

A Dynamic Programming Heuristic for the Quadratic Knapsack Problem

Franklin Djeumou Fomeni Adam N. Letchford*

March 2012

Abstract

It is well known that the standard (linear) knapsack problem can be solved exactly by dynamic programming in $\mathcal{O}(nc)$ time, where n is the number of items and c is the capacity of the knapsack. The *quadratic* knapsack problem, on the other hand, is \mathcal{NP} -hard in the strong sense, which makes it unlikely that it can be solved in pseudo-polynomial time. We show however that the dynamic programming approach to the linear knapsack problem can be modified to yield a highly effective constructive heuristic for the quadratic version. The lower bounds obtained by our heuristic are consistently within a fraction of a percent of optimal.

Keywords: knapsack problems, integer programming, dynamic programming.

1 Introduction

The *Quadratic Knapsack Problem* (QKP), introduced by Gallo *et al.* [7], is an optimisation problem of the following form:

$$\max \sum_{i=1}^n \sum_{j=1}^n q_{ij} x_i x_j \quad (1)$$

$$\text{s.t.} \quad \sum_{i=1}^n w_i x_i \leq c \quad (2)$$

$$x \in \{0, 1\}^n, \quad (3)$$

Here, n is the number of *items*, c is the (positive and integral) *capacity* of the knapsack, the w_i are the (positive and integral) *weights* of the items, and the q_{ij} are the (integral) *profits*. For $i = 1, \dots, n$, the binary variable x_i takes the value 1 if and only if the i th item is inserted into the knapsack.

*Department of Management Science, Lancaster University, Lancaster LA1 4YX, United Kingdom. E-mail: A.N.Letchford@lancaster.ac.uk

Note that, since $x_i^2 = x_i$ for all i , one can easily incorporate linear terms by adjusting the q_{ii} . Indeed, the QKP reduces to the standard (linear) knapsack problem (KP) when $q_{ij} = 0$ for all $i \neq j$. In terms of computational complexity, however, the QKP is more complex than the KP. Indeed, although the KP is \mathcal{NP} -hard, as shown by Karp [8], it is possible to solve it in $\mathcal{O}(nc)$ time using dynamic programming (DP), as shown by Bellman [2]. The QKP, on the other hand, is known to be \mathcal{NP} -hard in the *strong sense* (see, e.g., [5]), which makes it unlikely that a pseudo-polynomial time algorithm exists. In fact, there is evidence that, even in the special case in which $w_i = 1$ for all i , one cannot even *approximate* the optimal profit within a constant factor in polynomial time (e.g., Khot [10]). Even for this special case, the best approximation factor obtained to date is $\mathcal{O}(n^{1/4})$ (Bhaskara *et al.* [3]).

The QKP is a well-studied combinatorial optimisation problem, with a variety of important applications, for example in the location of satellites, airports, railway stations or freight terminals. A variety of bounding procedures, heuristics and exact algorithms are available for it. For details, we refer the reader to Chapter 12 of the book by Kellerer *et al.* [9], and the more recent survey by Pisinger [13]. One of the most effective exact algorithms for the QKP is that of Caprara *et al.* [5], which is based on Lagrangian relaxation, subgradient optimisation, and branch-and-bound. Recently, this algorithm was made even more effective by Pisinger *et al.* [14], using several powerful reduction techniques.

The purpose of this paper is to show that the DP approach to the KP can be modified to yield an effective *heuristic* for the QKP. Essentially, this is done by using the same stages and states as in the standard DP recursion for the KP, but using a quadratic objective in place of a linear objective. The reason that it is a heuristic, rather than an exact algorithm, is that the Bellman [2] principle of optimality does not hold in the quadratic case. This is illustrated later in Section 3. Nevertheless, our computational results demonstrate that the heuristic performs very well. In particular, if the items are ordered appropriately, and one breaks ties in an intelligent way, one can consistently obtain lower bounds that are within 0.05% of optimality.

The paper is structured as follows. In Section 2, we review the relevant literature on the KP and QKP. In Section 3, the new heuristic is described, and it is shown how to implement it to use $\mathcal{O}(n^2c)$ time and $\mathcal{O}(nc)$ memory space. In Section 4, we present the above-mentioned ordering and tie-breaking rules, and show that they considerably enhance the performance of the heuristic. In Section 5, we present more extensive computational results, obtained using the enhanced version of our heuristic. Finally, some concluding remarks are presented in Section 6.

2 Literature Review

Since the literature on knapsack problems is vast, we do not attempt to cover it all here. Instead, we refer the reader to the books [9, 11] and the survey [13]. Nevertheless, there are certain key concepts from the literature which are vital to what follows and are therefore covered in the following four subsections.

2.1 The continuous relaxation of the KP

The continuous relaxation of the KP is defined as follows:

$$\max \left\{ \sum_{i=1}^n q_{ii} x_i : w^T x \leq c, x \in [0, 1]^n \right\}.$$

Dantzig [6] showed that this relaxation can be solved quickly, in $\mathcal{O}(n \log n)$ time, as follows:

1. Temporarily set x_i to 0 for all i .
2. Sort the items in non-increasing order of *profit-to-weight ratio* q_{ii}/w_i .
3. Iteratively set variables to 1 in the given order, as long as this can be done while maintaining feasibility.
4. Set the next variable to the largest possible (typically fractional) value that it can take, while maintaining feasibility.

In Balas & Zemel [1] it is shown that, in fact, no sorting is necessary, and one can solve the continuous relaxation of the KP in $\mathcal{O}(n)$ time, using advanced median-finding techniques.

2.2 The classical dynamic programming approach to the KP

Next, we recall the classical DP approach to the KP. For any $k \in \{1, \dots, n\}$ and $r \in \{0, \dots, c\}$, let $f(k, r)$ be the maximum profit obtainable by packing a selection of the first k items whose total weight is equal to r . That is, let:

$$f(k, r) = \max \left\{ \sum_{i=1}^k p_i x_i : \sum_{i=1}^k w_i x_i = r, x_i \in \{0, 1\} (i = 1, \dots, k) \right\},$$

where $p_i = q_{ii}$.

If no such packing exists, then let $f(k, r) = -\infty$. Also, by convention, let $f(0, 0) = 0$. Now, observe that:

$$f(k, r) = \begin{cases} \max \{f(k-1, r), f(k-1, r-w_k) + p_k\} & \text{if } r \geq w_k, \\ f(k-1, r) & \text{otherwise.} \end{cases} \quad (4)$$

Algorithm 1 : The classical DP algorithm

Initialise $f(0,0)$ to 0 and $f(k,r) = -\infty$ for all other k,r .
for $k = 1, \dots, n$ **do**
 for $r = 0, \dots, c$ **do**
 if $f(k-1,r) > f(k,r)$ **then**
 Set $f(k,r) := f(k-1,r)$
 end if
 if $r + w_k \leq c$ and $f(k-1,r) + p_k > f(k,r + w_k)$ **then**
 Set $f(k,r + w_k) := f(k-1,r) + p_k$.
 end if
 end for
end for
Output $\max_{0 \leq r \leq c} f(n,r)$.

This is a classic dynamic programming recursive function, which indicates that the KP obeys the so-called ‘principle of optimality’ of Bellman [2]. It enables us to compute the $f(k,r)$ using Algorithm 1.

This algorithm clearly runs in $\mathcal{O}(nc)$ time. Note that it outputs only the optimal profit. If you want the optimal solution itself, you have to ‘trace back’ the path from the optimal state to the initial state $f(0,0)$. The time taken to do this is negligible.

2.3 Upper planes for the QKP

Now we move on to the QKP itself.

Gallo *et al.* [7] say that a vector $\pi \in \mathbb{Q}_+^n$ is an *upper plane* if the following holds:

$$\pi^T x \geq \sum_{i=1}^n \sum_{j=1}^n q_{ij} x_i x_j \quad (\forall x \in \{0,1\}^n).$$

Given an upper plane π , one can compute an upper bound for the QKP by solving the following (linear) KP:

$$\max \{ \pi^T x : w^T x \leq c, x \in \{0,1\}^n \}. \quad (5)$$

Moreover, the solution of this KP yields a feasible vector x , and therefore a lower bound for the QKP.

Gallo *et al.* [7] proposed four different upper planes, π^1, \dots, π^4 . These

are defined by setting, for all i :

$$\begin{aligned}\pi_i^1 &= \sum_{j=1}^n q_{ij}, \\ \pi_i^2 &= q_{ii} + \max \left\{ \sum_{j \neq i} q_{ij} x_j : \sum_{j=1}^n x_j \leq m, x \in \{0, 1\}^n \right\}, \\ \pi_i^3 &= q_{ii} + \max \left\{ \sum_{j \neq i} q_{ij} x_j : w^T x \leq c, x \in [0, 1]^n \right\}, \\ \pi_i^4 &= q_{ii} + \max \left\{ \sum_{j \neq i} q_{ij} x_j : w^T x \leq c, x \in \{0, 1\}^n \right\}.\end{aligned}$$

The quantity m appearing in the definition of π_i^2 is the maximum number of items that can be fitted into the knapsack. This can be computed by sorting the items in non-decreasing order of their weights, and then inserting the items into the knapsack in the given order, until no more can be inserted.

Clearly, for any i , we have $\pi_i^1 \geq \pi_i^2 \geq \pi_i^4$ and $\pi_i^1 \geq \pi_i^3 \geq \pi_i^4$. On the other hand, for any i , the quantities π_i^1 and π_i^2 can be computed easily in $\mathcal{O}(n)$ and $\mathcal{O}(n \log n)$ time, respectively, whereas computing π_i^4 amounts to solving a (linear) KP, which could be time consuming. As for π_i^3 , it is possible to compute it in $\mathcal{O}(n)$ time as well, using the median-finding technique of Balas & Zemel [1], mentioned in Subsection 2.1.

According to Gallo *et al.* [7], using π_i^3 gives the best trade-off between bound strength and computing time.

2.4 Improved upper planes for the QKP

Caprara *et al.* [5] noted that one can slightly improve the upper planes π_i^2 , π_i^3 and π_i^4 by replacing them with the following:

$$\begin{aligned}\tilde{\pi}_i^2 &= q_{ii} + \max \left\{ \sum_{j \neq i} q_{ij} x_j : \sum_{j \neq i} w_j x_j \leq m - 1, x_j \in \{0, 1\} (j \neq i) \right\}, \\ \tilde{\pi}_i^3 &= q_{ii} + \max \left\{ \sum_{j \neq i} q_{ij} x_j : \sum_{j \neq i} w_j x_j \leq c - w_i, 0 \leq x_j \leq 1 (j \neq i) \right\}, \\ \tilde{\pi}_i^4 &= q_{ii} + \max \left\{ \sum_{j \neq i} q_{ij} x_j : \sum_{j \neq i} w_j x_j \leq c - w_i, x_j \in \{0, 1\} (j \neq i) \right\}.\end{aligned}$$

They also show that one can obtain still stronger bounds using either linear programming or Lagrangian relaxation. These upper bounds, along

with others based on linear programming, Lagrangian decomposition and semidefinite programming, are surveyed and compared in [9, 13].

3 Description of the DP heuristic

Consider what will happen if we attempt to adapt the classical dynamic programming algorithm (see Algorithm 1) to the QKP. We can define $f(k, r)$ in a natural way, as:

$$\max \left\{ \sum_{i=1}^k \sum_{j=1}^k q_{ij} x_i x_j : \sum_{i=1}^k w_i x_i = r, x_i \in \{0, 1\} (i = 1, \dots, k) \right\}.$$

The problem is that there is no analogue of the recursive equation (4). Indeed, the Bellman principle of optimality does not hold for the QKP. The following simple example makes this clear.

Example: Suppose that $n = 3$, $c = 2$ and $w_1 = w_2 = w_3 = 1$, and suppose that the profit matrix is:

$$Q = \begin{pmatrix} 10 & 0 & 0 \\ 0 & 1 & 10 \\ 0 & 10 & 1 \end{pmatrix}$$

We have $f(2, 1) = 10$, which correspond to inserting item 1 into the knapsack, and we have $f(2, 2) = 11$, which corresponds to inserting items 1 and 2. On the other hand, we have $f(3, 2) = 22$, which corresponds to inserting items 2 and 3. Note that item 1 is inserted in the first two cases, but not in the third case. There is therefore no recursive formula that can express $f(3, 2)$ in terms of $f(2, 1)$ and $f(2, 2)$. (This is of course to be expected, given that the QKP is \mathcal{NP} -hard in the strong sense.) \square

The main point of this paper, however, is to show that one can devise an effective *heuristic* using dynamic programming. In order to do this, it is helpful to redefine $f(k, r)$ as the *profit of the best packing found by the heuristic* that uses a selection of the first k items and whose total weight is equal to r . It is also helpful to define $S(k, r)$ as the set of items (viewed as a subset of $\{1, \dots, k\}$) that gives the profit $f(k, r)$. Our heuristic then performs as described in Algorithm 2 below.

Example (cont.): Applying our heuristic to the small example given above, we obtain the results shown in Table 1 below.

In this table, the values $f(k, r)$ and the sets $S(k, r)$ are the ones obtained at the *end* of the iteration, for the given k and r . The maximum of $f(3, r)$

Algorithm 2 :Dynamic programming for the QKP

Initialise $f(0,0)$ to 0, and $f(k,r)$ to $-\infty$ for all other k,r .

Initialise $S(k,r) = \emptyset$ for all k,r .

for $k = 1, \dots, n$ **do**

for $r = 0, \dots, c$ **do**

if $f(k-1,r) > f(k,r)$ **then**

 Set $f(k,r) := f(k-1,r)$ and $S(k,r) := S(k-1,r)$.

end if

if $r + w_k \leq c$ **then**

 Let β be the profit of $S(k-1,r) \cup \{k\}$.

if $\beta > f(k,r + w_k)$ **then**

 Set $f(k,r + w_k) := \beta$,

 Set $S(k,r + w_k) := S(k-1,r) \cup \{k\}$.

end if

end if

end for

end for

Let $r^* = \arg \max_{0 \leq r \leq c} f(n,r)$.

Output the set $S(n,r^*)$ and the associated profit $f(n,r^*)$.

Table 1: Dynamic programming computations for the small example.

	$k = 1$			$k = 2$			$k = 3$		
	β	$f(1,r)$	$S(1,r)$	β	$f(2,r)$	$S(2,r)$	β	$f(3,r)$	$S(3,r)$
$r = 0$	0	0	\emptyset	0	0	\emptyset	0	0	\emptyset
$r = 1$	10	10	$\{1\}$	1	10	$\{1\}$	1	10	$\{1\}$
$r = 2$	0	0	\emptyset	11	11	$\{1,2\}$	11	11	$\{1,3\}$

over all r is 11, so $r^* = 2$ and $S(3,r^*) = \{1,3\}$. The heuristic finds the solution $\{1,3\}$, with a profit of 11. On the other hand, the optimal solution is $\{2,3\}$, with a profit of 22. \square

This heuristic runs in $\mathcal{O}(n^2c)$ time, since it takes linear time to compute the difference between β and $f(k-1,r)$. A drawback is that it also takes $\mathcal{O}(n^2c)$ memory space, since for all k and r we are storing the set $S(k,r)$. Our preliminary computational experiments indicated that Algorithm 2 can only handle problem instances of moderate size due to this large memory requirement.

However, it is possible to reduce the memory requirement to $\mathcal{O}(nc)$, without any worsening of the running time, by exploiting the fact that all of the computations that take place in a given stage depend only on the values that were computed in the preceding stage. Effectively, this enables us to

drop the index k from the $f(k, r)$ array.

Here are the details. For $r \in \{0, \dots, c\}$, let $f(r)$ denote the *current* value of the most profitable packing found so far, regardless of the stage k , having a total weight of r . Also, for $r \in \{0, \dots, c\}$ and $i = 1, \dots, n$, let $B(r, i)$ be a Boolean variable taking the value 1 if item i is *currently* in the set of items whose profit is $f(r)$, and 0 otherwise. The heuristic can then be implemented as described in Algorithm 3 below.

Algorithm 3 : Dynamic programming for the QKP with reduced memory

Initialise $f(r)$ to 0 for $r = 0, \dots, c$.

Initialise $B(r, i)$ to 0 for all $r = 0, \dots, c$ and $i = 1, \dots, n$.

for $k = 1, \dots, n$ **do**

for $r = c$ to 0 (going down in steps of 1) **do**

if $r \geq w_k$ **then**

 Let β be the profit of $S(k-1, r-w_k) \cup \{k\}$

 (This can be computed in $\mathcal{O}(n)$ time using the $B(r-w_k, i)$

values.)

if $\beta > f(r)$ **then**

 Set $f(r) := \beta$,

 Set $S(k, r) = S(k-1, r-w_k) \cup \{k\}$

end if

end if

end for

end for

Let $r^* = \arg \max_{0 \leq r \leq c} f(r)$.

Compute the final set of items $S(n, r^*)$.

(This can be done in $\mathcal{O}(n)$ time using the $B(r^*, i)$ values.)

Output $S(n, r^*)$ and the associated profit $f(r^*)$.

The reason that we iteratively decrease r , instead of increasing it, is that the value $f(r)$ at a given stage k depends on the values of $f(r)$ and $f(r-w_k)$ from the previous stage, but does not depend on the value of $f(r')$ from the previous stage, for any $r' > r$.

4 Enhancements

To gain a feeling for the potential of our DP heuristic, we performed some experiments on randomly generated QKP instances. These instances all had $n = 50$, and were created following the following scheme, which is standard in the literature [4, 5, 7, 12, 14]. For a given *density* $\Delta\%$, each profit q_{ij} is zero with probability $(100 - \Delta)\%$, and uniformly distributed between 1 and 100 with probability $\Delta\%$. Each weight w_i is uniformly distributed between 1 and 50. The capacity c is uniformly distributed between 50 and $\sum_{i=1}^n w_i$.

For each instance, we computed the average percentage gap between the lower bound found by our heuristic and the optimal profit. The average gap over 100 instances was 0.26% for $\Delta = 100\%$ and 16.74% for $\Delta = 25\%$. This is very good for the dense instances, but disappointing for the sparse ones. In the following two subsections, two simple enhancements to our heuristic are presented that, in combination, drastically improve the quality of the lower bounds obtained, especially for the sparse instances.

4.1 Ordering the items

Consider again the small example with $n = 3$, presented in the previous section. Observe that, if we had re-ordered the items so that either item 2 or 3 came first, the DP heuristic would have found the optimal solution. So, it is natural to consider the devising of some intelligent way to order the items. Following the papers of Gallo *et al.* [7] and Caprara *et al.* [5] mentioned in Subsections 2.3 and 2.4, it is natural to consider sorting the items in non-increasing order of one of the values π_i^1, \dots, π_i^4 , or one of the values $\tilde{\pi}_i^2, \dots, \tilde{\pi}_i^4$. This gives seven values in total. Moreover, following the idea behind the greedy algorithm for the linear KP, mentioned in Subsection 2.1, it is also natural to consider sorting the items in non-increasing order of any of these seven values divided by the weight w_i . This gives fourteen different possible ordering rules in total.

Accordingly, for each of the 200 random instances mentioned above, we computed 14 different upper bounds. Moreover, for comparison purposes, we also computed the corresponding 7 upper bounds and 7 lower bounds obtained from the upper-planes approach (i.e., by solving the knapsack problem (5)).

Tables 2 and 3 display the results obtained. Table 2 is concerned with dense instances and Table 3 with sparse ones. The rows labelled ‘UB UP’ and ‘LB UP’ give the average percentage gaps for the upper and lower bounds obtained from the upper-planes approach. The row labelled ‘LB DP’ gives the gaps for the lower bounds obtained from our DP heuristic, when the various different π values are used to sort the items. Finally, the row labelled ‘LB DP/W’ gives the same, when the π values divided by the weights w_i are used to sort the items.

It is clear that the upper bounds from the upper plane approach are rather poor in all cases. The corresponding lower bounds are surprisingly good, especially for dense instances. For our heuristic, the lower bounds for dense instances are very good, especially when the π values are divided by the weights. On the other hand, the various orderings do not help at all for sparse instances. Moreover, bizarrely, dividing the π values by the weights seems to make things worse. This anomalous behaviour is rectified by the second enhancement, described next.

	π_i^1	$\tilde{\pi}_i^2$	π_i^2	$\tilde{\pi}_i^3$	π_i^3	$\tilde{\pi}_i^4$	π_i^4
UB UP	226.24	34.01	36.50	13.54	15.24	13.13	14.84
LB UP	0.87	0.97	1.02	0.29	0.36	0.32	0.35
LB DP	0.19	0.25	0.25	0.17	0.29	0.17	0.35
LB DP/W	0.05	0.06	0.06	0.05	0.05	0.04	0.04

Table 2: Average percentage gaps for various bounds when $\Delta = 100\%$.

	π_i^1	$\tilde{\pi}_i^2$	π_i^2	$\tilde{\pi}_i^3$	π_i^3	$\tilde{\pi}_i^4$	π_i^4
UB UP	203.23	52.70	52.84	40.61	40.75	39.70	40.86
LB UP	4.06	4.01	4.01	3.39	3.42	3.33	3.43
LB DP	16.47	14.63	14.63	13.79	14.12	13.56	13.90
LB DP/W	25.54	25.13	25.13	24.42	24.61	24.38	24.51

Table 3: Average percentage gaps for various bounds when $\Delta = 25\%$.

4.2 Breaking ties intelligently

When inspecting the solutions obtained by the various algorithms on the sparse instances, we noticed an interesting phenomenon. The optimal solution to such instances almost always inserted more items into the knapsack than the DP heuristic did. This suggests adjusting the DP heuristic slightly, so that it gives a preference for (partial) solutions with more items. More precisely, suppose that, for a given k and r in Algorithm 3, we find that $\beta = f(r)$. Then, if $|S(k-1, r-w_k)| + 1 > S(k, r)$, we set $f(r) := \beta$ and set $S(k, r) = S(k-1, r-w_k) \cup \{k\}$.

We call this enhancement the ‘*tie-breaking*’ rule. When this rule is used on its own, without any re-ordering of the items, the average percentage gap over the same instances mentioned previously was 0.26% for $\Delta = 100\%$ and 3.93% for $\Delta = 25\%$. This represents a substantial improvement for the sparse instances, but has little or no effect on the dense instances.

The really interesting results, however, are obtained when re-ordering and tie-breaking are used in combination. Table 4 shows the results obtained when using this option. By comparing Table 4 with Tables 2 and 3, it is apparent that the addition of tie-breaking improves the performance dramatically for the sparse instances, without having any effect on the performance for the dense instances.

All things considered, the best option seems to be to order the items in non-increasing order of π_i^3/w_i , $\tilde{\pi}_i^3/w_i$, π_i^4/w_i or $\tilde{\pi}_i^4/w_i$, and to use the tie-breaking rule. This gives an average percentage error of 0.05 or less for the dense instances, and 0.30 or less for the sparse instances.

		π_i^1	$\tilde{\pi}_i^2$	π_i^2	$\tilde{\pi}_i^3$	π_i^3	$\tilde{\pi}_i^4$	π_i^4
LB DP	($\Delta = 100\%$)	0.19	0.25	0.25	0.17	0.29	0.17	0.35
LB DP/W	($\Delta = 100\%$)	0.05	0.06	0.06	0.05	0.05	0.04	0.04
LB DP	($\Delta = 25\%$)	1.50	1.61	1.61	1.42	1.72	1.62	1.63
LB DP/W	($\Delta = 25\%$)	0.37	0.41	0.41	0.30	0.30	0.30	0.30

Table 4: Average percentage gaps obtained when using the tie-breaking rule.

5 Additional Computational Results

In this section, we present more extensive computational results. The DP heuristic, the bounding procedures based on upper planes, and the sorting rules mentioned in Subsection 4.1 were all coded in the C language and compiled with *gcc 4.4.1 of Code::Blocks 10.05*, and run on a DELL desktop, with a 3.10 GHz processor and 4 GB of RAM, under the Windows 7 operating system.

The instances were generated using the scheme described at the start of the previous section, but this time we generated 20 instances for $n \in \{20, 40, \dots, 380\}$ and for $\Delta\% \in \{25\%, 50\%, 75\%, 100\%\}$. Also, for the sake of brevity, we consider only the four promising configurations mentioned in Subsection 4.2, i.e., using the tie-breaking rule and ordering the items in non-increasing order of π_i^3/w_i , $\tilde{\pi}_i^3/w_i$, π_i^4/w_i or $\tilde{\pi}_i^4/w_i$.

The results for the four different configurations are shown in Tables 5 to 8. It will be seen that, for each value of $\Delta\%$, the range of n varies. This is because we had to solve each instance to proven optimality in order to compute the percentage gaps, but this was possible only for certain values of Δ and n . (To solve the instances to optimality, we used the code of Caprara *et al.* [5], which was kindly given to us by Alberto Caprara.)

For each chosen pair of Δ and n , the tables give the average percentage gap from the optimum, the standard deviation of this gap, and the time taken to compute the bound (in seconds), for both the DP heuristic and the upper planes heuristic. Also, the last column displays what we call the ‘performance’ of the DP heuristic, which is defined as:

$$\frac{\text{Av. Gap from UP} - \text{Av. Gap from DP}}{\text{Av. Gap from UP}} \times 100\%$$

These results clearly demonstrate the excellent quality of the DP heuristic, since the average gap is a fraction of one percent in all cases. They also demonstrate the superiority of the DP heuristic over the upper planes heuristic, since the ‘performance’ is almost always well above 80%. Moreover, the running times are quite reasonable, despite the fact that the heuristic runs in pseudo-polynomial time, rather than polynomial time.

Table 5: Average percentage gaps when items are sorted according to π_i^3/w_i

Inst.		DP Heuristic			UP			Imp.
$\Delta(\%)$	n	Av. Gap	std	Av. Time	Av. Gap	std	Av. Time	perf (%)
100	20	0.002	0.01	0.003	0.30	0.61	0.000	71.20
	40	0.000	0.000	0.007	1.05	1.40	0.000	100.00
	60	0.01	0.03	0.047	0.08	0.13	0.001	78.36
	80	0.03	0.09	0.079	0.08	0.11	0.002	62.33
	100	0.01	0.04	0.176	0.10	0.19	0.005	86.60
	120	0.02	0.05	0.244	0.12	0.18	0.010	83.26
	140	0.04	0.12	0.120	0.15	0.36	0.011	68.26
	160	0.002	0.007	0.805	0.10	0.19	0.016	97.82
	180	0.005	0.01	0.521	0.06	0.07	0.020	91.67
	200	0.002	0.006	0.862	0.05	0.04	0.027	94.66
	220	0.006	0.01	1.090	0.03	0.05	0.035	82.79
	240	0.01	0.03	1.910	0.07	0.13	0.044	74.87
	260	0.04	0.15	1.440	0.09	0.18	0.055	50.73
	280	0.009	0.01	2.793	0.03	0.04	0.069	77.32
	300	0.005	0.01	6.142	0.08	0.14	0.140	93.99
	320	0.006	0.009	4.040	0.05	0.04	0.100	87.85
340	0.01	0.01	1.268	0.10	0.04	0.114	89.35	
360	0.003	0.005	5.647	0.04	0.04	0.140	93.10	
380	0.004	0.003	6.673	0.009	0.003	0.163	54.38	
75	20	0.21	0.42	0.002	0.63	1.21	0.000	66.44
	40	0.03	0.07	0.009	0.32	0.57	0.000	88.70
	60	0.17	0.62	0.019	0.86	1.14	0.001	79.31
	80	0.01	0.03	0.045	0.36	0.38	0.002	95.21
	100	0.01	0.03	0.076	0.30	0.49	0.004	94.29
	120	0.04	0.12	0.135	0.40	0.55	0.006	89.22
	140	0.01	0.02	0.270	0.27	0.53	0.010	96.33
	160	0.02	0.05	0.425	0.21	0.26	0.014	89.22
	180	0.01	0.02	0.510	0.18	0.21	0.020	93.56
200	0.03	0.05	0.700	0.23	0.28	0.026	86.99	
50	20	0.49	1.60	0.002	2.18	4.10	0.000	77.29
	40	0.16	0.36	0.008	1.08	1.77	0.000	84.49
	60	0.09	0.19	0.021	1.17	1.97	0.001	92.03
	80	0.05	0.16	0.046	0.73	1.17	0.002	92.52
	100	0.01	0.02	0.071	0.89	0.81	0.004	98.83
	120	0.06	0.13	0.149	0.74	1.17	0.006	90.90
	140	0.06	0.16	0.203	0.46	0.64	0.009	85.66
	160	0.02	0.07	0.387	0.40	0.27	0.014	92.59
25	20	0.15	0.47	0.002	3.98	5.60	0.000	96.15
	40	0.55	1.51	0.008	4.04	5.03	0.000	86.31
	60	0.24	0.54	0.018	2.91	3.28	0.001	91.68
	80	0.28	0.39	0.043	2.14	2.29	0.002	86.52
	100	0.29	0.39	0.074	2.71	2.81	0.003	89.24
	120	0.19	0.28	0.123	2.12	2.50	0.006	91.02
	140	0.14	0.21	0.231	1.78	1.70	0.009	92.07

Table 6: Average percentage gaps when items are sorted according to $\tilde{\pi}_i^3/w_i$

Inst.		DP Heuristic			UP			Imp.
$\Delta(\%)$	n	Av. Gap	std	Av. Time	Av. Gap	std	Av. Time	perf (%)
100	20	0.08	0.22	0.001	0.59	0.99	0.000	85.29
	40	0.002	0.01	0.008	0.53	0.69	0.000	99.52
	60	0.01	0.03	0.047	0.08	0.13	0.001	78.36
	80	0.03	0.09	0.080	0.11	0.18	0.002	73.17
	100	0.01	0.04	0.176	0.10	0.19	0.005	86.60
	120	0.02	0.05	0.252	0.12	0.18	0.007	83.60
	140	0.04	0.012	0.464	0.11	0.30	0.011	56.96
	160	0.002	0.007	0.803	0.10	0.19	0.017	97.82
	180	0.005	0.01	0.510	0.06	0.07	0.020	91.67
	200	0.001	0.006	0.899	0.05	0.04	0.027	94.74
	220	0.005	0.007	1.030	0.03	0.05	0.035	85.87
	240	0.02	0.03	1.500	0.08	0.14	0.045	74.50
	260	0.04	0.15	1.533	0.09	0.18	0.055	50.73
	280	0.009	0.01	2.607	0.03	0.04	0.069	77.32
	300	0.005	0.01	6.094	0.08	0.14	0.093	93.99
	320	0.006	0.01	3.364	0.06	0.06	0.099	88.95
340	0.01	0.01	1.355	0.10	0.04	0.113	89.35	
360	0.004	0.007	5.487	0.04	0.03	0.140	89.93	
380	0.004	0.003	6.316	0.009	0.006	0.163	54.38	
75	20	0.21	0.42	0.002	0.63	1.21	0.000	66.44
	40	0.03	0.07	0.008	0.25	0.47	0.000	85.42
	60	0.18	0.62	0.021	0.79	1.03	0.001	77.40
	80	0.01	0.03	0.045	0.36	0.38	0.002	95.49
	100	0.07	0.26	0.076	0.34	0.56	0.004	74.14
	120	0.04	0.12	0.132	0.40	0.55	0.006	89.54
	140	0.01	0.02	0.275	0.27	0.53	0.010	96.33
	160	0.02	0.05	0.420	0.21	0.28	0.015	89.44
	180	0.01	0.02	0.504	0.18	0.21	0.019	92.21
200	0.03	0.05	0.640	0.23	0.28	0.026	86.93	
50	20	0.55	1.57	0.002	1.68	3.36	0.000	67.20
	40	0.09	0.18	0.008	1.20	1.90	0.000	92.34
	60	0.09	0.19	0.021	1.39	2.05	0.001	93.31
	80	0.05	0.16	0.046	0.73	1.17	0.002	92.49
	100	0.01	0.02	0.072	1.04	1.33	0.003	99.99
	120	0.06	0.13	0.149	0.72	1.16	0.006	91.65
	140	0.06	0.16	0.200	0.46	0.64	0.009	85.66
160	0.02	0.07	0.380	0.38	0.27	0.014	92.36	
25	20	0.15	0.47	0.002	3.54	4.47	0.000	95.68
	40	0.55	1.51	0.008	3.86	5.11	0.000	85.65
	60	0.21	0.53	0.017	2.81	3.21	0.001	92.32
	80	0.28	0.39	0.045	2.69	4.13	0.002	89.28
	100	0.29	0.39	0.074	2.74	2.81	0.003	89.32
	120	0.19	0.27	0.126	2.00	2.48	0.006	90.41
	140	0.14	0.21	0.227	1.78	1.70	0.009	92.07

Table 7: Average percentage gaps when items are sorted according to π_i^4/w_i

Inst.		DP Heuristic			UP			Imp.
$\Delta(\%)$	n	Av. Gap	std	Av. Time	Av. Gap	std	Av. Time	perf (%)
100	20	0.08	0.22	0.004	0.43	0.81	0.000	79.76
	40	0.000	0.000	0.012	0.61	0.79	0.005	100.00
	60	0.01	0.04	0.093	0.07	0.12	0.048	76.22
	80	0.03	0.09	0.159	0.08	0.12	0.082	64.53
	100	0.01	0.04	0.351	0.10	0.19	0.179	86.60
	120	0.02	0.05	0.488	0.12	0.18	0.253	82.14
	140	0.04	0.12	0.910	0.11	0.30	0.445	56.25
	160	0.002	0.007	1.568	0.10	0.19	0.742	97.82
	180	0.005	0.07	0.996	0.07	0.49	0.496	92.03
	200	0.002	0.006	1.666	0.05	0.04	0.810	95.66
	220	0.006	0.01	2.000	0.03	0.05	0.953	82.41
	240	0.01	0.03	3.160	0.08	0.14	1.297	78.84
	260	0.04	0.15	2.775	0.09	0.18	1.298	50.35
	280	0.009	0.01	4.822	0.03	0.04	1.988	77.32
	300	0.005	0.01	10.96	0.08	0.14	6.065	93.99
	320	0.005	0.009	6.498	0.05	0.04	2.475	90.07
340	0.01	0.01	2.320	0.10	0.04	1.216	88.51	
360	0.003	0.005	9.199	0.04	0.04	3.549	93.03	
380	0.004	0.003	10.71	0.009	0.003	4.191	54.38	
75	20	0.21	0.42	0.003	0.63	1.21	0.000	66.44
	40	0.03	0.07	0.014	0.33	0.60	0.006	89.10
	60	0.15	0.48	0.040	0.83	1.14	0.019	80.86
	80	0.02	0.04	0.087	0.36	0.38	0.045	94.36
	100	0.07	0.26	0.150	0.32	0.51	0.077	76.83
	120	0.03	0.12	0.259	0.34	0.41	0.132	88.67
	140	0.01	0.02	0.538	0.27	0.53	0.271	96.21
	160	0.01	0.05	0.830	0.21	0.26	0.411	91.04
	180	0.01	0.02	1.004	0.18	0.21	0.511	93.56
200	0.02	0.05	1.296	0.23	0.28	0.610	90.45	
50	20	0.48	1.56	0.003	2.37	4.08	0.001	79.48
	40	0.16	0.36	0.015	0.89	1.50	0.007	81.15
	60	0.11	0.20	0.040	1.17	1.97	0.020	89.84
	80	0.05	0.16	0.089	0.73	1.17	0.046	92.48
	100	0.01	0.03	0.144	1.04	1.33	0.073	98.19
	120	0.06	0.13	0.287	0.74	1.17	0.146	90.90
	140	0.06	0.16	0.167	0.46	0.64	0.199	85.45
	160	0.02	0.07	0.766	0.40	0.27	0.375	92.95
25	20	0.05	0.25	0.003	3.98	5.60	0.000	98.54
	40	0.55	1.52	0.013	4.04	5.03	0.006	86.17
	60	0.22	0.53	0.032	2.91	3.28	0.016	91.26
	80	0.28	0.39	0.083	2.14	2.29	0.042	86.69
	100	0.31	0.40	0.074	2.43	2.32	0.071	86.88
	120	0.19	0.28	0.240	2.04	2.22	0.121	90.53
	140	0.14	0.21	0.457	1.78	1.70	0.230	92.07

Table 8: Average percentage gaps when items are sorted according to $\tilde{\pi}_i^4/w_i$

Inst.		DP Heuristic			UP			Imp.
$\Delta(\%)$	n	Av. Gap	std	Av. Time	Av. Gap	std	Av. Time	perf (%)
100	20	0.11	0.26	0.005	1.00	1.24	0.001	89.00
	40	0.000	0.000	0.010	0.46	0.49	0.005	100.00
	60	0.01	0.04	0.093	0.03	0.09	0.047	66.66
	80	0.04	0.09	0.156	0.07	0.09	0.080	42.85
	100	0.02	0.05	0.217	0.09	0.08	0.106	77.77
	120	0.01	0.01	0.496	0.07	0.05	0.256	85.71
	140	0.006	0.009	0.896	0.02	0.03	0.436	70.00
	160	0.001	0.003	1.502	0.17	0.25	0.738	99.41
	180	0.004	0.007	0.965	0.08	0.09	0.483	95.00
	200	0.001	0.003	1.687	0.04	0.03	0.793	97.50
	220	0.008	0.01	1.907	0.04	0.06	0.932	80.00
	240	0.02	0.03	2.879	0.11	0.19	1.273	81.81
	260	0.02	0.20	2.772	0.13	0.24	1.268	84.61
	280	0.008	0.01	4.907	0.05	0.05	1.964	84.00
	300	0.001	0.02	12.05	0.04	0.03	5.334	97.50
	320	0.007	0.01	4.245	0.04	0.06	1.881	82.50
340	0.01	0.01	2.302	0.11	0.04	1.182	90.90	
360	0.004	0.007	7.095	0.04	0.04	3.121	90.00	
380	0.004	0.003	9.982	0.09	0.06	3.953	93.03	
75	20	0.29	0.53	0.005	1.04	1.65	0.002	72.11
	40	0.04	0.07	0.014	0.20	0.28	0.006	80.00
	60	0.06	0.09	0.039	0.60	0.90	0.019	90.00
	80	0.02	0.05	0.087	0.29	0.31	0.044	93.10
	100	0.02	0.04	0.149	0.21	0.25	0.075	90.47
	120	0.008	0.01	0.255	0.41	0.52	0.129	98.04
	140	0.01	0.02	0.523	0.34	0.62	0.264	97.05
	160	0.01	0.02	0.826	0.15	0.13	0.405	93.33
	180	0.01	0.02	1.020	0.09	0.12	0.499	88.88
200	0.03	0.05	1.241	0.17	0.16	0.598	82.35	
50	20	0.29	0.41	0.003	2.48	3.93	0.001	88.30
	40	0.07	0.16	0.015	0.84	1.33	0.007	91.60
	60	0.09	0.21	0.040	2.79	0.68	0.020	96.77
	80	0.01	0.03	0.089	1.12	1.42	0.045	99.10
	100	0.02	0.03	0.139	1.51	0.70	0.071	98.67
	120	0.04	0.07	0.285	0.77	1.46	0.144	94.80
	140	0.01	0.02	0.381	0.24	0.16	0.194	95.83
	160	0.007	0.01	0.729	0.35	0.28	0.370	98.00
25	20	0.11	2.26	0.003	3.73	4.69	0.000	97.05
	40	0.99	2.03	0.003	3.81	4.62	0.000	74.01
	60	0.34	0.71	0.032	3.17	3.97	0.000	89.27
	80	0.34	0.40	0.083	2.89	2.88	0.040	88.23
	100	0.22	0.36	0.140	3.01	3.41	0.069	92.69
	120	0.29	0.30	0.236	2.68	2.60	0.118	89.17
	140	0.14	0.21	0.444	1.78	1.70	0.224	92.13

We remark that, whereas the exact algorithm could not be applied to larger instances without running into time and memory difficulties, the DP heuristic can be applied to much larger instances without any difficulty at all, provided that the knapsack capacity c is not excessively large.

6 Concluding Remarks

We have shown that the classical dynamic programming approach to the standard (linear) Knapsack Problem can be modified to yield an effective heuristic for the quadratic version. Specifically, our computational results demonstrate that, provided the items are ordered correctly and the tie-breaking rule is incorporated, one can consistently compute lower bounds that are within 0.05% of optimal. We are not aware of any other constructive heuristic for the QKP that is capable of consistently finding solutions of such high quality. This high performance is especially remarkable in light of the inapproximability result of Khot [10], mentioned in the introduction.

On the other hand, since our heuristic runs in $\mathcal{O}(n^2c)$ time and takes $\mathcal{O}(nc)$ space, it could not be used for instances with very large values of c . A possible way to proceed when c is large is to divide the weights w_i and the knapsack capacity c by a constant, round up each weight to the nearest integer, and round down the capacity. This could cause some feasible solutions to be lost, but would make the running time more manageable.

Another possible approach worth exploring is the following. A folklore trick in the literature on the standard KP (see, e.g., Sections 2.3 and 2.6 of [9]) is to define the following function:

$$f'(k, r) = \min \left\{ \sum_{i=1}^k w_i x_i : \sum_{i=1}^k p_i x_i \geq r, x_i \in \{0, 1\} (i = 1, \dots, k) \right\}.$$

Using a modified dynamic programming recursion, one can compute $f'(n, r)$ for $k = 1, \dots, n$ and $r = 0, \dots, p^*$, where p^* is the optimal profit of the KP instance. The running time then becomes $\mathcal{O}(np^*)$ rather than $\mathcal{O}(nc)$. It may be possible to adapt this idea to the QKP, and thereby obtain a DP heuristic that is suitable for instances in which the capacity c is large but the profits q_{ij} are small.

Finally, we remark that we examined the solutions obtained by our heuristic on some of the test instances. In some cases, one could easily improve the solution by applying a simple local search heuristic, involving the removal of one item and the insertion of one or two other items. Potentially, then, the performance of our heuristic could be improved even further with little computational effort.

Acknowledgement: The authors thank Alberto Caprara for kindly providing us with a copy of the exact QKP code described in the paper [5].

References

- [1] E. Balas & E. Zemel (1980) An algorithm for large zero-one knapsack problems. *Oper. Res.*, 28, 1130–1154.
- [2] R.E. Bellman (1957) *Dynamic Programming*. Princeton, NJ: Princeton University Press.
- [3] A. Bhaskara, M. Charikar, E. Chlamtac, U. Feige & A. Vijayaraghavan (2010) Detecting high log-densities: an $\mathcal{O}(n^{1/4})$ approximation for densest k -subgraph. In L.J. Schulman (ed.) *Proceedings of the 42nd ACM Symposium on Theory of Computing*, pp. 201–210. ACM Press.
- [4] A. Billionet & F. Calmels (1996) Linear programming for the quadratic knapsack problem. *Eur. J. Oper. Res.*, 92, 310–325.
- [5] A. Caprara, D. Pisinger & P. Toth (1998) Exact solution of the quadratic knapsack problem. *INFORMS J. Comput.*, 11, 125–137.
- [6] G.B. Dantzig (1957) Discrete-variable extremum problems. *Oper. Res.*, 5, 266–288.
- [7] G. Gallo, P.L. Hammer & B. Simeone (1980) Quadratic knapsack problems. *Math. Program. Stud.*, 12, 132–149.
- [8] R.M. Karp (1972) Reducibility among combinatorial problems. In R.E. Miller & J.W. Thatcher (eds.) *Complexity of Computer Computations*, pp. 85–103. New York: Plenum.
- [9] H. Kellerer, U. Pferschy & D. Pisinger (2004) *Knapsack Problems*. Berlin: Springer.
- [10] S. Khot (2006) Ruling out PTAS for graph min-bisection, dense k -subgraph, and bipartite clique. *SIAM J. Comput.*, 36, 1025–1071.
- [11] S. Martello & P. Toth (1990) *Knapsack Problems: Algorithms and Computer Implementations*. Chichester: Wiley.
- [12] P. Michelon & L. Veilleux (1996) Lagrangean methods for the 0-1 quadratic knapsack problem. *Eur. J. Oper. Res.*, 92, 326–341.
- [13] D. Pisinger (2007) The quadratic knapsack problem — a survey. *Discr. Appl. Math.*, 155, 623–648.
- [14] D. Pisinger, A.B. Rasmussen & R. Sandvik (2007) Solution of large quadratic knapsack problems through aggressive reduction. *INFORMS J. Comput.*, 19, 280–290.