Laboratoire d'Analyse et Modélisation de Systèmes pour l'Aide à la Décision
CNRS UMR 7024

# CAHIER DU LAMSADE

# 264

Juillet 2007

## Rewriting integer variables into zero-one variables : some guidelines for the integer quadratic multi-knapsack problem

Dominique Quadri, Eric Soutif

# Rewriting integer variables into zero-one variables: some guidelines for the integer quadratic multi-knapsack problem

**Dominique Quadri**

(LAMSADE, Université Paris Dauphine, Place du Maréchal de Lattre de Tassigny, 75775 Paris Cedex 16, France)

**Eric Soutif**

(CEDRIC, Conservatoire National des Arts et Métiers, 292 rue Saint Martin, 75141 Paris Cedex 03, France)

**Abstract:** This paper is concerned with the integer quadratic multidimensional knapsack problem (*QMKP*) where the objective function is separable. Our objective is to determine which expansion technique of the integer variables is the most appropriate to solve (*QMKP*) to optimality using the upper bound method proposed by Quadri et al. (2007). To the best of our knowledge the upper bound method previously mentioned is the most effective method in the literature concerning (*QMKP*). This bound is computed by transforming the initial quadratic problem into a 0-1 equivalent piecewise linear formulation and then by establishing the surrogate problem associated. The linearization method consists in using a direct expansion initially suggested by Glover (1975) of the integer variables and in applying a piecewise interpolation to the separable objective function. As the direct expansion results in an increase of the size of the problem, other expansions techniques may be utilized to reduce the number of 0-1 variables so as to make easier the solution to the linearized problem. We will compare theoretically the use in the upper bound process of the direct expansion (I) employed in Quadri et al. (2007) with two other basic expansions, namely: (II) a direct expansion with additional constraints and (III) a binary expansion. We show that expansion (II) provides a bound which value is equal to the one computed by Quadri et al (2007). Conversely, we provide the proof of the non applicability of expansion (III) in the upper bound method. More specifically, we will show that if (III) is used to rewrite the integer variables into 0-1 variables then a

linear interpolation can not be applied to transform (*QMKP*) into an equivalent 0-1 piecewise linear problem.

**Keywords:** integer quadratic knapsack problem; separable objective function; direct expansion; binary expansion; piecewise interpolation.

**Primary Code:**

Secondary Code:

## Introduction

This paper deals with the integer quadratic multi-knapsack problem (*QMKP*) where the objective function is separable. Problems of this structure arise in numerous industrial and economic situations, for instance in production planning [12], reliability allocation [10] and finance [5]. These include the main application of (*QMKP*) which is in the portfolio management area where the investments are independent, see [4] and [5]. Nevertheless, solving (*QMKP*) efficiently will constitute a starting point to solve the more general and realistic portfolio management problem where the investments are dependent, i.e. the objective function is non separable.

The integer quadratic multi-knapsack problem (*QMKP*) where the objective function is separable consists in maximizing a concave separable quadratic integer function subject to *m* linear capacity constraints. It may be stated mathematically as follows:

$$(QMKP)\begin{cases} \max f(x)=\sum_{j=1}^{n} c_j x_j - d_j x_j^2 = \sum_{j=1}^{n} f_j\left(x_j\right) \\ \text{s. t.} \begin{vmatrix} \sum_{j=1}^{n} a_{ij} x_j \le b_i, & i=1,\ldots,m \\ 0 \le x_j \le u_j, & j=1,\ldots,n \\ x_j \text{ integer}, & j=1,\ldots,n \end{vmatrix} \end{cases}$$

where the coefficients $c_j$, $d_j$, $a_{ij}$, $b_i$ are nonnegative. The bounds $u_j$ of variables $x_j$ are pure integers, with $u_j \leq \left(c_j / 2d_j\right)$. Indeed, the separable objective function is concave which implies that for all function $f_j$, $x_j^* \leq \left(c_j / 2d_j\right)$, where $x^*_j$ is the optimal solution of the program $\max_{x_j \geq 0} f_j\left(x_j\right)$.

The problem (*QMKP*) which is a NP-hard problem [3] is a generalization of both the integer quadratic knapsack problem [2] and the 0-1 quadratic knapsack problem where the objective function is subject to only one constraint [1].

Since, (*QMKP*) is NP-hard, one should not expect to find a polynomial time algorithm for solving it exactly. Hence, we are usually interested in developing branch-and-bound algorithms. A key step in designing an effective exact solution method for such a maximization problem is to establish a tight upper bound on the optimal value.
Basically, the available upper bound procedures for (*QMKP*) may be classified as attempting either to solve efficiently the LP-relaxation of (*QMKP*) (see [2] and [8]) or to find a good upper bound, of better quality than the LP-relaxation of (*QMKP*), transforming (*QMKP*) into a *0-1* linear problem easier to solve (see [4] and [9]). To the best of our knowledge, the upper bound method we have proposed in a previous work [11] is better than the existing methods (Djerdjour, Mathur and Salkin algorithm [4], a *0-1* linearization method, a classical LP-relaxation of (*QMKP*)) from both a qualitative and a computational standpoint. We have first used a direct expansion of the integer variables, originally suggested by Glover [7], and apply a piecewise interpolation to the objective function: an equivalent *0-1* linear problem is thus obtained. The second step of the algorithm consists in establishing and solving the surrogate relaxation problem associated to the equivalent linearized formulation.

Nevertheless the transformed linear formulation encounters numerous *0-1* variables because of the direct expansion used (denoted by expansion (I) in the following). Consequently, other expansions techniques may be utilized to reduce the number of *0-1* variables so as to make easier the solution to the linearized problem. Let us

consider the three basic expansions for rewriting the integer variables of (*QMKP*) into *0-1* variables:

- Expansion (I): direct expansion

$$\begin{cases} \sum_{k=1}^{u_j} y_{jk} = x_j , \\ x_j \in \{1 \ldots u_j\}, \; y_{jk} \in \{0,1\} \\ \forall j = 1 \ldots n , \; k = 1 \ldots u_j \end{cases}$$

- Expansion (II): direct expansion with additional constraint

$$\begin{cases} \sum_{k=1}^{u_j} k \, y'_{jk} = x_j , \\ \sum_{k=1}^{u_j} y'_{jk} \le 1 \\ x_j \in \{1 \ldots u_j\} \; , \; y'_{jk} \in \{0,1\} \\ \forall j = 1 \ldots n , \; k = 1 \ldots u_j \end{cases}$$

- Expansion (III): binary expansion.

$$\begin{cases} \sum_{k=1}^{\lceil \log(u_j) \rceil} 2^k z_{jk} = x_j , \\ \sum_{k=1}^{\lceil \log(u_j) \rceil} 2^k z_{jk} \le u_j \\ x_j \in \{1 \ldots u_j\}, \; z_{jk} \in \{0,1\} \; \forall j = 1 \ldots n , \quad k = 1 \ldots \lceil \log(u_j) \rceil \end{cases}$$

The purpose of this note is to evaluate the impact of the use of the above expansion techniques, on the upper bound computation developed in [11]. More specifically, we will determinate which expansion is the most appropriate to be used in the upper bound method for (*QMKP*) [11]. We will compare theoretically the use of the direct expansion (I) with the direct expansion with additional constraints (II) and with the binary expansion (III). We will show that the use of (II) provides a bound which value is equal to the one computed in [11]. Conversely, we provide the proof of the non applicability of both (III) and a linear interpolation to transform (*QMKP*) into an equivalent *0-1* piecewise linear problem.

The paper is organized as follows. The next section summarizes the upper bound method developed in [11] detailing the direct expansion (I) of integer variables and the piecewise interpolation. In Section 3, the direct expansion with additional constraints (II) is applied to (*QMKP*) so as to compute the upper bound suggested by Quadri et al. [11]. Section 4 is dedicated to the binary expansion (III). We finally conclude in Section 5.

In the remainder of this paper, we adopt the following notations: letting (*P*) be an integer or a 0-1 program, we will denote by ($\overline{P}$) the continuous relaxation problem of (*P*). We let *Z*[*P*] be the optimal value of the problem (*P*) and *Z*[$\overline{P}$] the optimal value of ($\overline{P}$). Finally $\lceil x \rceil$ (resp. $\lfloor x \rfloor$) will denote the smallest (resp. highest) integer greater (resp. lower) than or equal to *x*.

**Section 2. Direct expansion of the integer variables**

In this section we summarize the upper bound method for (*QMKP*) proposed by Quadri et al. [11]. First, an equivalent formulation is obtained by using a direct expansion (I) of the integer variables $x_j$ as originally proposed by Glover [7] and by applying a piecewise interpolation to the initial objective function as discussed in [4].

The direct expansion of the integer variables $x_j$ consists in replacing each variables $x_j$ by a sum of $u_j$ *0-1* variables $y_{jk}$ such that $\sum_{k=1}^{u_j} y_{jk} = x_j$. Since the objective function *f* is separable a linear interpolation can then be applied to each objective function term $f_j$. Consequently, (*QMKP*) is equivalent to the *0-1* piecewise linear program (*MKP*):

$$(MKP)\begin{cases} \max l(y) = \sum_{j=1}^{n}\left(\sum_{k=1}^{u_j} s_{jk} y_{jk}\right) \\ \text{s. t.} \begin{vmatrix} \sum_{j=1}^{n}\left(a_{ij}\sum_{k=1}^{u_j} y_{jk}\right) \le b_i, & i = 1,\ldots,m \\ y_{jk} \in \{0,1\}, & j = 1,\ldots,n,\ k = 1,\ldots,u_j \end{vmatrix} \end{cases}$$

where, $\sum_{k=1}^{u_j} y_{jk} = x_j$, $s_{jk} = f_{jk} - f_{j,k-1}$ and $f_{jk} = c_j k - d_j k^2$.

In the second step of the algorithm, a surrogate relaxation is applied to the LP-relaxation of (*MKP*). The resultant formulation (*KP*, *w*) is the surrogate relaxation problem of (*MKP*) and can be written as:

$$(KP, w)\begin{cases} \max \sum_{j=1}^{n}\left(\sum_{k=1}^{u_j} s_{jk} y_{jk}\right) \\ \text{s. t.} \left| \begin{array}{l} \sum_{j=1}^{n}\left[\sum_{i=1}^{m} w_i a_{ij}\right]\sum_{k=1}^{u_j} y_{jk} \le \sum_{i=1}^{m} w_i b_i \\ y_{jk} \in \{0,1\}, \quad j=1,\ldots,n,\ k=1,\ldots,u_j \end{array} \right. \end{cases}$$

As proved by Glover [7], (*KP, w*) is a relaxation of (*MKP*). For any value of $w \ge 0$ $Z\left[\overline{KP,w}\right]$ the optimal value of $\left(\overline{KP,w}\right)$ constitutes an upper bound for $Z\left[\overline{MKP}\right]$ . But the value of the bound $Z\left[\overline{KP,w}\right]$ depends on the choice of the surrogate multiplier employed. It is proved in [10] that if *w** is chosen as the optimal solution of the dual of $\left(\overline{MKP}\right)$ then the optimal value of (*KP,w**) is a tight upper bound for (*QMKP*) and it is obtained in a very competitive CPU time.

Nevertheless, the direct expansion used in the above upper bound procedure results in an increase of the size of the linearized problem. Indeed, the number of *0-1* variables is equal to $\sum_{j=1}^{n} u_j$ whereas it is well known that fewer variables are included in the program so less running time is consumed. The purpose of the next sections is to try to reduce the equivalent linearized problem size, using other expansion techniques and to evaluate the impact of such techniques on the computation of the upper bound proposed in [11].

**Section 3. Direct expansion of integer variables with additional constraints**

In this section we apply to the integer variables of (*QMKP*) the direct expansion with additional constraints (II). That is each variable $x_j$ is replaced by the following expression $\sum_{k=1}^{u_j} k\, y'_{jk}$ where $y'_{jk} \in \{0,1\}$, $k = 1,\ldots,u_j$ and $j=1,\ldots,n$. Since the integer variables are now replaced by 0-1 variables, we transform the resultant 0-1 quadratic

program into a 0-1 linear problem as follows: we replace each objective function term $f_j(x_j)$ by $\sum_{k=1}^{u_j} f_{jk} y'_{jk}$ where $f_{jk} = c_j k - d_j k^2$ .

The problem (*QMKP*) is thus equivalent to the following problem ($MKP_2$):

$$(MKP_2)\begin{cases} \text{maximize } h(y') = \sum_{j=1}^{n}\left(\sum_{k=1}^{u_j} f_{jk} y'_{jk}\right) \\ \text{s. t.} \begin{cases} \sum_{j=1}^{n}\left(a_{ij}\sum_{k=1}^{u_j} y'_{jk}\right) \le b_i, & i = 1,\ldots,m \\ \sum_{k=1}^{u_j} y'_{jk} \le 1 & j = 1,...,n \\ y'_{jk} \in \{0,1\}, & j = 1,\ldots,n,\ k = 1,\ldots,u_j \end{cases} \end{cases}$$

Following the upper bound method developed in [11] we then establish the surrogate problem ($KP_2,w$) associated to ($MKP_2$). The problem ($KP_2,w$) can be written as:

$$(KP_2, w)\begin{cases} \text{maximize } h(y') = \sum_{j=1}^{n}\left(\sum_{k=1}^{u_j} f_{jk} y'_{jk}\right) \\ \text{s. t.} \begin{cases} \sum_{i=1}^{m} \sum_{j=1}^{n}\left(w_i a_{ij}\sum_{k=1}^{u_j} y'_{jk}\right) \le \sum_{i=1}^{m} w_i b_i, & \\ \sum_{k=1}^{u_j} y'_{jk} \le 1 & j = 1,...,n \\ y'_{jk} \in \{0,1\}, & j = 1,\ldots,n,\ k = 1,\ldots,u_j \end{cases} \end{cases}$$

Utilizing the optimal solution $\omega^*$ of the dual of ($\overline{MKP_2}$) as the surrogate multiplier, the optimal value of the LP-relaxation of ($KP_2, \omega^*$) provides an upper bound of (*QMKP*). The following proposition and its corollary show that this upper bound, computed in this section through the use of expansion (II), is equal to the one computed in [11].

**Proposition 3.1** *Let w be any real surrogate multiplier, w ≥ 0. The following result holds:*

$$Z\left[\overline{KP,w}\right] = Z\left[\overline{KP2,w}\right]$$

**Proof 3.1**

- We first show that: $Z\left[\overline{KP2,w}\right] \le Z\left[\overline{KP,w}\right]$

  Let $\bar{y}'$ be an optimal solution for $\left(\overline{KP2,w}\right)$. We derive from $\bar{y}'$ a solution $\bar{y}$ feasible for $\left(\overline{KP,w}\right)$, such that $h(\bar{y}') \le g(\bar{y})$.

  Let $j \in \{1,\ldots,n\}$. We denote:

  - $\bar{x}_j = \sum_{k=1}^{u_j} k\,\bar{y}'_{jk}$. Obviously $\bar{x}_j \in [0,u_j]$;
  - $p_j = \lfloor \bar{x}_j \rfloor$, where $\lfloor x \rfloor$ stands for the highest integer lower than or equal to the real *x*. $p_j$ is the integer part of $\bar{x}_j$. Note that $p_j < u_j$ since $\bar{x}_j \le u_j$;
  - $\varepsilon_j = \bar{x}_j - p_j$, the fractional part of $\bar{x}_j$: $\bar{x}_j = p_j + \varepsilon_j$. Note that $\varepsilon_j \in [0,1[$.

  We define the solution $\bar{y}$ as follows:

$$\forall j \in \{1,\ldots,n\},\quad \begin{cases} \bar{y}_{jk} = 1 & \text{for } 1 \le k \le p_j \\ \bar{y}_{j,p_j+1} = \varepsilon_j & \\ \bar{y}_{jk} = 0 & \text{for } p_j + 2 \le k \le u_j \end{cases}$$

  We immediately get: $\forall j \in \{1,\ldots,n\}, \sum_{k=1}^{u_j} k\,\bar{y}'_{jk} = \bar{x}_j = \sum_{k=1}^{u_j} \bar{y}_{jk}$. This implies that $\bar{y}$ is feasible for $\left(\overline{KP,w}\right)$ since we have:

$$\begin{aligned} \sum_{i=1}^{m} w_i \left[ \sum_{j=1}^{n} a_{ij} \left( \sum_{k=1}^{u_j} \bar{y}_{jk} \right) \right] &= \sum_{i=1}^{m} w_i \left[ \sum_{j=1}^{n} a_{ij}\,\bar{x}_j \right] \\ &= \sum_{i=1}^{m} w_i \left[ \sum_{j=1}^{n} a_{ij} \left( \sum_{k=1}^{u_j} k\,\bar{y}'_{jk} \right) \right] \\ &\le \sum_{i=1}^{m} w_i b_i \quad (\text{since } \bar{y}' \text{ is feasible for } (\overline{KP2,w})) \end{aligned}$$

  Let us now show that $h(\bar{y}') \le g(\bar{y})$. It suffices to prove that:

$$\forall j \in \{1,\ldots,n\},\ \sum_{k=1}^{u_j} f_{jk}\ \bar{y}'_{jk} \le \sum_{k=1}^{u_j} s_{jk}\ \bar{y}_{jk}$$

Let $j \in \{1,\ldots,n\}$. We first note that:

$$\begin{aligned}\sum_{k=1}^{u_j} s_{jk}\ \bar{y}_{jk} &= \left[\sum_{k=1}^{p_j} s_{jk}\right] + s_{j,p_j+1}\varepsilon_j = \left[\sum_{k=1}^{p_j}\left(f_{jk} - f_{j,k-1}\right)\right] + \left(f_{j,p_j+1} - f_{j,p_j}\right)\varepsilon_j \\ &= \left(1-\varepsilon_j\right) f_{j,p_j} + \varepsilon_j\ f_{j,p_j+1}\end{aligned}$$

In the following we name $\Lambda_j$ the quantity $\left(1-\varepsilon_j\right) f_{j,p_j} + \varepsilon_j\ f_{j,p_j+1}$. We now show that $\Lambda_j$ constitutes an overestimation of $\sum_{k=1}^{u_j} f_{jk}\ \bar{y}'_{jk}$ if we suppose as previously mentioned that $\sum_{k=1}^{u_j} k\ \bar{y}'_{jk} = p_j + \varepsilon_j$. For a given index *j*, let us consider the following linear problem $(P_j)$ and its dual problem $(D_j)$:

$$(P_j)\begin{cases}\text{maximize} & h(y') = \sum_{k=1}^{u_j} f_{jk}\, y'_{jk} \\ \text{s. t.} & \sum_{k=1}^{u_j} y'_{jk} \le 1 \\ & \sum_{k=1}^{u_j} k\ y'_{jk} = p_j + \varepsilon_j \\ & y'_{jk} \ge 0 \qquad\qquad k = 1,\ldots,u_j\end{cases}$$

$$(D_j)\begin{cases}\text{minimize} & \alpha + \left(p_j + \varepsilon_j\right)\beta \\ \text{s. t.} & \alpha + k\beta \ge f_{jk} \qquad k = 1,\ldots,u_j \\ & \alpha \ge 0\end{cases}$$

The solution $\tilde{y}$ defined by ($\tilde{y}_{j,p_j} = 1-\varepsilon_j$, $\tilde{y}_{j,p_j+1} = \varepsilon_j$, $\tilde{y}_{jk} = 0$ for $k \ne p_j$ and $k \ne p_j + 1$) is clearly feasible for $(P_j)$ and of value $\Lambda_j$. The complementary slackness conditions suggest to consider the following solution $(\tilde{\alpha}, \tilde{\beta})$ for $(D_j)$:

$$\begin{cases} \tilde{\alpha} = (p_j + 1) f_{j,p_j} - p_j f_{j,p_j+1} \\ \tilde{\beta} = f_{j,p_j+1} - f_{j,p_j} \end{cases}$$

We now prove that $(\tilde{\alpha}, \tilde{\beta})$ is feasible for $(D_j)$. Since its value in $(D_j)$ is clearly $\Lambda_j$, the feasibility of $(\tilde{\alpha}, \tilde{\beta})$ for $(D_j)$ will imply by duality in linear programming that $\Lambda_j = Z[P_j]$ . $(\tilde{\alpha}, \tilde{\beta})$ is feasible for $(D_j)$ if and only if: $\tilde{\alpha} \geq 0$ and $\forall k \in \{1, \ldots, u_j\}$ $\tilde{\alpha} + k\tilde{\beta} \geq f_{jk}$. Noting that $f_{j,0} = 0$, we only have to prove the following result:

$$\forall k \in \{0, \ldots, u_j\}, \ \tilde{\alpha} + k\tilde{\beta} \geq f_{jk}$$

Let us denote $\Pi_j(k) = (p_j + 1 - k) f_{j,p_j} + (k - p_j) f_{j,p_j+1} - f_{jk}$. We thus have to prove that:

$$\forall k \in \{0, \ldots, u_j\}, \ \Pi_j(k) \geq 0$$

Using the definition of $f_{jk}$ ($f_{jk} = c_j k - d_j k^2$), we find with some easy algebra:

$$\Pi_j(k) = d_j k^2 - (2p_j + 1) d_j k + d_j p_j (p_j + 1)$$

Thus, $\Pi_j(k)$ is a degree 2 polynomial in *k*. Its discriminant is equal to $d_j^2$, so $\Pi_j(k)$ admits two roots: $p_j$ and $p_j + 1$. Since the coefficient of $k^2$ in $\Pi_j(k)$, $d_j$, is nonnegative, $\Pi_j(k)$ is also nonnegative for *k* outside the roots, namely for $k \leq p_j$ and $k \geq p_j + 1$. Since we only consider integer values of *k*, we get: $\forall k \in \{0, \ldots, u_j\}$, $\Pi_j(k) \geq 0$, which implies that $(\tilde{\alpha}, \tilde{\beta})$ is feasible for $(D_j)$ and therefore that $\Lambda_j$ is an overestimation of $\sum_{k=1}^{u_j} f_{jk} \, \bar{y}'_{jk}$ :

$$\sum_{k=1}^{u_j} f_{jk} \, \bar{y}'_{jk} \leq \Lambda_j = \sum_{k=1}^{u_j} s_{jk} \, \bar{y}_{jk} \ .$$

This last inequality implies that $h(\bar{y}') \leq g(\bar{y})$ and therefore that:

$$Z[\overline{KP2, w}] \leq Z[\overline{KP, w}].$$

- We now show that: $Z\left[\overline{KP,w}\right] \leq Z\left[\overline{KP2,w}\right]$

  The proof is analogous to the one of the previous point. Let $\bar{y}$ be an optimal solution for $\left(\overline{KP,w}\right)$. We derive from $\bar{y}$ a solution $\bar{y}'$ feasible for $\left(\overline{KP2,w}\right)$, such that $g(\bar{y}) = h(\bar{y}')$.

  Since $\left(\overline{KP,w}\right)$ is linear continuous knapsack problem, $\bar{y}$ admits at most one fractional component, $\bar{y}_{j_0,k_0}$. If such a fractional variable $\bar{y}_{j_0,k_0}$ exists, we denote:

$$\bar{x}_{j_0} = \sum_{k=1}^{u_{j_0}} \bar{y}_{j_0,k}\ ,\ p_{j_0} = \lfloor \bar{x}_{j_0} \rfloor \text{ and } \varepsilon_{j_0} = \bar{x}_{j_0} - p_{j_0}\ .$$

  $p_{j_0}$ and $\varepsilon_{j_0}$ are respectively the integer and fractional parts of $\bar{x}_{j_0}$. Since $\bar{x}_{j_0} \leq u_{j_0}$, $p_{j_0} < u_{j_0}$. We begin to define $\bar{y}'$:

$$\begin{cases} \bar{y}'_{j_0,p_{j_0}} = 1 - \varepsilon_{j_0} \\ \bar{y}'_{j_0,p_{j_0+1}} = \varepsilon_{j_0} \\ \bar{y}'_{j_0,k} = 0 \qquad \forall\, k \in \{1,\ldots,u_j\} - \{p_{j_0}, p_{j_0}+1\} \end{cases}$$

  For all $j \in \{1,\ldots,n\} - \{j_0\}$ (or for all $j \in \{1,\ldots,n\}$ if $\bar{y}$ has no fractional component, i.e. if $\bar{y}$ is integer), $\bar{y}_{jk}$ belongs to $\{0,1\}$ ($\forall k$) and therefore $\bar{x}_j = \sum_{k=1}^{u_j} \bar{y}_{jk}$ is an integer lower than or equal to $u_j$. We then finish to define $\bar{y}'$:

$$\forall j \in \{1,\ldots,n\} - \{j_0\}, \begin{cases} \text{if } \bar{x}_j \geq 1,\ \bar{y}'_{j,\bar{x}_j} = 1 \text{ and } \bar{y}'_{jk} = 0\ \forall k \in \{1,\ldots,u_j\} - \{\bar{x}_j\} \\ \text{if } \bar{x}_j = 0, \bar{y}'_{jk} = 0\ \forall k \in \{1,\ldots,u_j\} \end{cases}$$

  We immediately get:

$$\begin{cases} \forall j \in \{1,\ldots,n\} - \{j_0\}, \sum_{k=1}^{u_j} k\, \bar{y}'_{jk} = \bar{x}_j = \sum_{k=1}^{u_j} \bar{y}_{jk} \\ \text{for } j = j_0, \sum_{k=1}^{u_{j0}} k\, \bar{y}'_{j_0,k} = p_{j_0}(1-\varepsilon_{j_0}) + (p_{j_0}+1)\varepsilon_{j_0} = p_{j_0} + \varepsilon_{j_0} = \bar{x}_{j_0} = \sum_{k=1}^{u_{j0}} \bar{y}_{j_0,k} \end{cases}$$

which implies that $\bar{y}'$ verifies the surrogate constraint of $(\overline{KP2,w})$ (since $\bar{y}$ verifies the one of $(\overline{KP,w})$). Since $\forall j \in \{1,\ldots,n\}, \sum_{k=1}^{u_j} \bar{y}'_{jk} \le 1$, $\bar{y}'$ is feasible for $(\overline{KP2,w})$.

Let us now prove that $g(\bar{y}) = h(\bar{y}')$.

- For $j = j_0$, we have:

$$\begin{aligned} \sum_{k=1}^{u_{j0}} f_{j_0,k} \bar{y}'_{j_0,k} &= f_{j_0,p_{j_0}} \bar{y}'_{j_0,p_{j_0}} + f_{j_0,p_{j_0}+1} \bar{y}'_{j_0,p_{j_0}+1} \\ &= (1-\varepsilon_{j_0}) f_{j_0,p_{j_0}} + \varepsilon_{j_0} f_{j_0,p_{j_0}+1} = \Lambda_{j_0} \end{aligned}$$

  Now: $\sum_{k=1}^{u_{j0}} s_{j_0,k} \bar{y}_{j_0,k} = \sum_{k=1}^{p_{j_0}} s_{j_0,k} + s_{j_0,p_{j_0}+1} \varepsilon_{j_0} \left(= \Lambda_{j_0}\right)$

  This equality holds since $(\overline{KP,w})$ is a linear continuous knapsack problem: for a given index *j*, all the variables $y_{jk}$, $k \in \{1,\ldots,u_j\}$, have the same coefficient in the knapsack constraint ($\sum_{i=1}^{m} w_i a_{ij}$) and the coefficients of these variables in the objective function *g*(*y*) are in decreasing order (because of the concavity of function $f_j$): $s_{j1} \ge s_{j2} \ge \ldots \ge s_{j,u_j}$.

  Therefore the following result holds: $\sum_{k=1}^{u_{j0}} f_{j_0,k} \bar{y}'_{j_0,k} = \sum_{k=1}^{u_{j0}} s_{j_0,k} \bar{y}_{j_0,k}$.

- For $j \ne j_0$, we have: $\sum_{k=1}^{u_j} f_{jk} \bar{y}'_{jk} = f_{j,\bar{x}_j}$ and $\sum_{k=1}^{u_j} s_{jk} \bar{y}_{jk} = \sum_{k=1}^{\bar{x}_j} s_{jk} = f_{j,\bar{x}_j}$.

  This last equality holds for the same reason as in the previous

paragraph (the coefficients of variables $y_{jk}$ in the unique knapsack constraint are identical and the coefficients of these variables in the objective function are in decreasing order). This implies that: $\sum_{k=1}^{u_j} f_{jk}\bar{y}'_{jk} = \sum_{k=1}^{u_j} s_{jk}\bar{y}_{jk}$ .

We thus get: $g(\bar{y}) = h(\bar{y}')$, wich implies the result $Z\left[\overline{KP, w}\right] \leq Z\left[\overline{KP2, w}\right]$.

□

**Corollary 3.2** *The upper bound obtained by the surrogate linearization [11] (using expansion (I)) and the upper bound obtained by the surrogate linearization using expansion (II) are equal:* $Z\left[\overline{KP, w^*}\right] = Z\left[\overline{KP_2, \omega^*}\right]$ *, where* $w^*$ *and* $\omega^*$ *respectively stand for the optimal surrogate multiplier of the LP-relaxation of problems (MKP) and (MKP2).*

**Proof 3.2** The result of Proposition 3.1 is true whatever the surrogate multiplier *w* is used, so it remains true if we successively consider $w^*$ and $\omega^*$ as a surrogate multiplier. □

We thus have proved in this Section that solving (*QMKP*) with a linearization technique and a surrogate relaxation may be equivalently done with expansion (I) or with expansion (II) of the integer variables. We proved that expansions (I) and (II) are equivalent in the sense that they provide the same upper bound. However, expansion (II) involves *n* added constraints without improving the quality of the upper bound in comparison with expansion (I).

**Section 4. Binary expansion of integer variables**

This section is dedicated to the use of a binary expansion (referred as expansion (III)) of this integer variables in the upper bound procedure developed in [11]. Such expansion consists in rewriting each integer variable $x_j$ ($j$=1,...,$n$) as:

$$x_j = \sum_{k=0}^{\lceil \log(u_j) \rceil} 2^k z_{jk} \text{ where } z_{jk} \in \{0,1\} \qquad (1)$$

Using (1) involves only $\sum_{j=1}^{n} \lceil \log_2(u_j) \rceil$ 0-1 variables instead of $\sum_{j=1}^{n} u_j$ 0-1 variables when a direct expansion is applied.

As previously mentioned, the aim of this study is to compare other expansion techniques in the upper bound process developed in [11]. Since, the variables are now binary the next step of the algorithm of Quadri et al. [11] concerns with a piecewise linear interpolation of the objective function so as to obtain an equivalent 0-1 linear problem with only $\sum_{j=1}^{n} \lceil \log_2(u_j) \rceil$ 0-1 variables.

The following proposition shows the non applicability of both expansion (III) and a linear interpolation to transform (*QMKP*) into an equivalent *0-1* piecewise linear problem through the use of only $\sum_{j=1}^{n} \lceil \log_2(u_j) \rceil$ 0-1 variables.

**Proposition 4.1** *Considering the integer quadratic multi-knapsack problem (QMKP). If a binary expansion is used to rewrite the integer variables of (QMKP) as* $\sum_{j=1}^{n} \lceil \log_2(u_j) \rceil$ *0-1 variables then an equivalent 0-1 piecewise linear problem can not be established applying a linear interpolation.*

**Proof 4.1** Assume that it is possible to replace each function $f_j(x_j) = c_j x_j - d_j x_j^2$, for all $j$ from 1 to $n$, by a linear function $g_j(z_j)$ using exactly $\lceil \log_2(u_j) \rceil$ 0-1 variables. We denote by (*H*) this assumption.

If (*H*) is true then it should exist coefficients $g_{jk}$ such that:

$$f_j(x_j) = c_j x_j - d_j x_j^2 = \sum_{k=1}^{\lceil \log_2(u_j) \rceil} g_{jk} z_{jk}, \forall x_j \in \{1 \ldots u_j\} \quad (2)$$

with $x_j = \sum_{k=0}^{\lceil \log(u_j) \rceil} 2^k z_{jk},\ z_{jk} \in \{0,1\}$ (3).

The assumption (*H*) must be satisfied for all problem data. Let us set $u_j = 3,\ c_j = 10$ and $d_j = 4$.

Consequently,

- if $x_j = 1$ (cf. (2) and (3)) then $g_{j1} = 6$ (i), $z_{j1} = 1$ and $z_{j2} = 0$.
- if $x_j = 2$ (cf. (2) and (3)) then $g_{j2} = 4$ (ii), $z_{j1} = 0$ and $z_{j2} = 1$.
- if $x_j = 3$ then (cf. (3)) then $z_{j1} = z_{j2} = 1$ which implies (cf. (2)) that $3c_j - 9\ d_j = g_{j1} + g_{j2} = -6$ (iii).

The equations system (I), (II) and (III) has clearly no solution. Consequently there is a contradiction with (*H*). □

Proposition 4.1 shows the non applicability of a binary expansion of the integer variables for (*QMKP*) so as to transform the initial problem into a 0-1 linear program. Consequently, the upper bound method proposed in [11] can not be applied together with expansion (III).

**Section 5. Concluding remarks**

In this paper we have theoretically compared the use of three techniques to rewrite integer variables into zero-one variables in an upper bound procedure for (*QMKP*) developed by Quadri et al. [11], which provides, to the best of our knowledge a bound closer to the optimum than the existing methods. More specifically, we have compared a direct expansion of the integer variables, originally employed in [11] with a direct expansion with additional constraints (II) and with a binary expansion (III). We have proved that (II) provides the same upper bound as the one computed in [11] whereas it involves *n* additional constraints. We therefore do not expect an improvement of the upper bound computational time. Finally, we provide a proof of

the non applicability of both (III) and a linear interpolation to transform (*QMKP*) into an equivalent *0-1* piecewise linear problem through the use of only $\sum_{j=1}^{n}\lceil \log_2(u_j) \rceil$ 0-1 variables.

A possible way to get a further decrease of the number of 0-1 variables would be the use of a classical linearization technique as suggested by Foret (1959) [6]. This linearization scheme would first consist in applying (III) to the integer variables and then using the basic linearization technique for 0-1 variables. Nevertheless, preliminary computational experiments have shown that the bound provided by this technique is of worst quality than the one computed in [11]. Moreover, this process seems to be more time consuming than [11].

**References**


[1] A. Billionnet, A. Faye, E. Soutif, A new upper bound for the 0-1 quadratic knapsack problem. *European Journal of Operational Research*, 112 (3) (1999) 664-672.

[2] K. Bretthauer, B. Shetty, The nonlinear knapsack problem – algorithms and applications. *European Journal of Operational Research*, 138 (3) (2002) 459-472.

[3] P. Chaillou, P. Hansen, Y. Mathieu, Best network flow bounds for the quadratic knapsack problem. *Lecture Notes in Mathematics*, 1403 (1986) 226-235.

[4] M. Djerdjour, K. Mathur, H. Salkin, A surrogate-based algorithm for the general quadratic multidimensional knapsack. *Operations Research Letters*, 7 (1988) 253-257.

[5] B. Faaland, An integer programming algorithm for portfolio selection. *Management Science*, 20 (10) (1974) 1376-1384.

[6] R. Foret, Applications de l'algèbre de Boole en recherché opérationnelle. *Revue française d'automatique d'informatique et de recherche opérationnelle*, 4 (1959) 5-36.

[7] F. Glover, Improved linear programming formulations of nonlinear integer problems. *Management Science*, 22 (4) (1975) 455-460.

[8] F. Körner, Integer quadratic programming. *European Journal of Operational Research*, 19 (2) (1985) 268-273.

[9] K. Mathur, H. Salkin, S. Morito, A branch and search algorithm for a class of nonlinear knapsack problems. *Operations Research Letters*, 2 (4) (1983) 155-160.

[10] Y. Nakagawa, Optimal reliability allocation by branch and bound techniques. *IIE Trans. On Reliability*, 27 (31) (1978) 27-35.

[11] D. Quadri, E. Soutif, P. Tolla, Upper bounds for large scale integer quadratic multidimensional knapsack problems. *International Journal of Operations Research*, 4 (3) (2007) 1-9.

[12] K. Stecke, Nonlinear integer production planning problems. *Management Science*, 29 (3) (1983) 273-288.