

Exact Solution Methods for the k -Item Quadratic Knapsack Problem

Lucas Létocart¹ and Angelika Wiegele²(✉)

¹ Université Paris 13, Sorbonne Paris Cité, LIPN, CNRS, (UMR 7030),
93430 Villetaneuse, France

lucas.letocart@lipn.univ-paris13.fr

² Alpen-Adria-Universität Klagenfurt, 9020 Klagenfurt am Wörthersee, Austria
angelika.wiegele@aau.at

Abstract. The purpose of this paper is to solve the 0–1 k -item quadratic knapsack problem ($kQKP$), a problem of maximizing a quadratic function subject to two linear constraints. We propose an exact method based on semidefinite optimization. The semidefinite relaxation used in our approach includes simple rank one constraints, which can be handled efficiently by interior point methods. Furthermore, we strengthen the relaxation by polyhedral constraints and obtain approximate solutions to this semidefinite problem by applying a bundle method. We review other exact solution methods and compare all these approaches by experimenting with instances of various sizes and densities.

Keywords: Quadratic programming · 0–1 knapsack · k -cluster · Semidefinite programming

1 Introduction

The 0–1 k -item quadratic knapsack problem consists of maximizing a quadratic objective function subject to a linear capacity constraint with an additional equality cardinality constraint:

$$(kQKP) \begin{cases} \max f(x) = \sum_{i=1}^n \sum_{j=1}^n c_{ij} x_i x_j \\ \text{s.t.} \quad \sum_{j=1}^n a_j x_j \leq b & (1) \\ \sum_{j=1}^n x_j = k & (2) \\ x_j \in \{0, 1\} \quad j = 1, \dots, n \end{cases}$$

where n denotes the number of items, and all the data, k (number of items to be filled in the knapsack), a_j (weight of item j), c_{ij} (profit associated with the selection of items i and j) and b (capacity of the knapsack) are nonnegative integers. Without loss of generality, matrix $C = (c_{ij})$ is assumed to be symmetric. Moreover, we assume that $\max_{j=1, \dots, n} a_j \leq b < \sum_{j=1}^n a_j$ in order to avoid either trivial solutions or variable fixing via constraint (1). Let us denote by k_{max} the largest number of items which could be filled in the knapsack, that is the largest number of the smallest a_j whose sum does not exceed b . We can assume that

$k \in \{2, \dots, k_{max}\}$, where k_{max} can be found in $O(n)$ time [2]. Otherwise, either the value of the problem is equal to $\max_{i=1, \dots, n} c_{ii}$ (for $k = 1$), or the domain of $(kQKP)$ is empty (for $k > k_{max}$).

$(kQKP)$ is an NP-hard problem as it includes two classical NP-hard subproblems, the k -cluster problem [6] by dropping constraint (1), and the quadratic knapsack problem [18] by dropping constraint (2). Even more, the work of Bhaskara et al. [4] indicates that approximating k -cluster within a polynomial factor might be a harder problem than Unique Games. Rader and Woeginger [19] state negative results concerning the approximability of QKP if negative cost coefficients are present.

Applications of $(kQKP)$ cover those found in previous references for k -cluster or classical quadratic knapsack problems (e.g., task assignment problems in a client-server architecture with limited memory), but also multivariate linear regression and portfolio selection. Specific heuristic and exact methods including branch-and-bound and branch-and-cut with surrogate relaxations have been designed for these applications (see, e.g., [3, 5, 9, 17, 22]).

The purpose of this paper is twofold.

1. We introduce a new algorithm for solving $(kQKP)$ and
2. we briefly review other state of the art methods and compare the methods by running numerical experiments.

Our new algorithm consists of a branch-and-bound framework using

- a combination of a semidefinite relaxation and polyhedral cutting planes to obtain tight upper bounds and
- fast hybrid heuristics [16] for computing high quality lower bounds.

This paper is structured as follows. In Sect. 2 a semidefinite relaxation is derived, followed by a discussion of solving the semidefinite problems in Sect. 3. The relaxation is used inside a branch-and-bound framework, the various components of this branch-and-bound algorithm are discussed in Sect. 4. Other methods for solving $(kQKP)$ and numerical results are presented in Sect. 5 and Sect. 6 concludes.

Notation. We denote by e the vector of all ones of appropriate size. $\text{diag}(X)$ refers to diagonal of X as a vector and $\text{Diag}(v)$ is the diagonal matrix having diagonal v .

2 A Semidefinite Relaxation of $(kQKP)$

In order to develop a branch-and-bound algorithm for solving $(kQKP)$ to optimality we aim in finding strong upper bounds. Semidefinite optimization proved to provide such strong bounds, see e.g. [1, 20, 21].

A straightforward way to obtain a semidefinite relaxation is the following. Express all functions involved as quadratic functions, i.e. functions in xx^t , replace

the product xx^t by a matrix X and get rid of non-convexities by relaxing $X = xx^t$ to $X \succeq xx^t$.

Hence, we apply the following changes:

- Replace the constraint $e^t x = k$ by the constraint $(e^t x - k)^2 = 0$.
- As for the capacity constraint, define b' to be the sum of the weights of the k smallest items. Clearly, $b' \leq a^t x$ is a valid constraint for $(kQKP)$. Combining this redundant constraint with the capacity constraint we obtain $(b' - a^t x)(b - a^t x) \leq 0$.
- Transform the problem to a ± 1 problem by setting $y = 2x - e$.
- Relax the problem by relaxing $Y = yy^t$ to $Y \succeq yy^t$, i.e., dropping the constraint Y being of rank one.

This procedure yields the following semidefinite problem:

$$\begin{aligned}
 \max \quad & \langle \tilde{C}, Y \rangle \\
 \text{s.t.} \quad & \text{diag}(Y) = e \\
 & \langle \tilde{E}, Y \rangle = 0 \\
 & \langle \tilde{A}, Y \rangle \leq (b - b')^2 \\
 & Y \succeq 0
 \end{aligned} \tag{SDP_1}$$

with $\tilde{E} = \tilde{e}\tilde{e}^t$, $\tilde{e} = \begin{pmatrix} n - 2k \\ e \end{pmatrix}$, $\tilde{A} = \tilde{a}\tilde{a}^t$, $\tilde{a} = \begin{pmatrix} a^t e - (b + b') \\ a \end{pmatrix}$, and appropriate \tilde{C} .

Observation 1. (SDP_1) has no strictly feasible point and thus Slater’s condition does not hold.

Proof. Note that $\langle \tilde{E}, Y \rangle = \tilde{e}^t Y \tilde{e} = 0$ together with $Y \succeq 0$ implies Y being singular and thus every feasible solution is singular. □

Observe that $\tilde{e} = \begin{pmatrix} n - 2k \\ e \end{pmatrix}$ is an eigenvector to the eigenvalue 0 of every feasible Y . Now consider matrix $V = \begin{pmatrix} \frac{1}{2k-n} e^t \\ I_n \end{pmatrix}$. V spans the orthogonal complement of the span of eigenvector \tilde{e} . Set $Y = V X V^t$ to “project out” the 0-eigenvalue and consider the $n \times n$ matrix X instead of the $(n + 1) \times (n + 1)$ matrix Y . The relationship between X and Y is simply given by

$$Y = V X V^t = \begin{pmatrix} \frac{1}{(2k-n)^2} e^t X e & \frac{1}{2k-n} (X e)^t \\ \frac{1}{2k-n} X e & X \end{pmatrix}.$$

Looking at the effect of the constraints of (SDP_1) on matrix X , we derive the following conditions.

– From $\text{diag}(Y) = e$ we obtain the constraints

$$\begin{aligned} e^t X e &= (2k - n)^2 \\ \text{diag}(X) &= e \end{aligned}$$

– The left-hand side of constraint $\langle \tilde{E}, Y \rangle = 0$ translates into

$$\langle \tilde{E}, Y \rangle = \langle \tilde{E}, V X V^t \rangle = \langle V^t \tilde{E} V, X \rangle = \langle 0, X \rangle = 0$$

and the constraint becomes obsolete.

– Constraint $\langle \tilde{A}, Y \rangle \leq (b - b')^2$ yields the following.

$$\begin{aligned} \langle \tilde{A}, Y \rangle &= \langle \tilde{A}, V X V^t \rangle = \langle V^t \tilde{A} V, X \rangle = \\ &= \left\langle \left(\frac{a^t e - (b + b')}{2k - n} e + a \right) \left(\frac{a^t e - (b + b')}{2k - n} e + a \right)^t, X \right\rangle \end{aligned}$$

Hence,

$$\left\langle \left(\frac{a^t e - (b + b')}{2k - n} e + a \right) \left(\frac{a^t e - (b + b')}{2k - n} e + a \right)^t, X \right\rangle \leq (b - b')^2$$

Defining $\bar{a} = \left(\frac{a^t e - (b + b')}{2k - n} e + a \right)$ we finally obtain

$$\begin{aligned} \max \quad & \langle \bar{C}, X \rangle \\ \text{s.t.} \quad & \text{diag}(X) = e \\ & \langle E, X \rangle = (2k - n)^2 \\ & \langle A, X \rangle \leq (b - b')^2 \\ & X \succeq 0 \end{aligned} \tag{SDP}$$

where $E = e e^t$, $A = \bar{a} \bar{a}^t$, and appropriate cost matrix \bar{C} .

Strengthening the Relaxation. Since we derived a relaxation from a problem in ± 1 variables, we can further tighten the bound by adding the well known *triangle inequalities* to the semidefinite relaxation (SDP). These are for any triple $1 \leq i < j < k \leq n$:

$$\begin{aligned} x_{ij} + x_{ik} + x_{jk} &\geq -1 \\ -x_{ij} - x_{ik} + x_{jk} &\geq -1 \\ -x_{ij} + x_{ik} - x_{jk} &\geq -1 \\ x_{ij} - x_{ik} - x_{jk} &\geq -1 \end{aligned} \tag{1}$$

For several problems formulated in ± 1 variables adding these constraints significantly improves the bound, see e.g. [20]. The set of matrices satisfying all triangle-inequalities is called the *metric polytope* and is denoted by *MET*. Thus, the strengthened semidefinite relaxation reads

$$\begin{aligned}
 & \max \quad \langle \bar{C}, X \rangle \\
 & \text{s.t.} \quad \text{diag}(X) = e \\
 & \quad \quad \langle E, X \rangle = (2k - n)^2 \\
 & \quad \quad \langle A, X \rangle \leq (b - b')^2 \\
 & \quad \quad X \in MET \\
 & \quad \quad X \succeq 0
 \end{aligned} \tag{SDP_{MET}}$$

3 Solving the Semidefinite Relaxations

3.1 Solving the Basic Relaxation (SDP)

The most prominent methods for solving semidefinite optimization problems are interior point methods. The interior point method is an iterative algorithm where in each iteration Newton’s method is applied in order to compute new search directions.

Consider the constraints $\mathcal{A}(X) = (\cdot)$ with $\mathcal{A}(X) = \begin{pmatrix} \langle A_1, X \rangle \\ \langle A_2, X \rangle \\ \vdots \\ \langle A_m, X \rangle \end{pmatrix}$. In each iter-

ation we determine a search direction Δy (y are variables in the dual semidefinite problem) by solving the system $M\Delta y = rhs$ where

$$m_{ij} = \text{trace}(Z^{-1}A_jXA_i).$$

Z denotes the (positive definite) matrix variable of the dual semidefinite program.

Forming this system matrix requires $O(mn^3 + m^2n^2)$ steps and is among the most time-consuming operations inside the interior point algorithm. (The other time-consuming steps are maintaining positive definiteness of the matrices X and Z and linear algebra operations such as forming inverse matrices.)

The primal-dual pair of (SDP) in variables (X, s, y, Z, t) is given as follows.

$$\begin{aligned}
 & \max \{ \langle \bar{C}, X \rangle \\
 & \text{s.t.} \quad \text{diag}(X) = e, \langle E, X \rangle = (2k - n)^2, \langle A, X \rangle + s = (b - b')^2, X \succeq 0, s \geq 0 \}
 \end{aligned}$$

$$\begin{aligned}
 & \min \{ e^t y_{1:n} + (n - 2k)^2 y_{n+1} + (b - b')^2 y_{n+2} \\
 & \text{s.t.} \quad \text{Diag}(y_{1:n}) + y_{n+1}E + y_{n+2}A - Z = \bar{C}, y_{n+2} - t = 0, Z \succeq 0, t \geq 0 \}
 \end{aligned}$$

Hence, the set of constraints is rather simple and the system matrix M reads

$$\begin{pmatrix} Z^{-1} \circ X & \text{diag}(Z^{-1}EX) & \text{diag}(Z^{-1}AX) \\ \text{diag}(Z^{-1}EX)^t & \langle E, Z^{-1}EX \rangle & \langle E, Z^{-1}AX \rangle \\ \text{diag}(Z^{-1}AX)^t & \langle A, Z^{-1}EX \rangle & \langle A, Z^{-1}AX \rangle + \frac{s}{t} \end{pmatrix}.$$

Even more, all data matrices have rank one which can be exploited when computing the inner products, e.g.,

$$\langle A, Z^{-1}AX \rangle = \text{trace}(\bar{a}\bar{a}^t Z^{-1}\bar{a}\bar{a}^t X) = (\bar{a}^t Z^{-1}\bar{a})(\bar{a}^t X \bar{a})$$

Thus, the computation of the inner products of the matrices simplifies and computing the system matrix can be reduced from $O(mn^3 + m^2n^2)$ to $O(mn^2 + m^2n)$. And since $m = n + 2$ in our case, we end up with $O(n^3)$.

Hence, (SDP) can be solved efficiently by interior point methods.

3.2 Solving the Strengthened Relaxation (SDP_{MET})

Problem (SDP_{MET}) has a considerably larger number of constraints than (SDP). Remember that $X \in MET$ is described by $4\binom{n}{3}$ linear inequalities and thus solving (SDP_{MET}) by interior point methods is intractable. An alternative has been proposed in [12]. Therein the concept of *bundle methods* is used, in order to obtain an approximate optimizer on the dual functional and thus getting a valid upper bound on (SDP_{MET}), leading to a valid upper bound on ($kQKP$).

Bundle methods have been developed to minimize nonsmooth convex functions. To characterize the problem to be solved, an oracle has to be supplied that evaluates the function at a given point and computes an ϵ -subgradient. The set of points, function values, and subgradients is collected in a “bundle”, which is used to construct a cutting plane model minorizing the function to be minimized. By doing a sequence of *descent steps* the cutting plane model is refined and one gets closer to the minimizer of the function.

We will apply the bundle method to minimize the dual functional of (SDP_{MET}). Let

$$\mathcal{X} = \{X \succeq 0: \text{diag}(X) = e, \langle E, X \rangle = (2k - n)^2, \langle A, X \rangle \leq (b - b')^2\}$$

i.e., the feasible region of (SDP). We introduce the dual functional

$$\begin{aligned} f(\gamma) &= \max_{X \in \mathcal{X}} \{ \langle \bar{C}, X \rangle + \gamma^t (e - \mathcal{T}(X)) \} \\ &= e^t \gamma + \max_{X \in \mathcal{X}} \langle \bar{C} - \mathcal{T}^t(\gamma), X \rangle \end{aligned} \tag{2}$$

where $\mathcal{T}(X) \leq e$ denotes the triangle inequalities (1). Minimizing $f(\gamma)$ over $\gamma \geq 0$ gives a valid upper bound on ($kQKP$). In fact, any $\tilde{\gamma} \geq 0$ gives a valid upper bound

$$z^* = \min_{\gamma \geq 0} f(\gamma) \leq f(\tilde{\gamma}) \text{ for any } \tilde{\gamma} \geq 0.$$

Since we use this bound inside a branch-and-bound framework, this allows us to stop early and prune a node as soon as $f(\tilde{\gamma})$ is smaller than some known lower bound. Furthermore, we do not rely on getting to the optimum. We will stop once we are “close” to optimum and branch, rather than investing time in dropping the bound by a tiny number.

Evaluating function (2) (the most time consuming step in the bundle method) amounts in solving (SDP) (with varying cost matrix), which can be done efficiently as discussed in the previous section. Having the maximizer X^* of (SDP), i.e. the function evaluation, a subgradient is given by $g^* = e - \mathcal{T}(X^*)$.

Dynamic Version of the Bundle Method. The number of variables γ in (2) is $4\binom{n}{3}$. This number is substantially larger than the dimension of the problem and we are interested only in those inequalities that are likely to be active at the optimum. Thus, we do not consider *all* triangle inequalities but work with a subset that is updated on a regular basis, say every fifth descent step. The update consists of

1. adding the m inequalities being most violated by the current iterate X and
2. removing constraints with γ close to 0 (an indicator for an inactive constraint).

In this way we are able to efficiently run the bundle algorithm by keeping the size of the variable vector γ reasonably small.

4 Branch and Bound

We develop an exact solution method for solving ($kQKP$) by designing a branch-and-bound framework using relaxation (SDP_{MET}) discussed above for getting upper bounds.

The remaining tools of our branch-and-bound algorithm are described in this section.

4.1 Heuristics for Obtaining Lower Bounds

We use two heuristics to obtain a global lower bound inside our algorithm: one that is executed at the root node and another one that is called at each other node in the branch-and-bound tree.

As a heuristic method at the root node we chose the primal heuristic denoted by H_{pri} in [16], which is an adaption of a well-known heuristic developed by Billionnet and Calmels [7] for the classical quadratic knapsack problem (QKP). This primal heuristic combines a greedy algorithm with local search.

At each node of the branch-and-bound tree, we apply a variable fixation heuristic inspired from H_{sdp} [16]. This heuristic method uses the solution of the semidefinite relaxation obtained at each node, it fixes variables under some threshold $\epsilon > 0$ to zero and applies the primal heuristic over the reduced problem. It updates the solution by performing a fill-up and exchange procedure over the

unreduced problem. This procedure iterates, increasing ϵ at each iteration, until the reduced problem is empty.

Both heuristics, the primal and the variable fixation one, are very fast and take only hundredths of a second for sizes of our interest.

4.2 Branching Rule and Search Strategy

As a branching variable we choose the “most fractional” variable, i.e., $v = \operatorname{argmin}_i |\frac{1}{2} - x_i|$. The vector x is extracted from matrix X given by the semidefinite relaxation.

We traverse the search tree in a *best first search* manner, i.e., we always consider the node in the tree having the smallest upper bound.

4.3 Speed up for Small k

Whenever k , the number of items to be filled in the knapsack, is small, a branch-and-prune algorithm is triggered in order to speed-up the approach. No relaxation is performed at each node of the branch-and-prune tree and a fast depth first search strategy, in priority fixing variables to one, is implemented. We only check the feasibility of the current solution through the cardinality and capacity constraints.

This branch-and-prune approach is very fast, at most a few seconds, for very small k . So we embedded it into our branch-and-bound algorithm and run it at nodes where the remaining number of items to be filled in the current knapsack is very small (less or equal than 5 in practice). To solve the original problem, we can also replace the global branch-and-bound method using this branch-and-prune approach for small initial values of k , in practice we choose $k \leq 10$.

5 Numerical Results

We coded the algorithm in C++. For the function evaluation (i.e., solving (SDP)) we implemented a predictor-corrector variant of an interior point algorithm [14]. We use the ConicBundle Library of Ch. Helmberg [13] as framework for the bundle method to solve (SDP_{MET}).

We compare our method ($B\&C$) to:

- (Cplex): IBM CPLEX solver, version 12.6.2 [11], with default settings. The original nonconvex 0–1 quadratic problem is given directly to CPLEX which is now able to deal with such a formulation.
- (MIQCR+Cplex): our implementation of the MIQCR method [8]. MIQCR uses a semidefinite relaxation in order to obtain a problem having a convexified objective function; the resulting convex integer problem can then be solved by standard solvers. We use the CSDP solver [10] for solving the semidefinite relaxation to convexify the objective function, and IBM CPLEX 12.6.2 [11] with default settings to solve the reformulated convex problem.

– (BiqCrunch): Also BiqCrunch [15] is an algorithm based on semidefinite and polyhedral relaxations within a branch-and-bound framework. In BiqCrunch a quadratic regularization term is added to the objective function of the semi-definite problem and a quasi-Newton method is used to compute the bounds. We use the BiqCrunch solver enhanced with our primal and variable fixation heuristics described in Sect. 4.1.

All experiments have been performed on an Intel i7-2600 quad core 3.4 GHz with 8 GB of RAM, using only one core. The computational results have been obtained for randomly generated instances from [16] with up to 150 variables. We choose $k \in \{1, \dots, \lfloor \frac{n}{4} \rfloor\}$, b , a_j and c_{ij} are positive integers. The time limit for each approach is 3h.

In Table 1 we display the run time of the overall algorithm, the gap at the root node, and the number of nodes produced during the branch-and-bound algorithm for each method. Each line of Table 1 represents average values over 10 instances for each n and δ where n is the number of variables and δ is the density of the profit matrix C . We put the number of instances solved within the time limit into brackets in case not all 10 instances could be solved. Average values are computed only over instances solved within the time limit.

The numerical experiments demonstrate that the methods having semidefinite optimization inside clearly outperform Cplex. In fact, Cplex already fails to solve all instances of size $n = 50$ within the time limit.

Table 1. Numerical results comparing four approaches. (Time limit: 3h)

n	δ	Cplex			MIQCRC+Cplex			BiqCrunch			B&C		
		Gap root %	Time (s)	#Nodes	Gap root %	Time (s)	#Nodes	Gap root %	Time (s)	#Nodes	Gap root %	Time (s)	#Nodes
50	25	102.7	3.7	3426.9	30.5	1.0	621.2	7.4	21.4	79.6	0.9	72.4	11.6
50	50	150.8	150.8	77407.9	1.0	1276.6	136.8	4.9	26.0	136.8	1.3	110.3	88.1
75	25	230.3	213.1	104419.5	102.0	0.7	656.7	56.1	26.6	98.6	0.6	3.6	0.1
100	25	356.5	(8) 53.1	(8) 14228.8	62.7	1.5	3620.0	31.4	23.0	89.6	0.9	73.0	38.1
60	25	60.8	3.0	917.1	127.4	0.9	621.2	123.0	32.2	85.2	0.6	18.3	18.4
50	50	92.7	282.4	134246.3	15.1	1.4	1280.3	4.7	39.7	136.8	2.0	110.3	88.1
75	25	212.7	(9) 50.9	(9) 8258.3	137.5	3.3	7594.5	131.4	71.3	123.0	1.3	75.8	28.2
100	25	284.5	(8) 188.6	(8) 55411.8	61.2	3.2	5808.1	47.8	63.0	147.2	0.3	21.5	18.4
70	25	130.2	23.7	12065.8	37.9	3.4	2884.9	13.8	109.6	147.7	4.5	259.3	42.0
50	50	177.1	(6) 213.8	(6) 63859.7	71.7	8.4	13221.8	59.2	141.0	207.4	2.2	128.2	139.7
75	25	382.4	(8) 873.0	(8) 105465.6	56.1	16.2	33821.0	17.4	196.2	811.7	3.5	246.8	114.9
100	25	252.2	(4) 60.2	(4) 10867.5	59.6	14.6	25809.6	53.0	153.3	243.2	4.0	319.7	338.8
80	25	111.2	226.6	89013.4	33.5	7.8	6115.3	13.0	149.6	195.2	7.5	390.9	86.1
50	50	271.6	(8) 872.9	(8) 181325.9	55.0	26.8	36346.3	20.9	373.9	369.2	8.6	544.8	213.6
75	25	313.3	(5) 278.7	(5) 14838.5	82.0	47.8	96543.3	70.8	61.1	745.2	2.6	413.4	359.0
100	25	473.0	(6) 1469.5	(6) 98024.5	43.0	99.5	216700.0	17.1	717.5	804.6	5.4	1849.4	1219.7
90	25	118.5	(9) 585.9	(9) 630635.0	111.5	23.3	22656.9	107.3	188.6	390.4	3.6	430.6	94.1
50	50	248.6	(6) 3708.5	(6) 312105.5	82.2	67.8	99574.6	73.3	532.3	810.0	3.4	729.1	404.9
75	25	388.7	(2) 2850.5	(2) 62190.5	37.9	735.1	1348558.3	14.2	1281.2	970.8	8.7	(7) 3234.1	(7) 2233.0
100	25	390.0	(3) 146.2	(3) 5947.5	26.6	180.4	282996.1	19.4	1094.7	5644.1	6.5	2749.9	1357.9
100	25	169.4	2308.1	623731.5	74.4	65.4	71448.9	61.6	392.5	617.4	10.9	1583.1	284.1
50	50	145.7	(6) 1724.3	(6) 122716.0	17.5	308.0	465749.5	7.6	986.9	882.0	8.0	(9) 3379.6	(9) 1488.6
75	25	270.9	(2) 4243.5	(2) 88176.5	21.8	856.8	1322390.0	6.8	980.0	967.8	14.8	(7) 2613.0	(7) 855.6
100	25	473.0	(5) 2658.8	(5) 120959.0	98.6	649.7	977246.9	94.0	(9) 723.8	(9) 3166.7	6.0	(7) 318.1	(7) 115.4
110	25	124.0	(6) 277.4	(6) 36270.2	72.2	327.0	288129.4	64.4	848.8	1003.6	13.5	(8) 2602.9	(8) 810.4
50	50	117.5	(3) 661.5	(3) 55327.0	14.3	1188.1	1089556.0	4.7	1010.2	727.8	5.7	(7) 2065.1	(7) 652.3
75	25	580.7	(6) 908.8	(6) 35891.8	138.7	(7) 27.4	(7) 37498.8	118.8	(8) 2305.7	(8) 4523.3	10.7	(7) 1062.7	(7) 297.7
100	25	332.2	(1) 1911.6	(1) 118532.0	19.6	(8) 758.0	(8) 295886.3	7.1	(5) 1438.1	(5) 1375.0	7.0	(6) 1789.2	(6) 511.2
120	25	55.1	(6) 320.1	(6) 94936.5	95.3	1771.4	1644176.9	111.8	424.3	447.6	5.7	(8) 1872.3	(8) 317.5
50	50	288.1	(3) 2995.9	(3) 81429.7	90.3	(7) 1888.9	(7) 1554792.4	82.5	(8) 1073.9	(8) 821.1	10.2	(6) 1725.3	(6) 427.5
75	25	507.6	(5) 305.0	(5) 11101.2	133.8	(6) 484.9	(6) 177043.8	126.7	(9) 3001.6	(9) 7996.3	7.9	(5) 10.8	(5) 10.0
100	25	179.7	(3) 411.9	(3) 41666.5	66.2	(6) 61.6	(6) 36073.9	68.7	(9) 1522.6	(9) 962.0	3.6	(6) 683.4	(6) 145.4
130	25	129.1	(6) 2014.5	(6) 383386.2	24.7	1256.6	698989.0	10.6	3341.8	2520.8	7.1	(8) 4194.0	(8) 850.7
50	50	411.8	(4) 3246.9	(4) 245787.0	67.3	(7) 493.0	(7) 384516.5	50.5	(8) 1719.6	(8) 2330.7	6.9	(7) 2590.8	(7) 2340.8
75	25	207.3	(6) 203.0	(6) 1075.3	12.2	(5) 4973.3	(5) 31338017.2	8.0	(9) 2630.7	(9) 1000.3	11.9	(2) 2813.5	(2) 4430.0
100	25	385.5	(6) 305.0	(6) 118532.0	21.1	(4) 2250.4	(4) 1170285.3	8.8	(6) 4965.0	(6) 2437.4	14.1	(3) 2012.0	(3) 533.5
140	25	207.9	(5) 15.0	(5) 1180.5	48.8	(8) 1770.8	(8) 654605.8	44.4	2625.2	44.4	13.3	(4) 2561.8	(4) 298.5
50	50	306.2	(1) 2401.8	(1) 106348.0	36.8	(5) 2692.3	(5) 2306379.0	16.4	(7) 4809.5	(7) 1134.1	24.0	(3) 3360.3	(3) 213.3
75	25	259.0	(6) 305.0	(6) 118532.0	19.6	(4) 2263.3	(4) 1520163.5	8.0	(5) 4095.2	(5) 1773.4	12.5	(1) 431.0	(1) 0.0
100	25	647.8	(2) 64.1	(2) 483.0	49.4	(4) 1042.9	(4) 1238929.3	37.1	(6) 6123.7	(6) 17745.3	13.3	(4) 2561.7	(4) 298.5
150	25	103.7	(4) 1744.3	(4) 202004.0	67.4	(6) 587.1	(6) 552495.8	69.1	(8) 3203.9	(8) 994.8	12.6	(3) 1458.0	(3) 164.3
50	50	105.6	(5) 91.8	(5) 5591.7	98.2	(7) 2240.0	(7) 905027.4	101.4	(7) 2761.5	(7) 2349.4	8.8	(5) 2797.7	(5) 3453.7
75	25	496.9	(6) 305.0	(6) 118532.0	7.0	(5) 997.2	(5) 3004045.4	23.1	(3) 3908.7	(3) 876.3	17.5	(2) 155.0	(2) 0.0
100	25	171.1	(1) 1039.3	(1) 24546.0	43.7	(3) 3493.0	(3) 5264462.0	3.3	(4) 4320.1	(4) 2171.0	8.1	(3) 5391.7	(3) 554.3
Avg		258.5	(236) 787.0	(236) 127524.4	59.0	(372) 570.0	(372) 902502.7	45.7	(394) 1215.4	(394) 1487.2	7.4	(328) 1250.5	(328) 433.8

Instances with up to $n = 100$ variables can be solved most efficiently by the MIQCR approach, i.e., finding a convexified problem via semidefinite optimization and then solve the resulting convex problem using Cplex.

For $n > 100$, BiqCrunch performs best in terms of overall run time, but the dominance to MIQCR and our approach is not significant.

Our new approach provides by far the smallest gap at the root node. The high quality of our bound is also reflected in the number of nodes in the branch-and-bound tree. Our method explores a substantial smaller number of nodes than the other approaches.

Our approach is not superior to MIQCR or BiqCrunch in terms of overall computation time, however, the implementation is a prototype and there is room for speeding up the approach by experimenting with different settings in the branch-and-bound framework (such as branching strategies) as well as parameter settings in the bundle algorithm and in the update of the set of triangle inequalities. This is currently under investigation.

6 Conclusion

The 0–1 k -item quadratic knapsack problem is a challenging problem, as it includes two NP-hard problems, namely quadratic knapsack and k -cluster. We review approaches to solve this problem to optimality and introduce a new method, where the bound computation is based on a semidefinite relaxation. The derived basic semidefinite relaxation has only simple constraints, in fact all constraints are of rank one. This can be exploited in interior point methods to efficiently compute the system matrix. We strengthen the relaxation using triangle inequalities and solve the resulting semidefinite problem by a dynamic version of the bundle method.

To have a comparison with state of the art algorithms we implement the convexification algorithm MIQCR [8], use BiqCrunch [15] enhanced with our primal heuristics, and run Cplex. The numerical results prove that CPLEX is clearly outperformed by all the methods based on semidefinite programming. Our new method provides the tightest bound at the root node, while the overall computation time is smallest for MIQCR for $n \leq 100$ and BiqCrunch for larger n . An optimized implementation and a study of the best parameter settings for the various components inside our code is subject of further study.

References

1. Anjos, M.F., Ghaddar, B., Hupp, L., Liers, F., Wiegele, A.: Solving k -way graph partitioning problems to optimality: the impact of semidefinite relaxations and the bundle method. In: Jünger, M., Reinelt, G. (eds.) Facets of combinatorial optimization, pp. 355–386. Springer, Heidelberg (2013)
2. Balas, E., Zemel, E.: An algorithm for large zero-one knapsack problems. *Oper. Res.* **28**, 1130–1154 (1980)
3. Bertsimas, D., Shioda, R.: Algorithm for cardinality-constrained quadratic optimization. *Comput. Optim. Appl.* **43**, 1–22 (2009)

4. Bhaskara, A., Charikar, M., Guruswami, V., Vijayaraghavan, A., Zhou, Y.: Polynomial integrality gaps for strong SDP relaxations of densest k -subgraph. In: Proceedings of the Twenty-Third Annual ACM-SIAM Symposium on Discrete Algorithms, pp. 388–405 (2012)
5. Bienstock, D.: Computational study of a family of mixed-integer quadratic programming problems. *Math. Programm.* **74**, 121–140 (1996)
6. Billionnet, A.: Different formulations for solving the heaviest k -subgraph problem. *Inf. Syst. Oper. Res.* **43**(3), 171–186 (2005)
7. Billionnet, A., Calmels, F.: Linear programming for the 0–1 quadratic knapsack problem. *Eur. J. Oper. Res.* **92**, 310–325 (1996)
8. Billionnet, A., Elloumi, S., Lambert, A.: Extending the QCR method to general mixed-integer programs. *Math. Programm.* **131**(1–2), 381–401 (2012)
9. Bonami, P., Lejeune, M.: An exact solution approach for portfolio optimization problems under stochastic and integer constraints. *Oper. Res.* **57**, 650–670 (2009)
10. Borchers, B.: CSDP, a C library for semidefinite programming. *Optim. Meth. Softw.* **11**(1), 613–623 (1999)
11. IBM ILOG CPLEX Callable Library version 12.6.2. <http://www-03.ibm.com/software/products/en/ibmilogcpleoptistud/>
12. Fischer, I., Gruber, G., Rendl, F., Sotirov, R.: Computational experience with a bundle approach for semidefinite cutting plane relaxations of Max-Cut, equipartition. *Math. Programm. Ser. B* **105**(2–3), 451–469 (2006)
13. Helmberg, C.: The conicbundle library for convex optimization, August 2015. <https://www-user.tu-chemnitz.de/~helmberg/ConicBundle/Manual/index.html>
14. Helmberg, C., Rendl, F., Vanderbei, R.J., Wolkowicz, H.: An interior-point method for semidefinite programming. *SIAM J. Optim.* **6**(2), 342–361 (1996)
15. Krislock, N., Malick, J., Roupin, F.: Improved semidefinite bounding procedure for solving max-cut problems to optimality. *Math. Program. Ser. A* **143**(1–2), 61–86 (2014)
16. Létocart, L., Plateau, M.-C., Plateau, G.: An efficient hybrid heuristic method for the 0–1 exact k -item quadratic knapsack problem. *Pesquisa Operacional* **34**(1), 49–72 (2014)
17. Mitra, G., Ellison, F., Scowcroft, A.: Quadratic programming for portfolio planning: insights into algorithmic and computational issues. *J. Asset Manage.* **8**, 249–258 (2007). Part ii: Processing of Portfolio Planning Models with Discrete Constraints
18. Pisinger, D.: The quadratic knapsack problem: a survey. *Discrete Appl. Math.* **155**, 623–648 (2007)
19. Rader Jr., D.J., Woeginger, G.J.: The quadratic 0–1 knapsack problem with series-parallel support. *Oper. Res. Lett.* **30**(3), 159–166 (2002)
20. Rendl, F., Rinaldi, G., Wiegele, A.: Solving max-cut to optimality by intersecting semidefinite, polyhedral relaxations. *Math. Program. Ser. A* **121**(2), 307–335 (2010)
21. Rendl, F., Sotirov, R.: Bounds for the quadratic assignment problem using the bundle method. *Math. Program. Ser. B* **109**(2–3), 505–524 (2007)
22. Shawa, D.X., Liub, S., Kopmanb, L.: Lagrangean relaxation procedure for cardinality-constrained portfolio optimization. *Optim. Meth. Softw.* **23**, 411–420 (2008)