G_t(x^*), [f(x^*)]^p) \in E_p$.

(iii) *For any* $(a_i, c_i^p) \in E_p$, *there exists* $x^* \in X$ *such that* $(a_i, c_i^p) = (G_t(x^*), [f(x^*)]^p)$.

(iv) *There exists at least one* $\hat{x} \in X$ *such that* $\hat{x}$ *solves* (P) *and* $(G_t(\hat{x}), [f(\hat{x})]^p) \in E_p$.

Proof. (i) Since $G_t(x^*) \leqslant y$, by the monotonicity of $\phi_p(y)$, we have $[f(x^*)]^p = \phi_p(y) \leqslant \phi_p(G_t(x^*))$. On the other hand, since x^* is feasible in the perturbated problem $\phi_p(y) = \min\{[f(x)]^p \mid G_t(x) \leqslant G_t(x^*);\ x \in X\}$, we have $\phi_p(G_t(x^*)) \leqslant [f(x^*)]^p$. Thus, $\phi_p(G_t(x^*)) = [f(x^*)]^p$, that is, $(G_t(x^*), [f(x^*)]^p) \in \Phi_p$.

(ii) Suppose that there exists an $\hat{x}$ such that $[f(\hat{x})]^p = \phi_p(G_t(x^*))$ with $G_t(\hat{x}) \leqslant G_t(x^*)$. By the definition of ϕ_p, we have, $[f(\hat{x})]^p \leqslant [f(x^*)]^p$. If $[f(\hat{x})]^p < [f(x^*)]^p$, then

$$[f(\hat{x})]^p + \mu(G_t(\hat{x}) - b) < [f(x^*)]^p + \mu(G_t(x^*) - b),$$

which contradicts to the optimality of x^* in $(P_p R_\mu)$. We therefore have $\phi_p(G_t(x^*)) = [f(\hat{x})]^p = [f(x^*)]^p$ and thus $(G_t(x^*), [f(x^*)]^p) \in \Phi_p$. If, on the contrary, $(G_t(x^*), [f(x^*)]^p) \notin E_p$ and $G_t(x^*) > a_0$. Then, there exists a $(y, y_0) \in \Phi_p$ such that $y < G_t(x^*)$ and $y_0 = \phi_p(y) = [f(x^*)]^p$. Suppose that $[f(\tilde{x})]^p = \phi_p(y)$ with $G_t(\tilde{x}) \leqslant y$. Then, we have

$$[f(\tilde{x})]^p + \mu(G_t(\tilde{x}) - b) < [f(x^*)]^p + \mu(G_t(x^*) - b),$$

which is a contradiction to that x^* solves $(P_p R_\mu)$.

(iii) Suppose that x^* solves the perturbated problem $\{[f(x)]^p \mid G_t(x) \leqslant a_i;\ x \in X\}$, then $[f(x^*)]^p = c_i^p$ and $G_t(x^*) \leqslant a_i$. From part (i), we know that $(G_t(x^*), [f(x^*)]^p) \in \Phi_p$. By the definition of E_p, we must have $G_t(x^*) = a_i$ and so $(G_t(x^*), [f(x^*)]^p) = (a_i, c_i^p)$.

(iv) Let S^* denote the set of optimal solutions of (P). For any $x^* \in S^*$, it follows from part (i) that $(G_t(x^*), [f(x^*)]^p) \in \Phi_p$. Let $\hat{x} = \arg\min\{G_t(x^*) \mid x^* \in S^*\}$.

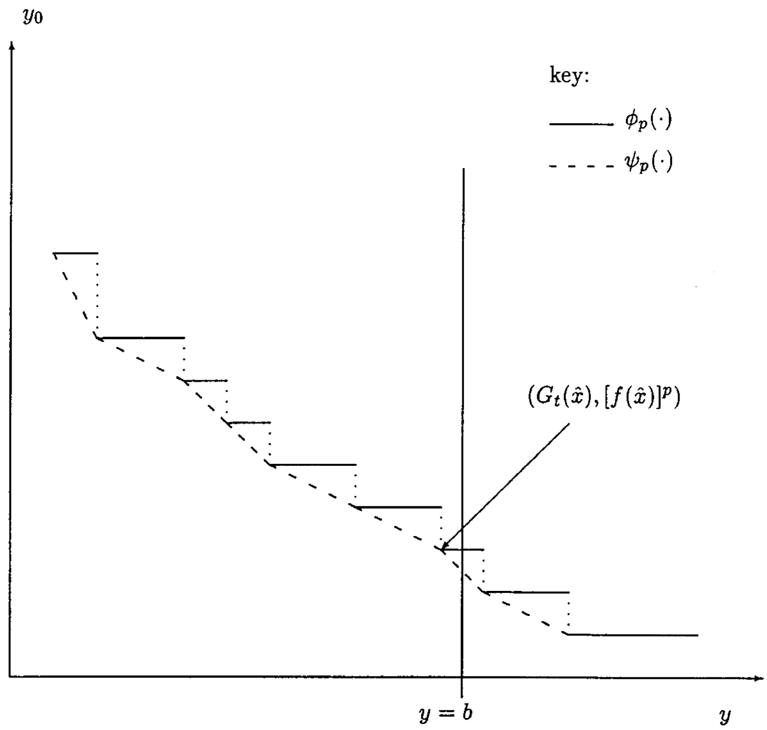

Figure 1. Perturbation function and its lower envelope.

Then, for any $z < G_t(\hat{x})$, we have $\phi_p(z) > [f(\hat{x})]^p$ and hence $(z, [f(\hat{x})]^p) \notin \Phi_p$. Therefore, $(G_t(\hat{x}), [f(\hat{x})]^p) \in E_p$. □

Now we define the lower envelope function of $\phi_p(y)$ as

$$\psi_p(y) = \begin{cases} c_0^p - \mu_0(p)(y - a_0), & a_0 \leqslant y \leqslant a_1 \\ c_1^p - \mu_1(p)(y - a_1), & a_1 \leqslant y \leqslant a_2 \\ \cdots & \cdots \\ c_{N-1}^p - \mu_{N-1}(p)(y - a_{N-1}), & a_{N-1} \leqslant y \leqslant a_N \\ c_N^p, & a_N \leqslant y < \infty \end{cases} \tag{3.20}$$

where

$$\mu_i(p) = -\frac{c_{i+1}^p - c_i^p}{a_{i+1} - a_i} > 0, \quad i = 0, 1, \ldots, N-1. \tag{3.21}$$

It is clear that $\phi_p(y) \geqslant \psi_p(y)$ for all $y \in Y$ and $\phi_p(a_i) = \psi_p(a_i) = c_i^p$ for $i = 0, 1, \ldots, N$. See Figure 1 for graphical illustration. The lower envelope function $\psi_p(y)$ is continuous and piecewise linear. We have the following convexification result for $\psi_p(y)$.

THEOREM 3.1. *Let*

$$p_0 = \frac{\ln[(\alpha+1)/\alpha]}{\ln(\beta)} \tag{3.22}$$

with

$$\alpha = \min_{0\leqslant i\leqslant N-2}\left(\frac{a_{i+2}-a_{i+1}}{a_{i+1}-a_i}\right), \tag{3.23}$$

$$\beta = \min_{0\leqslant i\leqslant N-2}\left(\frac{c_i}{c_{i+1}}\right). \tag{3.24}$$

Then $\psi_p(y)$ is a convex function of y when $p \geqslant p_0$.

Proof. We first observe from (3.13) and (3.14) that $\alpha > 0$ and $\beta > 1$. Thus, $p_0 > 0$. By the definition of $\psi_p(y)$ (cf. (3.20)), the convexity of $\psi_p(y)$ is equivalent to the decreasing monotonicity of the sequence $\{\mu_0(p), \mu_1(p), \ldots, \mu_{N-1}(p)\}$. From (3.21), the inequality $\mu_{i+1}(p) < \mu_i(p)$ is equivalent to

$$\frac{c_{i+1}^p - c_{i+2}^p}{a_{i+2}-a_{i+1}} < \frac{c_i^p - c_{i+1}^p}{a_{i+1}-a_i},$$

which is in turn equivalent to

$$\frac{1-(c_{i+2}/c_{i+1})^p}{(c_i/c_{i+1})^p - 1} < \frac{a_{i+2}-a_{i+1}}{a_{i+1}-a_i}. \tag{3.25}$$

Note that

$$\frac{1-(c_{i+2}/c_{i+1})^p}{(c_i/c_{i+1})^p - 1} < \frac{1}{(c_i/c_{i+1})^p - 1} \leqslant \frac{1}{\beta^p - 1} \tag{3.26}$$

and

$$\alpha \leqslant \frac{a_{i+2}-a_{i+1}}{a_{i+1}-a_i}. \tag{3.27}$$

If $p \geqslant p_0$, then, by (3.22), we have $1/(\beta^p - 1) \leqslant \alpha$. Thus, from (3.26) and (3.27), we imply that (3.25) holds for each $i = 0, \ldots, N-2$ when $p \geqslant p_0$. □

The implication of Theorem 3.1 is clear. When $p \geqslant p_0$, $\psi_p(y)$ becomes a convex function. Thus, a subgradient of $\psi_p(y)$ exists at every $y = a_i$, $i = 0, 1, \ldots, N$. Specially, by Lemma 3.2 (iv), a subgradient of ψ_p exists at $y = G_t(\hat{x})$, where $\hat{x}$ is an optimal solution of (P_p) and $G_t(\hat{x}) = \min\{G_t(x^*) \mid x^* \in S^*\}$, if $p \geqslant p_0$. In summary, the existence of an optimal generating Lagrangian multiplier is guaranteed when $p \geqslant p_0$. This convexification result will further lead to the existence of an optimal primal-dual pair.

THEOREM 3.2. *Let $\hat{x}$ be such that $\hat{x}$ solves (P), or equivalently $\hat{x}$ solves (P_p), and $G_t(\hat{x}) = \min\{G_t(x^*) \mid x^* \in S^*\}$, where S^* is the set of the optimal solutions of (P). Assume that $(G_t(\hat{x}), [f(\hat{x})]^p) = (a_k, c_k^p)$. Then $\{\hat{x}, \mu_k(p)\}$ is an optimal primal-dual pair of problem (P_p) when $p \geqslant p_0$.*

Proof. We first prove that $\hat{x}$ solves problem $(P_p R_{\mu_k(p)})$. From the feasibility and optimality of $\hat{x}$, we have $a_{k+1} > b$. By Theorem 3.1, $\psi_p(y)$ is a convex function of y when $p \geqslant p_0$ and $-\mu_k(p)$ is a subgradient of $\psi_p(y)$ at $y = G_t(\hat{x}) = a_k$. We have

$$[f(\hat{x})]^p + \mu_k(p)[G_t(\hat{x}) - y] \leqslant \psi_p(y) \leqslant \phi_p(y), \quad \forall\, y \in Y. \tag{3.28}$$

For any $x \in X$, let $y = G_t(x)$. Then $\phi_p(y) \leqslant [f(x)]^p$. It follows from (3.28) that

$$[f(x)]^p \geqslant \phi_p(y) \geqslant [f(\hat{x})]^p + \mu_k(p)[G_t(\hat{x}) - G_t(x)],$$

which in turn yields

$$[f(x)]^p + \mu_k(p)[G_t(x) - b] \geqslant [f(\hat{x})]^p + \mu_k(p)[G_t(\hat{x}) - b]. \tag{3.29}$$

Since $x \in X$ is arbitrary, (3.29) implies that $\hat{x}$ solves problem $(P_p R_{\mu_k(p)})$.

We now turn to prove that $\mu_k(p)$ solves (D_p). For any fixed $\mu \geqslant 0$, if x_μ solves $(P_p R_\mu)$, then by the definition of $\phi_p(y)$, we have $\phi_p(G_t(x_\mu)) \leqslant [f(x_\mu)]^p$. For any $y \in Y$, suppose that $[f(x)]^p = \phi_p(y)$ with $G_t(x) \leqslant y$, then

$$\begin{aligned}
\phi_p(y) &= [f(x)]^p \\
&\geqslant [f(x)]^p + \mu[G_t(x) - y] \\
&= [f(x)]^p + \mu[G_t(x) - b] + \mu(b - y) \\
&\geqslant [f(x_\mu)]^p + \mu[G_t(x_\mu) - b] + \mu(b - y) \\
&= [f(x_\mu)]^p + \mu[G_t(x_\mu) - y].
\end{aligned} \tag{3.30}$$

Setting $y = a_i$ $(i = k,\ k+1)$ in (3.30) and noting that $c_i^p = \psi_p(a_i) = \phi_p(a_i)$, we have

$$c_i^p \geqslant [f(x_\mu)]^p + \mu[G_t(x_\mu) - a_i], \quad i = k,\ k+1. \tag{3.31}$$

Moreover, since $b \in [a_k, a_{k+1})$, there exists a $\gamma \in (0, 1]$ such that $b = \gamma a_k + (1 - \gamma)a_{k+1}$. We thus obtain from (3.21) and (3.31) that

$$\begin{aligned} v(P_p R_{\mu_k(p)}) &= [f(\hat{x})]^p + \mu_k(p)[G_t(\hat{x}) - b] \\ &= c_k^p - \frac{c_{k+1}^p - c_k^p}{a_{k+1} - a_k}[a_k - (\gamma a_k + (1-\gamma)a_{k+1})] \\ &= c_k^p + (1-\gamma)(c_{k+1}^p - c_k^p) \\ &= \gamma c_k^p + (1-\gamma)c_{k+1}^p \\ &\geqslant \gamma\{[f(x_\mu)]^p + \mu[G_t(x_\mu) - a_k]\} + (1-\gamma)\{[f(x_\mu)]^p + \\ &\quad \mu[G_t(x_\mu) - a_{k+1}]\} \\ &= [f(x_\mu)]^p + \mu(G_t(x_\mu) - b) \\ &= v(P_p R_\mu). \end{aligned}$$

Hence $\mu_k(p)$ solves (D_p) when $p \geqslant p_0$. □

COROLLARY 3.1. *Suppose that f and g_i, $i = 1, 2, \ldots, m$, are integer-valued functions, e.g., polynomial functions with integer coefficients and b is a positive integer. Take a positive integer t greater that t_1 defined in (2.11). Under the same assumptions of Theorem 3.2, $\{\hat{x}, \mu_k(p)\}$ is an optimal primal-dual pair of (P_p) for all $p \geqslant p_1$, where*

$$p_1 = \frac{\ln(\overline{g})}{\ln[(1+\overline{f})/\overline{f}]} \tag{3.32}$$

with

$$\begin{aligned} \overline{g} &= \max\{(\|g(x)\|_t)^t \mid x \in X\}, \\ \overline{f} &= \max\{f(x) \mid x \in X\}. \end{aligned}$$

Proof. Note that the constraint $G_t(x) \leqslant b$ in (P_p) is equivalent to $(\|g(x)\|_t)^t \leqslant mb^t$. Replacing $G_t(x)$ and b by $(\|g(x)\|_t)^t$ and mb^t in (P_p), respectively, we get an integer-valued constraint function $G_t(x)$ in (P_p). It follows from (3.23) and (3.24) that

$$\begin{aligned} \alpha &\geqslant \frac{1}{\overline{g} - 1}, \\ \beta &\geqslant \min_{1 \leqslant i \leqslant N-2} \frac{c_{i+1} + 1}{c_{i+1}} \geqslant \frac{1 + \overline{f}}{\overline{f}}. \end{aligned}$$

Thus

$$p_0 \leqslant \frac{\ln(\overline{g})}{\ln[(1+\overline{f})/\overline{f}]} = p_1.$$

The corollary then follows from Theorem 3.2. □

The implication of Theorem 3.2 and Corollary 3.1 is significant. If the value of p is selected to be equal to or larger than p_0 or p_1, then an optimal solution of problem (P_p) is guaranteed to be generated by the dual search. In other words, an optimal solution of (P) can be generated by applying the conventional Lagrangian method to problem (P_p), i.e., the existence of an optimal primal-dual pair is ensured for (P_p) when $p \geqslant p_0$.

4. p-th power Lagrangian method

Recognizing prominent features of problem (P_p), the following special dual search method is devised to facilitate the solution process. Set

$$w = \frac{\mu}{1+\mu}.$$

Problem $(P_p R_\mu)$ can be recast to the following equivalent form,

$$(A_w) \qquad \min_{x \in X} l(x, w) = (1-w)[f(x)]^p + w(G_t(x) - b), \tag{4.33}$$

where $w \in [0,1]$.

On the basis of the previous discussion, a solution algorithm of the p-th power Lagrangian method is now proposed as follows. Geometrically, the algorithm performs on the noninferior points of the perturbation function ϕ_p. The algorithm starts to determine the first and the last noninferior points in E_p (cf. (3.18)). At each iteration, the Lagrangian relaxation (A_w) is solved with $w = \mu/(1+\mu)$, where $-\mu$ is the slope of the line connecting the two noninferior points of Φ_p that are corresponding to the best feasible solution and the least infeasible solution up to the current iteration, respectively. A new noninferior point will be generated if the optimal solution has not been reached. Eventually, the algorithm will terminate at two noninferior points of Φ_p that are nearest to the line $y = b$ on the left and right, respectively.

p-th Power Lagrangian Method (pPLM)

Step 1. Set $w = 1$. Solve (A_1). Denote the optimal solution by x^0. If $G_t(x^0) - b > 0$, stop. There is no feasible solution. Otherwise set $f_0^- = [f(x^0)]^p$ and $d_0^- = G_t(x^0)$.

Step 2. Set $w = 0$. Solve (A_0). Denote the optimal solution by z^0. If $G_t(z^0) - b \leqslant 0$, stop, z^0 is the optimal solution. Otherwise set $f_0^+ = [f(z^0)]^p$ and $d_0^+ = G_t(z^0)$.

Step 3. Set $k = 0$.

Step 4. Compute a w_k satisfying

$$l(x^k, w_k) = l(z^k, w_k). \tag{4.34}$$

Step 5. Solve (A_{w_k}). Denote the optimal solution by y^k. If x^k solves (A_{w_k}), stop, x^k is an optimal solution to (P). Otherwise, go to Step 6.

Step 6. If $G_t(y^k) - b \leqslant 0$, set

$$\begin{aligned}&f_{k+1}^- = [f(y^k)]^p,\ f_{k+1}^+ = f_k^+,\\ &d_{k+1}^- = G_t(y^k),\ d_{k+1}^+ = d_k^+,\\ &x^{k+1} = y^k,\ z^{k+1} = z^k.\end{aligned}$$

Otherwise if $G_t(y^k) - b > 0$, set

$$\begin{aligned}&f_{k+1}^- = f_k^-,\ f_{k+1}^+ = [f(y^k)]^p,\\ &d_{k+1}^- = d_k^-,\ d_{k+1}^+ = G_t(y^k),\\ &x^{k+1} = x^k,\ z^{k+1} = y^k.\end{aligned}$$

Set $k := k + 1$. Return to Step 4.

THEOREM 4.1. *If $p \geqslant p_0$, where p_0 is defined by (3.22), then the algorithm (pPLM) stops at an optimal solution of (P) within a finite number of steps.*

Proof. Suppose that the algorithm goes through Step 1 and Step 2. We first observe from Step 1 and Step 6 that

$$l(x^k, w_k) = (1 - w_k)f_k^- + w_k(d_k^- - b), \quad \forall k \geqslant 0, \tag{4.35}$$
$$l(z^k, w_k) = (1 - w_k)f_k^+ + w_k(d_k^+ - b), \quad \forall k \geqslant 0. \tag{4.36}$$

Thus, by (4.34), we have

$$w_k = \frac{f_k^- - f_k^+}{(f_k^- - f_k^+) + (d_k^+ - d_k^-)}, \quad k \geqslant 0. \tag{4.37}$$

From the algorithm, $f_k^- > f_k^+$ and $d_k^+ > d_k^-$ for all $k \geqslant 0$. Equation (4.37) then implies that $w_k \in (0, 1)$ for all $k \geqslant 0$.

We now show that if the algorithm stops at k-th iteration, i.e., x^k solves (A_{w_k}), then x^k is an optimal solution to (P). From Step 1 and Step 6, we know that x^k is a feasible solution and z^k is an infeasible solution to (P). Thus, we have $d_k^- \leqslant b < d_k^+$. By Lemma 3.2 (ii), points (d_k^-, f_k^-) and (d_k^+, f_k^+) belong to set E_p. We claim that there is no $(a_i, c_i^p) \in E_p$ such that a_i lies between d_k^- and d_k^+ and hence we can conclude by Lemma 3.1 that x^k is an optimal solution of (P). Suppose on the contrary, there exists $(a_i, c_i^p) \in E_p$ such that $d_k^- < a_i < d_k^+$, then $a_i = \lambda d_k^- + (1 - \lambda)d_k^+$ for some $\lambda \in (0, 1)$. Since $\psi_p(y)$ is a convex function of y and is strictly decreasing in $[a_0, a_N]$, we have

$$\begin{aligned} c_i^p &= \psi_p(a_i)\\ &< \lambda\psi_p(d_k^-) + (1 - \lambda)\psi_p(d_k^+)\\ &= \lambda f_k^- + (1 - \lambda)f_k^+. \end{aligned} \tag{4.38}$$

By Lemma 3.2 (iii), there exists $\tilde{x} \in X$ such that $(a_i, c_i^p) = (G_t(\tilde{x}), [f(\tilde{x})]^p)$. We thus obtain from (4.35), (4.36) and (4.38) that

$$\begin{aligned} v(A_{w_k}) &\leqslant (1-w_k)[f(\tilde{x})]^p + w_k(G_t(\tilde{x}) - b) \\ &= (1-w_k)c_i^p + w_k(a_i - b) \\ &< (1-w_k)[\lambda f_k^- + (1-\lambda) f_k^+] + w_k[\lambda d_k^- + (1-\lambda)d_k^+ - b] \\ &= \lambda\, l(x^k, w_k) + (1-\lambda)\, l(z^k, w_k) \\ &= l(x^k, w_k), \end{aligned}$$

where the last equality follows from (4.34). This contradicts to the assumption that x^k solves (A_{w_k}).

Next we prove the finite termination of the algorithm. We notice from (4.34) that if the algorithm does not stop at k-th iteration, then neither x^k nor z^k solves (A_{w_k}). Let

$$u_k = l(x^k, w_k) = l(z^k, w_k). \tag{4.39}$$

Then

$$l(y^k, w_k) < u_k. \tag{4.40}$$

By Lemma 3.2 (iii), there exists an $i \in \{0, 1, \ldots, N\}$ such that $(a_i, c_i^p) = (G_t(y^k), [f(y^k)]^p)$. We will show by contradiction that $d_k^- \leqslant G_t(y^k) \leqslant d_k^+$. Suppose that $G_t(y^k) > d_k^+$, then $d_k^+ = \lambda a_i + (1-\lambda)d_k^-$ for some $\lambda \in (0, 1)$. By (4.35), (4.36), (4.39) and the convexity of $\psi_p(y)$, we obtain

$$\begin{aligned} u_k &= (1-w_k) f_k^+ + w_k(d_k^+ - b) \\ &= (1-w_k)\psi_p(d_k^+) + w_k(d_k^+ - b) \\ &\leqslant (1-w_k)[\lambda\psi_p(a_i) + (1-\lambda)\psi_p(d_k^-)] + w_k(d_k^+ - b) \\ &= \lambda\{(1-w_k)[f(y^k)]^p + w_k(G_t(y^k) - b)\} + (1-\lambda)u_k \\ &= \lambda l(y^k, w_k) + (1-\lambda)u_k \end{aligned}$$

This contradicts to (4.40). Therefore, $G_t(y^k) \leqslant d_k^+$. Similarly, we can prove $G_t(y^k) \geqslant d_k^-$. Since neither x^k nor z^k solves (A_{w_k}), we must have $d_k^- < G_t(y^k) < d_k^+$. Thus, by the updating rules in Step 6, the range of $[d_k^-, d_k^+]$ is strictly decreasing as k increases. Since the set E_p is finite, an optimal solution must be reached at Step 5 within a finite number of steps. □

Now we demonstrate the solution algorithm by applying (pPLM) to the example problem, Ex. 5.12 in Parker and Rardin (1988). The Lagrangian relaxation of (2.12) is

$$\min\ L_p(x, \mu) = [f(x)]^p + \mu\{[(12/19)g_1(x)]^t + [(7/19)g_2(x)]^t - 2 \times (84/19)^t\}$$

$$\text{s.t.}\quad x \in X = \left\{ \begin{array}{c} x \text{ integer} \\ 0 \leqslant x_1 \leqslant 1,\ 0 \leqslant x_2 \leqslant 2 \\ 8x_1 + 8x_2 \geqslant 1 \end{array} \right\}.$$

Table 1. Solution process of the example problem ($t = p = 9$)

k	1	2	3	4
w_k	1	0	0.1103	0.16856
y^k	(0,2)	(0,1)	(1,0)	(0,2) & (1,0)
$[f(y^k)]^p$	262,144	512	19,683	262,144 & 19,683
$G_t(y^k) - b$	$-1.1291E + 6$	$9.8069E + 5$	$6.6862E + 4$	$-1.1291E + 6$ & $6.6862E + 4$
$v(A_{w_k})$	$-1.1291E + 6$	512	$2.4888E + 4$	$2.7635E + 4$
x^k	(0,2)	(0,2)	(0,2)	
f_k^-	262,144	262,144	262,144	
d_k^-	$1.6139E + 5$	$1.6139E + 5$	$1.6139E + 5$	
z^k		(0,1)	(1,0)	
f_k^+		512	19,683	
d_k^+		$2.2712E + 6$	$1.3574E + 6$	

Table 1 shows the solution process using (pPLM) with $t = p = 9$. At iteration 4, both $x^4 = (0, 2)$ and $z^4 = (1, 0)$ solve problem (A_{w_4}) with $w_4 = 0.16856$. The optimal solution $x^* = (0, 2)$ has thus been successfully identified through the dual search and is equal to x^4 at iteration 4.

In addition to the existence guarantee of a primal-dual pair and the success guarantee of the dual search associated with the p-th power Lagrangian method, the reduction in the dimension of the Lagrangian multiplier greatly facilitates the solution process. The Lagrangian multiplier is a scalar in the the p-th power Lagrangian method, while it is of an m-dimension in the conventional Lagrangian method.

The emphasis of this paper is to provide a theoretical foundation in characterizing the existence of optimal generating Lagrangian multiplier vectors and the existence of optimal primal-dual pairs. The computational aspects of the proposed p-th power Lagrangian method need to be further explored. Compared to the conventional Lagrangian method, a major disadvantage of the p-th power method is the nonlinearity inherent in the p-th power transformation (3.15) as well as in the t-th norm surrogate transformation (2.8). When the original problem is of a linear form, the p-th power method makes it nonlinear. When the original problem is of a separable form, the p-th power transformation makes it nonseparable. Promising application areas of the p-th power method thus seem to be in nonlinear nonseparable integer programming problems. For example, notice that any power of a zero-one variable is itself. Polynomial zero-one programming problem thus is an area where the p-th power Lagrangian method could show its computational promise in problem solving practice. Two classes of nonlinear integer programming problems are investigated in the next section to show the the potential applications of the proposed p-th power Lagrangian method.

5. Applications

Problem (A_w) at the lower level in the dual search can be viewed as a type of 'unconstrained' integer programming problem. The computational aspect in solving (A_w) is largely dependent on the problem structure. In this section we illustrate potential applications of the proposed p-th power Lagrangian method to two classes of nonlinear integer programming problems.

PROBLEM 1. Consider a network system consisting of n subsystems. Let x_i denote the number of the same redundancy components in parallel in i-th subsystem. The reliability of the i-th subsystem is $R_i(x_i) = 1 - (1 - r_i)^{x_i}$, where $r_i \in (0, 1)$ is the reliability of a fixed component in i-th subsystem. Also, denote by $C_i(x_i)$ the resource consumed in the i-th subsystem and by b the total available resource. The constrained redundancy optimization problem in a complex network (Tzafestas (1980)) can be formulated as

$$\min \quad Q(x) = 1 - g(R_1, R_2, \dots, R_n) \tag{5.41a}$$

$$\text{s. t.} \quad C(x) = h(C_1(x_1), C_2(x_n), \dots, C_n(x_n)) \leqslant b, \tag{5.41b}$$

$$x \in X = \{x \mid L_i \leqslant x_i \leqslant U_i,\ x_i \text{ integer},\ i = 1, \dots, n\}, \tag{5.41c}$$

where $Q(x)$ and $C(x)$ represent the overall unreliability of the system and the total resource consumed, respectively, and f and g are in general nonconvex functions on $\mathbb{R}^n$. Inherent properties in the complex reliability system are that $0 < Q(x) < 1$ and $C(x) > 0$, for all $x \in X$. Solution methods in the literature for (5.41) are mainly heuristic, see Tzafestas (1980), Tillman et al. (1980) and Ohtagaki et al. (1995). When branch and bound approach is used to solve (5.41), one has to obtain at each node of the search tree a global optimal solution to a nonlinear constrained nonconvex optimization problem, for which few efficient methods are known (Horst and Tuy (1993)). Now we apply the p-th power Lagrangian method to (5.41) by incorporating the nonlinear constraint $C(x) \leqslant b$ into the objective function. In consideration of computational stability, we take exponential to the objective function $Q(x)$. Since (5.41) is a singly inequality constrained problem, its p-th power Lagrangian relaxation (A_w) (cf. (4.33)) is

$$\min_{x \in X} \ (1 - w)\exp(pQ(x)) + w(C(x) - b). \tag{5.42}$$

This problem is much more tractable than (5.41) as the branch and bound method for (5.42) now involves solving an unconstrained global optimization problem over a box set at each node of the search tree. A number of computational implementable algorithms have been developed to globally minimize a nonconvex function over a box set, based on both deterministic approaches (see, e.g., Barhen et al. (1997), Ge (1990) and Horst and Tuy (1993)) and stochastic approaches (see, e.g., Cvijović and Klinowski (1995) and Rinnoy Kan and Timmer (1987a, b)).

Now consider an instance of problem (5.41) for a bridge network with 5 elements (Tzafestas (1980)):

$$\begin{aligned}\min\ & Q(x) = 1 - R_1R_2 - Q_2R_3R_4 - Q_1R_2R_3R_4 \\ & \qquad\qquad - R_1Q_2Q_3R_4R_5 - Q_1R_2R_3Q_4R_5 \\ \text{s. t. } & C(x) = x_1x_2 + 3x_2x_3 + 3x_2x_4 + x_1x_5 \leqslant 28, \\ & 1 \leqslant x_i \leqslant 6,\ x_i \text{ integer},\ i = 1, \cdots, 5,\end{aligned}$$

where $Q_i = 1 - R_i := 1 - R_i(x_i)$, $r_1 = 0.7$, $r_2 = 0.85$, $r_3 = 0.75$, $r_4 = 0.8$, $r_5 = 0.9$. The optimal solution of this example is $x^* = (2, 1, 4, 4, 1)$ with $Q(x^*) = 0.006569$. Take $p = 10$ in algorithm (pPLM). At each iteration of the algorithm, the box-constrained integer programming (A_w) (cf. (5.42)) is solved by a branch and bound procedure. The algorithm stops at iteration 7 with the optimal solution $x^7 = (2, 1, 4, 4, 1)$.

PROBLEM 2. Consider the following integer convex programming:

$$\min\ f(x) \tag{5.43a}$$
$$\text{s.t. } g_i(x) \leqslant 0,\ i = 1, \dots, m, \tag{5.43b}$$
$$x \in X = \{x \mid Ax \leqslant b,\ Bx = c,\ x \text{ is integral}\}, \tag{5.43c}$$

where f and g_i $(i = 1, \dots, m)$ are nonlinear convex functions, $A \in \mathbb{R}^{l_1 \times n}$, $B \in \mathbb{R}^{l_2 \times n}$, $b \in \mathbb{R}^{l_1}$, $c \in \mathbb{R}^{l_2}$, and X is a finite set.

Optimization problems with the structure given in (5.43) arise in many areas of practical interests (see, e.g., Cooper (1981), Kraay et al. (1991) and Sung and Cho (1999)). Computational difficulty, however, may be caused by the nonlinear constraints (5.43b) when outer approximation algorithm (Fletcher and Leyffer (1994)) or branch and bound method (Gupta and Ravindran (1985)) is adopted to solve (5.43). The p-th power Lagrangian method presented in this paper provides an approach to reduce (5.43) to a sequence of linearly constrained convex integer programming problems. In fact, by Theorem 2.1, (5.43) has the following equivalent form for a suitable $t > 0$ (cf. (3.15)):

$$\min\ \exp(pf(x)) \tag{5.44a}$$
$$\text{s.t. } \sum_{i=1}^{m} \exp(tg_i(x)) \leqslant m, \tag{5.44b}$$
$$x \in X, \tag{5.44c}$$

where we have taken exponential transformations to $f(x)$ and $g_i(x) \leqslant 0$ $(i = 1, \dots, m)$. Thus, the p-power Lagrangian relaxation (A_w) of (5.44) is

$$\min\ (1 - w)\exp(pf(x)) + w\Big(\sum_{i=1}^{m} \exp(tg_i(x)) - m\Big) \tag{5.45a}$$
$$\text{s.t. } x \in X. \tag{5.45b}$$

It is clear that for any $p > 0$ and $t > 0$, (5.45) is a linearly constrained convex integer programming, for which various algorithms have been developed by exploiting the polyhedral nature of the constraint set X (see Cooper (1981), Gupta and Ravindran (1985), Michelon and Maculan (1991) and Skorin-Kapov and Granot (1987)).

Consider the following example of problem (5.43):

$$
\begin{aligned}
\min\ & f(x) = 1/200\left(\sum_{i=1}^{4}(x_i^2-2)^2+(x_5^2-1)^2\right)\\
\text{s. t.}\ \ & g_1(x) = x_1^2+(x_2-5)^4+x_3^2+(x_4+5)^4-\\
& \qquad -3x_1x_2-x_3x_4+x_5-700\leqslant 0,\\
& x\in X=\left\{\begin{array}{c} x \text{ integer}\\ -x_1+2x_2-2x_3+5x_4-2\leqslant 0,\\ 2x_1-x_2-2x_3+x_5-2\leqslant 0,\\ -4\leqslant x_i\leqslant 4,\ i=1,\ldots,5.\end{array}\right\}
\end{aligned}
$$

The optimal solution of this example is $x^* = (2, 2, 1, 0, 1)$ with $f(x^*) = 0.065$. We take $p = 8$ in algorithm (pPLM). The linearly constrained convex integer programming (A_w) at each iteration is solved by a branch and bound procedure. After 8 iterations, the algorithm stops at the optimal solution $x^8 = (2, 2, 1, 0, 1)$.

6. Conclusions

Theoretical breakthroughs have been made in this paper to guarantee the success of the dual search by ensuring the existence of a primal-dual pair in integer programming. Two key equivalent transformations are involved in the solution process, the t-norm surrogate constraint formulation that converts an integer problem with multiple Lagrangian constraints into an equivalent one with a single surrogate Lagrangian constraint, and the p-th power transformation that takes p-th power on the objective function. When the values of t and p are selected large enough, these two equivalent transformations ensure the existence of an optimal primal-dual pair, in a new equivalent setting, for problems where an optimal primal-dual pair may not exist in the original setting.

The results presented in this paper can be also viewed as a companion of Li (1995) in which a convexification scheme using p-th power transformation is developed for general nonconvex nonlinear programming problems. It reveals, in Li (1995), that a saddle point can be generated for a class of nonconvex optimization problems in an equivalent representation space and the primal-dual method is then guaranteed to succeed with a zero duality gap. In a similar manner, the p-th power transformation is applied in this paper to the perturbation function of integer programming problems. Less can be achieved, however, in the integer programming case than in the nonlinear programming case, as we observe in the paper. We can only convexify the lower envelope function of the perturbation function, ψ_p, not the perturbation function itself. While the p-th power Lagrangian method can guarantee to generate an optimal solution of problem (P_p) and problem (P) via dual search, the duality gap is in general nonzero. Specifically, the duality gap is given by the following when $\{\hat{x}, \mu_k(p)\}$ is an optimal primal-dual pair of problem (P_p),

$$v(P_p) - v(P_p R_{\mu_k(p)}) = \mu_k(p)[b - G_t(\hat{x})].$$

The focus of this paper is to provide fresh theoretical insights into the dual search in integer programming. The resulting p-th power Lagrangian method yields an optimal solution of the primal problem in a convergent dual-search solution process.

7. Acknowledgements

The authors very much appreciate the discussion from Prof. Douglas J. White on dual search in integer programming.

References

Barhen, J., Protopopescu, V. and Reister, D. (1997), TRUST: A deterministic algorithm for global optimization, *Science* 276: 1094–1097.

Bell, D.E. and Shapiro, J.F. (1977), A convergent duality theory for integer programming, *Operations Research* 25: 419–434.

Cooper, M.W. (1981), A survey of methods for pure nonlinear integer programming, *Management Science* 27: 353–361.

Cvijović, D. and Klinowski, J. (1995), Taboo search: An approach to the multiple minima problem, *Science* 267: 664–666.

Fisher, M.L. and Shapiro, J.F. (1974), Constructive duality in integer programming, *SIAM Journal on Applied Mathematics* 27: 31–52.

Fisher, M.L. (1981), The Lagrangian relaxation method for solving integer programming problems, *Management Science* 27: 1–18.

Fletcher R. and Leyffer, S. (1994), Solving mixed integer nonlinear programs by outer approximation, *Mathematical Programming* 66: 327–349.

Ge, R. (1990), A filled function method for finding a global minimizer of a function of several variables, *Mathematical Programming* 46: 191–204.

Geoffrion, A.M. (1974), Lagrangean relaxation for integer programming, *Mathematical Programming Study* 2: 82–114.

Glover, F. (1968), Surrogate constraints, *Operations Research* 16: 741–749.

Gupta, O.K. and Ravindran, A. (1985), Branch and bound experiments in convex nonlinear integer programming, *Management Science* 31: 1533–1546.

Horst, R. and Tuy, H. (1993), *Global Optimization: Deterministic Approaches*, Springer-Verlag, Berlin.

Karwan, M.H. and Rardin, R.L. (1979), Some relationships between Lagrangian and surrogate duality in integer programming, *Mathematical Programming* 17: 320–334.

Karwan, M.H. and Rardin, R.L. (1980), Searchability of the composite and multiple surrogate dual functions, *Operations Research* 28: 1251–1257.

Kraay, D., Harker, P.T. and Chen, B.T. (1991), Optimal pacing of trains in freight railroads-model formulation and solution, *Operations Research* 39: 82–99.

Li, D. (1995), Zero duality gap for a class of nonconvex optimization problems, *J. of Optimization Theory and Applications* 85: 309–324.

Li, D. (1999), Zero duality gap in integer programming: P-norm surrogate constraint method, *Operations Research Letters* 25: 89–96.

Michelon, P. and Maculan, N. (1991), Lagrangean decomposition for integer nonlinear programming with linear constraints, *Mathematical Programming* 52: 303–313.

Ohtagaki, H., Nakagawa, Y., Iwasaki, A. and Narihisa, H. (1995), Smart greedy procedure for solving a nonlinear knapsack class of reliability optimization problems, *Mathematical and Computer Modelling* 22: 261–272.

Parker, R.G. and Rardin, R.L. (1988), *Discrete Optimization*, Academic, Boston.

Rinnoy Kan, A.H.G. and Timmer, G.T. (1987a), Stochastic global optimization methods, Part I: Clustering method, *Mathematical Programming* 39: 27–56.

Rinnoy Kan, A.H.G. and Timmer, G. T. (1987b), Stochastic global optimization methods, Part II: Multi-level method, *Mathematical Programming* 39: 57–78.

Shapiro, J.F. (1979), A survey of Lagrangian techniques for discrete optimization, *Annals of Discrete Mathematics* 5: 113–138.

Skorin-Kapov, J. and Granot, F. (1987), Non-linear integer programming: sensitivity analysis for branch and bound, *Operations Research Letters* 6: 269–274.

Sung, C.S. and Cho, Y.K. (1999), Branch-and-bound redundancy optimization for a series system with multiple-choice constraints, *IEEE Transactions on Reliability* 48: 108–117.

Tillman, F.A., Hwuang, C.L. and Kuo, W. (1980), *Optimization of System Reliability*, Marcel Dekker.

Tzafestas, S.G. (1980), Optimization of system reliability: A survey of problems and techniques, *International J. of Systems Science* 11: 455–486.