

# Analysis of Baseline Evolutionary Algorithms for the Packing While Travelling Problem

Vahid Roostapour<sup>1</sup>, Mojgan Pouthassan<sup>1</sup>, and Frank Neumann<sup>1</sup>

<sup>1</sup>Optimisation and Logistics, School of Computer Science  
The University of Adelaide, SA, Australia

{vahid.roostapour,mojgan.pourhassan,frank.neumann}@adelaide.edu.au

February 14, 2019

## Abstract

Although the performance of base-line Evolutionary Algorithms (EAs) on linear functions has been studied rigorously, the same theoretical analyses on non-linear objectives are still far behind. In this paper, variations of the Packing While Travelling (PWT), also known as a non-linear knapsack problem, is considered to address this gap. We investigate PWT for two cities with correlated weights and profits using single-objective and multi objective algorithms. Our results show that RLS finds the optimal solution in  $O(n^3)$  expected time while the GSEMO enhanced with a specific selection operator to deal with exponential population size, calculates all the Pareto front solutions in the same expected time. In the case of uniform weights, (1+1) EA is able to find the optimal solution in expected time  $O(n^2 \log(\max\{n, p_{max}\}))$ , where  $p_{max}$  is the largest profit of the given items. We also validate the theoretical results using practical experiments and present estimation for expected running time according to the experiments.

**Keywords**— Evolutionary algorithms, Packing while travelling problem, Runtime analysis, (1+1) EA

## 1 Introduction

Being a general problem solver combined with the quality of found solutions causes the evolutionary algorithms to be widely applied to many combinatorial optimisation and engineering problems. This type of algorithms, similar to other bio-inspired ones, aim to optimise the current solution iteratively by imitating the evolution process in nature.

Random decision making plays an essential role in EAs' basic operations such as selection and mutation. The area of runtime analysis aims to theoretically

investigate the performance of EAs as a specific class of randomised algorithms and many significant contributions has been achieved along this 20 year old path [Pop14, Jan13, NW10]. The baseline evolutionary algorithms, such as RLS, (1+1) EA and GSEMO, suite well for the theoretical analyses, since they can present a general insight for the performance of more complicated EAs. Interestingly, it has been proven that they perform efficiently in many cases [DJW02, NW10].

Many of the studies have considered the linear class of problems in which the objective function is linear [FKL<sup>+</sup>18]. Friedrich et al. proved that (1+1)EA finds the optimal solution of any linear function under the uniform constraint in  $O(n^2 \log(B \cdot p_{max}))$ , where  $B$  is the constraint and  $p_{max}$  is the maximum profit of the given items.

However, the literature is not equally rich for investigations of EAs on non-linear functions. Moreover, there are many single-component and multi-component problems that challenge EAs. Travelling salesman problem (TSP) for instance, is a well-considered single-component problem in which the only goal is to minimise the cost of the tour. The multi-component problems, on the other hand, gained much attention during the recent years, since they include many real-world problems [BMWN19]. The actual interactions between different components build a more complicated cases in a way that the solutions must be acceptable in all the components. Travelling thief problem (TTP) is the extension of TSP which considers the Knapsack problem as the second component [BMB13]. TTP combines two difficult problems such that the effect of the packing plan on benefit function is nonlinear. Using the same formulation, Polyakovskiy and Neumann introduced the Packing While Travelling problem (PWT) which is actually the packing version on TTP. They prove that this problem is also NP-hard. In PWT there are some items distributed between the cities. A vehicle with a specified capacity wants to visit cities in predefined order and pick the items such that maximise the benefit function and does not violate the capacity. Each item has a profit and weight. The benefit function has a direct relationship with the profit of packed item while the relationship with velocity is inverse. The velocity between each of the two cities is determined by the weight of the vehicle such that the higher weight, the less velocity.

According to the benefit function of PWT, we use this problem to investigate the performance of EAs on nonlinear functions. In this paper we theoretically analyse the performance of three baseline evolutionary algorithms, RLS, (1+1) EA and GSEMO, on the packing while travelling problem and prove an upper bound for the expected running time. We prove that for the instances with correlated weights and profits, RLS and GSEMO find the optimal solution in expected time  $O(n^3)$ . Furthermore, we consider the instances with uniform weights and prove that (1+1)EA finds the optimal solution in expected time  $O(n^2 \log(\max\{n, p_{max}\}))$ . We also investigate the performance of these algorithms in addition to two other multi-objective algorithms from the practical point of view. We present a more precise bounds from our empirical analysis.

The rest of the paper is organised as follows. Section 2 presents the detailed definition of considered Packing While Travelling problem and the algorithms

used in this study. In Section 3, we investigate the problem theoretically for correlated weights and uniform weights in Sections 3.1 and 3.2, respectively. Our experimental analyses are presented in Section 4, followed by a conclusion in Section 5.

## 2 Preliminaries

In this section we present the definition of packing while travelling problem and the detail of algorithms we analyse in this paper.

### 2.1 Problem Definition

The general Packing While Traveling problem proposed on 2016 can be seen as a Traveling Thief Problem when the tour is fixed [BMB13, PN17, PN15]. Given a set of  $m + 1$  ordered cities, distances  $d_i$  from city  $i$  to  $i + 1$  ( $1 \leq i \leq m$ ) and a set of items  $N = \cup_{i=1}^m N_i$  distributed over first  $m$  cities such that city  $1 \leq i \leq m$  contains items  $N_i$ . Let  $|N_i| = n_i$  denote the number of solutions in city  $i$ . Positive integer profit  $p_{ij}$  and weight  $w_{ij}$  are assigned to each item  $e_{ij} \in N_i$ ,  $1 \leq j \leq n_i$ . Moreover, route  $M = (1, 2, \dots, m + 1)$  is travelled by a vehicle with velocity  $v = [v_{min}, v_{max}]$  and capacity  $C$ . A solution vector  $s = (x_{11}x_{12} \dots x_{1n_1} \dots x_{mn_m})$  represents a set of selected items  $S \subseteq N$  such that variable  $x_{ij} \in \{0, 1\}$  indicate whether item  $e_{ij}$  is selected or not. Let  $W(s)$  denote the sum of weights of items in  $s$  and  $s$  is feasible if  $W(s) \leq C$ . Let

$$P(s) = \sum_{i=1}^m \sum_{j=1}^{n_i} p_{ij} x_{ij}$$

be the total profit and

$$T(s) = \sum_{i=1}^m \frac{d_i}{v_{max} - \nu \sum_{k=1}^i \sum_{j=1}^{n_k} w_{kj} x_{kj}}$$

where  $\nu = \frac{v_{max} - v_{min}}{C}$  is a constant, depicts the total travel time for vehicle carrying a subset of selected items  $s$ . The denominator of  $T(s)$  is such that picking an item in city  $i$  only affects the time to travel from city  $i$  to the end. Then the benefit function of  $s$  is computed as:

$$B(s) = P(s) - R \cdot T(s)$$

where  $R$  is a given renting rate. The aim is to find a feasible solution  $s^* = \max_{s \in \{0,1\}^n} B(s)$ .

The effect on different values for  $R$  on the benefit function has been considered previously [WPN16]. Generally speaking,  $R = 0$  changes PWT to the 0-1 knapsack problem while a larger  $R$  forces the optimal solution to pick fewer items. This problem is proven to be  $\mathcal{NP}$ -hard by reducing the subset sum problem to the decision variant of unconstrained PWT [PN17].

---

**Algorithm 1: RLS**

---

```
1 Choose  $s \in \{0, 1\}^n$  uniformly at random;
2 Let  $|s|_1$  denote the number of items in  $s$ ;
3 while stopping criterion not met do
4    $p \leftarrow$  a random real number in  $[0, 1]$ ;
5   if  $|s|_1 = 0 \vee |s|_1 = n \vee p < 1/2$  then
6     Create  $s'$  by flipping a randomly chosen bit of  $s$  ;
7   else
8     Create  $s'$  by flipping a randomly chosen zero bit and a randomly
      chosen one bit of  $s$  ;
9   if  $F(s') \geq F(s)$  then
10     $s \leftarrow s'$ ;
```

---

However, in this paper we consider another version of this problem where we have only two cities and  $n$  items that are located in the first city. Moreover, the weights of the given items are favourably correlated with the profits i.e. for any two item  $e_i$  and  $e_j$ ,  $p_i \geq p_j$  implies  $w_i \leq w_j$ . Hence, for a bit string solution  $s = (x_1 \cdots x_n) \in \{0, 1\}^n$ , we have:

$$B(s) = \sum_{i=1}^n x_i \cdot p_i - \frac{Rd}{v_{max} - \nu \sum_{i=1}^n x_i \cdot w_i}.$$

Without loss of generality, in the rest of the paper we assume that items are indexed in a way that  $p_1 \geq p_2 \cdots \geq p_n$  and  $w_1 \leq w_2 \cdots \leq w_n$ . We also consider another version of the problem in which all the weights are uniform and equal to one.

## 2.2 Algorithms

In this paper we study the behaviour of three algorithms. The first one, described in Algorithm 1, is the Random Local Search (RLS) which is also able to do a swap (flip a zero bit and a one bit simultaneously), in addition to the usual one-bit flip. In each iteration, if there are no ones or no zeros in the current solution, it flips a randomly chosen bit of the solution. Otherwise, with probability of  $1/2$ , either does a one-bit flip as described or chooses a one and a zero uniformly at random and flips both of them. The generated offspring replaces the current solution if it is better with respects to the fitness function. This swap mutation is added to RLS because there are some situations in optimising PWT that no one-bit flip is able to pass the local optima.

Another algorithm we consider is (1+1) EA (Algorithm 2). This algorithm flips each bit of the current solution with probability of  $1/n$  in mutation step. Similar to RLS, it compares the parent and the offspring and picks the better one for the next generation.

---

**Algorithm 2:** (1+1) EA

---

```
1 Choose  $s \in \{0, 1\}^n$  uniformly at random;
2 while stopping criterion not met do
3   Create  $s'$  by flipping each bit of  $s$  independently with probability of
      $1/n$ ;
4   if  $F(s') \geq F(s)$  then
5      $s \leftarrow s'$ ;
```

---

---

**Algorithm 3:** GSEMO

---

```
1 Choose  $s \in \{0, 1\}^n$  uniformly at random;
2  $P \leftarrow \{s\}$ ;
3 while stopping criterion not met do
4   Let  $P_i = \{s \in P \mid |s|_1 = i\}$ ,  $0 \leq i \leq n$  and  $I = \{i \mid P_i \neq \emptyset\}$ ;
5   Choose  $j \in I$  uniformly at random;
6    $s \leftarrow \arg \max\{B(x) \mid x \in P_j\}$ ;
7   Create  $s'$  by flipping each bit of  $s$  independently with probability of
      $1/n$ ;
8   if  $\{z \in P : z \succ s'\} = \emptyset$  then
9      $P \leftarrow P \setminus \{z \in P \mid s' \succeq z\} \cup \{s'\}$ ;
```

---

In RLS and (1+1) EA, which are single-objective algorithms, the comparisons between solutions is based on the fitness function

$$F(s) = (q(s), B(s))$$

where  $q(s) = \min\{C - w(s), 0\}$ . According to  $q(s)$ ,  $s$  is infeasible if and only if  $q(s) < 0$  and the absolute value of  $q(s)$  denote the amount of constraint violation. The goal is to maximise  $F(s)$  with respect to lexicographic order, i.e.  $s_1$  is better than  $s_2$  ( $F(s_1) \geq F(s_2)$ ) if and only if  $(q(s_1) > q(s_2)) \vee (q(s_1) = q(s_2) \wedge B(s_1) \geq B(s_2))$ . This implies that any feasible solution has better fitness than any infeasible solution. Moreover, between two infeasible solutions, the one with smaller constraint violation is better.

We also consider PWT with a multi-objective algorithm using a variant of GSEMO, which uses a specific selection function, to deal with the exponential size of the population (Algorithm 3). Neumann and Sutton suggested this version of GSEMO for Knapsack problem with correlated weights and profit to avoid exponential population size [NS18]. We use the same approach since PWT easily changes to KP when  $R = 0$ . As the objectives, we use the weight function ( $W(s)$ ) and the previously defined fitness function ( $F(s)$ ). The aim is to minimise  $W(s)$  while maximising  $F(s)$ . Between two solutions  $s_1$  and  $s_2$ , we say  $s_1$  (weakly) dominates  $s_2$ , denoted by  $s_1 \succeq s_2$ , if and only if  $W(s_1) \leq W(s_2) \wedge F(s_1) \geq F(s_2)$ . The dominance is called strong, denoted by

$s_1 \succ s_2$ , when at least one of the inequalities strictly holds. Note that based on this definition, similar to the single-objective fitness function, each feasible solution dominates all infeasible solutions and an infeasible solution closer to the constraint bound dominates the farther ones.

### 3 Theoretical Analysis

In this section we investigate the performance of RLS, GSEMO and (1+1) EA on different versions of PWT problem from theoretical aspects. We first consider RLS and GSEMO on the PWT with correlated weights and profits. Next, the behaviour of (1+1) EA on the PWT with uniform weights is analysed.

To study the PWT problem, we need to investigate the properties of an optimal solution for this problem, and the impact of adding or removing an item on the benefit function. For this reason, in the following lemma we prove that the weight of the current solution and  $p_i$  determine if item  $e_i$  is worth to be added to or removed from the current solution.

**Lemma 1.** *For each item  $e_i$  there is a unique weight  $w_{e_i}$  such that adding  $e_i$  to the current solution  $s$  improves the benefit function only if  $W(s) < w_{e_i}$ . Moreover, only if  $W(s) > w_{e_i} + w_i$  then removing  $e_i$  increases the benefit function.*

*Proof.* Assume that the current solution  $s$  does not include item  $e_i$  and  $s' = s \cup e_i$ . Hence, we have:

$$\begin{aligned} B(s') - B(s) &= (P(s') - RT(s')) - (P(s) - RT(s)) \\ &= p_i - Rd \left( \frac{1}{v_{max} - \nu(W(s) + w_i)} - \frac{1}{v_{max} - \nu W(s)} \right) \\ &= p_i - \frac{Rd\nu w_i}{(v_{max} - \nu(W(s) + w_i)) \cdot (v_{max} - \nu W(s))}. \end{aligned} \quad (1)$$

Adding  $e_i$  to  $s$  increases  $B(s)$  only if the Expression 1 is greater than zero. Solving this equation we have:

$$B(s') - B(s) > 0 \iff W(s) < \overbrace{\frac{v_{max}}{\nu} - \frac{w_i}{2}}^{w_{e_i}} \left( 1 + \sqrt{\frac{4Rd}{\nu w_i p_i}} \right) \quad (2)$$

As it can be observed for item  $e_i$ ,  $w_{e_i}$  only depends on the weight of the current solution. Hence, by solving the Equation 1 for each  $e_i$ ,  $1 \leq i \leq n$ , we can find  $w_{e_i}$  such that adding  $e_i$  to  $s$  improves  $B(s)$  only if  $W(s) < w_{e_i}$ . Considering the case that  $e_i \in s$  and  $s' = s \setminus \{e_i\}$ , similar calculations implies that:

$$\begin{aligned} B(s') - B(s) > 0 \iff W(s) &> \frac{v_{max}}{\nu} + \frac{w_i}{2} \left( 1 - \sqrt{\frac{4Rd}{\nu w_i p_i}} \right) \\ &> w_{e_i} + w_i. \end{aligned} \quad (3)$$

In other words, removing  $e_i$  from  $s$  increases the benefit function only if  $W(s) > w_{e_i} + w_i$ , which completes the proof.  $\square$

A straight result from Equations 2 and 3 is that:

$$p_i \geq p_j \wedge w_i \leq w_j \implies w_{e_i} \geq w_{e_j}$$

and the equality holds only if  $p_i = p_j \wedge w_i = w_j$ . Therefore, we have  $w_{e_1} \geq w_{e_2} \geq \dots \geq w_{e_n}$ . Moreover, if weight of solution  $s$  equals to  $w_{e_i}$  then  $B(s) = B(s \cup e_i)$ .

Now we argue on the optimal solution of the PWT problem. Let  $s_i = \langle 1^i, 0^{n-i} \rangle$  denote the solution that only includes the first  $i$  items and  $s_0 = \langle 0^n \rangle$ . Consider the case that for some  $i$ , we have  $w_{e_i} > W(s_{i-1})$ . Hence, by Lemma 1, adding  $e_i$  to  $s_{i-1}$  improves the benefit function and  $B(s_i) > B(s_{i-1})$ . Moreover, since  $w_{e_1} \geq \dots \geq w_{e_i}$ , we have  $B(s_i) \geq \dots \geq B(s_1)$ . The claim is that for a given set of items, there is a unique  $k$  such that the optimal solution of PWT problem is packing  $k$  or  $k \wedge (k+1)$  items with the least indices i.e  $s_k$  or  $s_{k+1}$ . The following lemma proves this claim.

**Lemma 2.** *The optimal solution of Packing While Traveling problem is the set of  $k$  or  $k \wedge (k+1)$  items with highest profits and least weights where  $k$  is unique and depends on the given set of items.*

*Proof.* First, we assume that  $w_{e_n} > W(s_{n-1})$ . In this case, we have  $B(s_n) > B(s_{n-1})$  and  $s_n$  is the optimal solution. On the other hand, if  $w_{e_1} < 0$  then the optimal solution is  $s_0$ . In the rest of the proof, we assume that none of these cases happen.

Let  $o = \min\{i \mid W(s_i) > w_{e_{i+1}}, 0 \leq i \leq n-1\}$ . According to this definition  $w_{e_{o+1}}$ , adding any item to  $s_o$  decreases the benefit function. Moreover, we have  $W(s_{o-1}) < w_{e_o}$  that implies  $W(s_{o-1}) + w_o = W(s_o) < w_{e_o} + w_o$ . Hence, removing any item from  $s_o$  also reduces the benefit function. Therefore, the following equation holds:

$$B(s_0) < \dots < B(s_o) \geq B(s_{o+1}) > \dots > B(s_n).$$

The equality  $B(s_o) = B(s_{o+1})$  holds only when  $w_{e_{o+1}} = W(s_o)$ . However, all other inequalities are strict according to Lemma 1.

To prove that  $s_o$  or  $s_o \wedge s_{o+1}$  are the only optimums, it is now enough to show that for any  $i$ ,  $s_i$  has the highest benefit value among other solutions with  $i$  items. This is also true since the items in  $s_i$  have the highest profits and the least weights which results in the highest benefit value.

Involving the capacity constraint  $C$ , however, may change the optimal solution. We define  $s_k$ ,  $k = \max\{j \mid W(s_j) \leq C, 0 \leq j \leq o\}$ , the feasible solution with the highest benefit function which is the actual optimal solution with respect to  $C$ . This finalises the proof.  $\square$

Note that there could be more than one item with the same weights and profits as  $e_k$  and there is no priority among them to be in the optimal solution. From this point, we denote the optimal solution, its weight and its profit by  $s_k$ ,  $w^k$  and  $p^k$ , respectively.

### 3.1 Correlated weights and profits

In this section, we consider the instances of packing while travelling problem in which the weights are strongly correlated with the profits. We calculate the performance of RLS and GSEMO for this type of PWT.

#### 3.1.1 RLS

Using the result of Lemma 2, we prove the performance of RLS in finding the optimal solution of PWT problem in terms of the number of evaluations. We refer to the first  $k$  bits of  $s_k$  as the first block and the rest as the second block. Moreover, let  $l$  and  $r$  denote the number of zeros in the first block and the number of ones in the second block, respectively. For technical reason assume item  $e_0$  exists where  $w_{e_0} > \sum_{i=1}^n w_i$  and  $x_0 = 1$ . We define

$$h_s = \max\{0 \leq i \leq k \mid x_0 = x_1 = \dots = x_i = 1 \wedge W(s) \leq w_{e_i}\},$$

index of a specific bit in solution  $s$ . Assume  $s^t$  denote the solution achieved by RLS after  $t$  generations and  $h_t = \max\{h_{s^i} \mid 1 \leq i \leq t\}$ . The following lemma and theorem investigate the performance of RLS on the PWT problem with correlated weights.

**Lemma 3.** *Having achieved a solution  $s$  with  $h_t = h_s$ , RLS does not accept a solution  $s' = \langle x'_1, \dots, x'_n \rangle$  in which  $\exists i \leq h_t : x'_i = 0$ .*

*Proof.* Since the weights and profits are correlated, it is enough to prove that RLS does not remove  $e_{h_t}$ . Let  $s'$  denote the first solution in generation  $t+1$  that  $W(s') \geq w_{e_{h_t}} + w_{h_t}$ , i.e. RLS is able to remove  $e_{h_t}$  from  $s'$ . Note that all the items  $e_i, i \leq h_t$ , are still in  $s'$ . Moreover, no swap mutation is able to remove  $e_{h_t}$  since it has higher profit and less weight than other missed items. Hence, there exist the last item  $e_x, x > h_t$ , added to  $s'$  such that  $W(s') - w_x < w_{e_{h_t}} + w_{h_t}$ ,  $C \geq W(s') \geq w_{e_{h_t}} + w_{h_t}$  and  $B(s') > B(s' \setminus e_x)$ . From the benefit inequality and Lemma 1, we have

$$W(s') - w_x < w_{e_x} \Rightarrow W(s') < w_{e_x} + w_x.$$

This contradicts with the condition of  $s'$  to remove  $e_{h_t}$  and completes the proof.  $\square$

According to the definition of  $h_t$ , the weight of the accepted solutions after generation  $t$  is lower bounded by  $W(s_{h_t})$  in which the first  $h_t$  items are selected only. It illustrates that the maximum possible value for  $h_t$  is  $k$ , otherwise  $s_k$  is not the optimal solution (Lemma 2). Let  $x_{h_t+1} = 0$  and  $y = \min\{i \mid i > h_t \wedge x_i = 1\}$ , be the index of first one bit after a sequence of zeros after  $h_t$ . Similar to the proof of Lemma 3, the following corollary holds.

**Corollary 4.** *Achieving a solution  $s$  with  $W(s) < w_{e_y}$ , RLS does not remove  $e_y$  unless by reducing the value of  $y$ .*



**Theorem 5.** *RLS finds the optimal solution of Packing While Travelling problem with correlated weights and profits in  $O(n^3)$  expected time.*

*Proof.* Let RLS start with a random solution  $s^0$ . We analyse the optimisation process in three main phases. In the first phase, RLS achieves a feasible solution. The second phase is to achieve  $h_t = k$  and the third phase is to remove the remaining items from the second block to achieve the optimal solution.

If  $W(s^0) \leq C$  the first phase is already finished. Therefore, let  $W(s^0) \geq C$ . Using a fitness level argument, Neumann and Sutton proved that (1+1) EA, which uses a fitness function with strictly higher priority in weight constraint satisfaction, finds a feasible solution for KP with correlated weights and profits in  $O(n^2)$  expected time (Theorem 3 in [NS18]). Their proof also holds for RLS since the constraint is linear and RLS is able to do a one-bit flip in  $O(n)$  expected time. Hence, RLS achieves a feasible solution  $s$  in  $O(n^2)$  expected time and completes the first phase.

In this phase, we analyse the expected time needed to increase the value of  $h_t$  by at least one. We need to calculate the expected time to find a solution  $s$  such that  $x_{h_t+1} = 1$  and  $W(s) \leq w_{e_{h_t+1}}$ . Let  $s$  denote the current solution. According to Lemma 1 and 2, if  $W(s) \geq w_{e_{k'+1}} + w_{k'+1}$ ,  $k' \leq k$ , removing any item  $e_i$  where  $i \geq k'$  is beneficial with respect to the benefit function. The probability of such one-bit flis is at least  $1/(2n)$ . Since there are at most  $n$  items to be removed, using the additive drift, RLS achieves  $s$  with  $W(s) < w_{e_{k'+1}} + w_{k'+1}$  in  $O(n^2)$  expected time. Note that after this point, no item can be removed from  $s$  with a one-bit flip.

Now we consider the item  $e_{h_t+1}$ . According to the order of  $w_{e_i}$ ,  $1 \leq i \leq n$ ,  $e_{h_t+1}$  is the first item that could be added a solution. Note that  $h_t < k$ , otherwise RLS has already achieved the optimal solution. We show that after  $n^2$  expected time, RLS achieves a solution such that  $x_{h_t+1} = 1$  and the value of  $h_t$  is increased by one. Remind that  $y > h_t$  denotes the index of the first one bit after  $h_t$  and we assume  $n > y \neq h_t + 1$ . Since no item can be removed with one-bit flips, then the only possible moves are two-bit flips. Each accepted two-bit flip is a swap that reduces the weight of  $s$ . Let  $\beta = |\{e_i | i \geq h + 1 \wedge x_i = 1\}|$  be the number of items in the solution with higher index than  $h$ . Since the first  $h$  items can not be removed,  $\alpha = \min\{n - h - \beta, \beta\}$  is the number of possible two-bit flips. The  $i$ th swap happens with the probability of  $\frac{(\alpha-i+1)^2}{2n^2}$ . Therefore, all the two-bit flips take place in expected time

$$2n^2 \cdot \sum_{i=1}^{\alpha} \frac{1}{i^2} = O(n^2).$$

After this time, RLS cannot reduce the weight of the solution anymore. Moreover, since the  $x_k = 1$ , RLS has done a two bit flip which added  $e_{h_t+1}$ . Moreover,  $e_{h_t+1}$  is better than  $e_k$  with respect to the benefit function and Corollary 4 shows that  $e_k$  cannot be removed. Hence,  $e_{h_t+1}$  cannot be removed.

At this point, using  $O(n^2)$  expected time, RLS has achieved a solution such that  $x_{h_t+1} = 1$  and the weight cannot be reduced anymore. This guarantees

that either  $W(s) < w_{e_{h_t+1}}$  and  $h_{t+1} = h_t + 1$ ; or RLS has achieved the optimal solution in which first  $h_t$  items are in the solution and adding or removing any item decreases the benefit function (Lemma 2). Thus in expected time  $O(n^2)$  RLS increases  $h_t$  by one. Since the maximum value of  $h_t$  is  $k \leq n$ , RLS achieves  $h_t = k$  in expected time  $O(n^3)$ .

Finally in the third phase, RLS needs to remove the remaining items from the second block. Note that the first  $k$  items are now selected and cannot be removed anymore. Each item can be removed with a one-bit flip which results in  $O(n^2)$  expected time for this phase.

Therefore we can conclude that RLS finds the optimal solution of the PWT with correlated weights and profits in  $O(n^3)$  expected time.  $\square$

### 3.1.2 GSEMO

In this section we consider the time performance of GSEMO on the PWT problem with correlated weights and profits. The two objectives used in this algorithm are the weight function and the lexicographical fitness function, denoted by  $W(s)$  and  $F(s)$ , respectively. We say solution  $s_1$  dominates solution  $s_2$ , denoted by  $s_1 \succeq s_2$ , if  $W(s_1) \leq W(s_2) \wedge F(s_1) \geq F(s_2)$ . In case of at least one strict inequalities, it is called strongly dominance. In our analysis which is inspired from [NS18], we use a fitness level argument on the weights of the solution and compute the expected time needed to find at least one Pareto solution. Next, we calculate the time for finding all the Pareto front. Due to lemma 2, we can observe the following corollary that describes the Pareto front structure.

**Corollary 6.** *The Pareto set corresponding to the Packing While Travelling problem with correlated weights and profits is the solution set  $\{s_0, \dots, s_k\}$ .*

*Proof.* As it is explained in the proof of Lemma 1  $s_i, i \leq n$ , has the highest profit among all the solutions with size  $i$ . Moreover, since the weights correlated, it also has the least weight. Hence,  $s_i, i \leq n$ , dominates all the solutions with size  $i$ . On the other hand, we have  $B(s_0) \leq \dots \leq B(s_k)$  while  $W(s_0) \leq \dots \leq W(s_k)$  which implies that  $\{s_0, \dots, s_k\}$  do not dominate each other. Moreover,  $s_k$  dominates every solution  $s_i, i > k$ , since it has higher fitness and less weight. Thus, there exist no solution dominating  $\{s_0, \dots, s_k\}$  which completes the proof.  $\square$

The following theorem proves that the expected time for GSEMO to find the whole Pareto front of PWT problem with correlated weights and profits is  $O(n^3)$ .

**Theorem 7.** *GSEMO finds all the non-dominated solutions of Packing While Travelling problem with correlated weights and profits in  $O(n^3)$  expected time.*

*Proof.* The proof consists of two phases. In the first phase, GSEMO finds at least one of the Pareto solutions,  $\{0\}^n$ . In the second phase, GSEMO finds other Pareto solutions based on the assumption of having at least one.

Note that according to the definition of  $F(s)$ , an infeasible solution with less constraint violation always dominates the other ones, even in this two-objective space. Hence, while GSEMO has not achieved a feasible solution, the size of its population remains one. Therefore, it behaves exactly the same as (1+1) EA until it finds a feasible solution which is proven to take  $O(n^2)$  expected time. In the rest of the analysis of the first phase, we assume that GSEMO has found a feasible solution.

To analyse the first phase, we define  $n$  fitness levels  $A_i$ ,  $0 \leq i \leq n-1$ , based on the weight objective  $W(s)$  as follows:

$$\begin{aligned} A_i &= \{x \mid x = s_0\} & i = 0 \\ A_i &= \{x \mid W(s_{i-1}) < W(x) \leq W(s_i)\} & 1 \leq i \leq n. \end{aligned}$$

According to the correlation of the weights, this definition guarantees that if a solution  $s$  belongs to level  $A_i$ ,  $1 \leq i \leq n$ , then  $s$  includes at least one of the items  $e_j$  where  $j \geq i$ . Moreover, removing  $e_j$  from  $s$  takes it to a lower level since  $w_1 \leq \dots \leq w_{i-1} \leq w_i \leq \dots \leq w_n$ . Now assume that  $P$  denotes the population in step  $t$  of GSEMO. Let  $A_u$  denote the least level that has been achieved until step  $t$  and  $s \in A_u$  has the highest fitness among the solutions in level  $A_u$ . GSEMO chooses this solution by choosing  $|s|$  from  $I$ , which happens with the probability of  $1/|I| \geq 1/n$ . Furthermore, there exists item  $e_x$  in  $s$  such that  $x \geq u$  and removing it produces a solution in a lower level. This solution will be accepted since it has the least weight. The one-bit flip that removes  $e_x$  from  $s$  happens with the probability of  $1/(en)$ . Thus, GSEMO reduces  $u$  with probability of  $1/(en^2)$  in step  $t$ . In other words, GSEMO reduces  $u$ , at least by one, in expected time  $O(n^2)$ . On the other hand,  $s_0$ , which is the only solution of level  $A_0$  is a Pareto solution since it has the least weight. Thus, if the algorithm achieves  $A_0$  then the first phase is accomplished. Since the maximum possible value of  $u$  is  $n$  and  $u$  is reduced by one in every  $O(n^2)$  expected iterations, GSEMO finishes the first phase in expected time of  $O(n^3)$ .

In the second phase, we assume  $s_i \in P$ ,  $i \leq k$  exists such that  $s_i$  is Pareto solution by Corollary 6. Moreover, we assume either  $s_{i-1}$  or  $s_{i+1}$ , which are also Pareto solutions, do not exist in  $P$ . Otherwise, the second phase is already finished. Adding  $e_{i+1}$  to or removing  $e_{i-1}$  from  $s_i$  result in  $s_{i+1}$  or  $s_{i-1}$ , respectively. For such a step, the algorithm must choose  $i \in I$  and flip the correct bit which happens with the probability  $1/n$  and  $1/(en)$ , respectively. Hence, the algorithm finds a new Pareto solution from  $s_i$  in  $O(n^2)$  expected iterations. Based on Corollary 6, the size of the Pareto set is  $k+1 \leq n$ . Therefore, GSEMO finds all the  $k+1$  Pareto solutions of the PWT with correlated weights and profits in  $O(n^3)$  expected time.  $\square$

### 3.2 Uniform Weights

In this section, we analyse the performance of (1+1) EA on another version of Packing While Travelling problem. Here we assume that weights of all the items are one and the profits are arbitrary. Similar to the previous instances,

we assume items  $\{e_1, \dots, e_n\}$  are indexed such that  $p_1 \geq \dots \geq p_n$ . Note that the results of Lemmata 1 and 2 also holds for the uniform weights. Moreover, since correlated weights and profits are actually the more general version of the uniform weights, the results for GSEMO and RLS also holds. Additionally, we have  $w^k = k$ .

The analysis of (1+1) EA is based on the following Lemma which proves the existence of a set of one-bit flips and two-bit flips which transform an arbitrary solution to the optimal solution.

**Lemma 8.** *Let  $w^k$  be the weight of the optimal solution and the current solution  $s$  is feasible. There exists a set of one-bit flips and two-bit flips that transform  $s$  to an optimal solution if happen in any order.*

*Proof.* Assume that  $W(s) \leq w^k$ . This implies that  $i \leq k$ , where  $i$  is the number of items in  $s$ . In this case, any two-bit flips that swap a one bit from the second block and a zero bit from the first block will be accepted by the algorithm. If there is no one bit in the second block, then all the one-bit flips that change a zero bit in the first block is accepted and this set leads  $s$  to the optimal solution  $s_k$ . Note that all defined the one-bit flips will be accepted (Lemma 1 and Lemma 2) since  $i \leq k$  and the weights are one. Hence, it is not necessary that the bit flips happen in a special order. The other case that  $C \geq W(s) > w^k$  is similar to the first one. Any two-bit flips that swap a zero bits of the first block with a one bit on the second block and any one-bit flips that remove ones from the second block are accepted by the algorithm. Thus, in this case, there also exists a set of one-bit flips and two-bit flips that if take place in any order, transform  $s$  into the optimal solution.  $\square$

Here we present some more definitions that help us with analysing the performance of (1+1) EA on the PWT problem. Assume  $M$  is the set of  $m_1$  one-bit flips and  $m_2$  two-bit flips that transform the current solution  $s$  to an optimal solution and  $|M| = m_1 + m_2$ . Moreover, let  $g(s) = B(s_k) - B(s)$  be the difference between the benefit of  $s$  and the optimal solution. Therefore, we can denote the contribution of one-bit flips and two-bit flips in  $g(s)$  by  $g_1(s)$  and  $g_2(s)$ , respectively, such that  $g(s) = g_1(s) + g_2(s)$ . In the next theorem, we calculate the expected time for (1+1) EA to find the optimal solution of PWT problem with uniform weights.

**Theorem 9.** *(1+1) EA finds the optimal solution for the Packing While Traveling problem with uniform weights in  $O(n^2 \max\{\log n, \log p_{max}\})$  expected time.*

*Proof.* Let  $\Delta_t = g(s) - g(s') = B(s') - B(s)$  denote the improvement of (1+1) EA in iteration  $t$  which transforms  $s$  to  $s'$ . We partition the proof into two cases. Firstly, let the overall contribution of one-bit flips in  $g(s)$  be more than the total improvement that could be achieved by two-bit flips. Hence, we have  $g_1(s) \geq \frac{g(s)}{2}$ . Since there are  $m_1$  one-bit flips in  $M$ , each of them happens with the probability of  $\frac{m_1}{en}$  and improves the solution in average by  $\frac{g_1(s)}{m_1}$ . Thus, in

this case we have :

$$E[\Delta_t] \geq \frac{g_1(s)}{m_1} \cdot \frac{m_1}{en} \geq \frac{g_1(s)}{en} \geq \frac{g(s)}{2en}.$$

In the second case, the sum of improvement by two-bit flips in  $M$  is more than the total improvement of one-bit flips and we have  $g_2(s) \geq \frac{g(s)}{2}$ . Here, the average improvement of each two-bit flip is  $\frac{g_2(s)}{m_2}$  that takes place with the probability of  $\frac{m_2}{2en^2}$ . Therefore, we have:

$$E[\Delta_t] \geq \frac{g_2(s)}{m_2} \cdot \frac{m_2}{2en^2} \geq \frac{g_2(s)}{2en^2} \geq \frac{g(s)}{4en^2}.$$

We can conclude that the expected improvement at step  $t$  is at least  $E[\Delta_t] \geq \frac{g(s)}{4en^2}$ . On the other hand,  $B(s_k) \leq n \cdot p_{max}$  is the maximum possible benefit by ignoring the cost function. To use the multiplicative drift, it is only needed to calculate the minimum possible amount of  $g(s)$ .

Let  $s^l$  be the last solution that turned into an optimal solution with a bit flip. We need to find the minimum of  $B(s_k) - B(s^l)$  which happens when  $B(s^l)$  is maximized. There are three possibilities that  $s^l$  be the closest solution to the optimal solution. If the last bit flip is a two-bit flip, to maximize  $B(s^l)$ ,  $s^l = s_{k-1} \cup e_{k+1}$ . In this case, we have  $g(s^l) = p_k - p_{k+1} \geq 1$ , since all the profits are integers. If the last bit flip is a one-bit and it adds an item to  $s^l$ , then the maximum benefit of  $s^l$  is achieved when  $s^l = s_{k-1}$ . Therefore, we have  $g(s^l) = p_k - \frac{Rd\nu}{(v_{max}-\nu(k-1))(v_{max}-\nu k)}$ . Considering  $R, d, v_{min}$  and  $v_{max}$  as constants,  $g(s) = O(n^{-q})$  for some constant  $q \geq 1$ . The same results holds for the third case where the final bit flip removes  $e_{k+1}$  from  $s^l = s_{k+1}$ .

Finally, using the multiplicative drift with  $X_0 = n \cdot p_{max}$ ,  $x_{min} = n^{-q}$  and  $\delta = \frac{1}{4en^2}$ , the expected first hitting that (1+1) EA finds the optimal solution of unconstrained PWT problem is:

$$E[T|X_0] \leq 4en^2 \ln(n \cdot p_{max} \cdot n^q) = O(n^2 \max\{\log n, \log p_{max}\}).$$

□

## 4 Experiments

In this section, we compare the performance algorithms on uniform and correlated instance. Moreover, according to the theoretical results, we consider the gap between the proven upper bounds and the practical performance of the algorithms. Finally, we estimate the optimisation time for each algorithm based on the experiments. In this section, we also consider two new multi-objective algorithms called SEMO and RLS\_SEMO. They are different from GSEMO (Algorithm 3) only in the mutation step. The mutation step in RLS\_SEMO is the same as Algorithm 1. SEMO, on the other hand, only choose a uniformly random bit and flips it. We show in previous section that it is essential for an

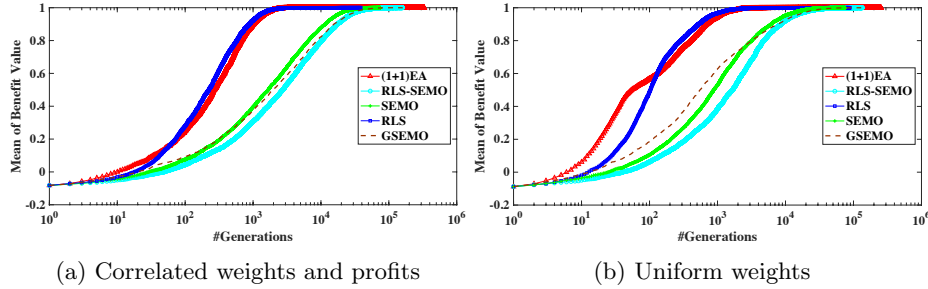


Figure 1: The average of best solution in each generation for RLS, (1+1) EA, GSEMO, SEMO and RLSSEMO in thirty different instances

algorithm to do a two-bit flip to escape from the local optimas, however, our experiments show that SEMO covers this weakness by using the population and perform better than other multi-objective algorithms.

#### 4.1 Benchmark and Experimental setting

To compare the practical performance of different algorithms in different types of PWT, we used thirty different instances where each instance consists of 300 items. Each instance for the correlated PWT is generated by choosing integers uniformly at random within  $[1, 1000]$  such that for any item  $e_i, e_j : i < j$  we have  $p_i \geq p_j$  and  $w_i \leq w_j$ . For the instances with uniform weights, we use the same profits as before but change the weights to one. We set the constant values in all the instances as follow:  $d = 50$ ,  $R = 70$ ,  $v_{max} = 1$  and  $v_{min} = 0.1$ .  $C = 8000$  is the capacity chosen for correlated instances. To calculate the proper capacity for uniform instances, we use the average of the maximum number of items with correlated weights that fit in  $C = 8000$  which results in  $C = 72$  for the uniform instances. Furthermore, for all the experiments, algorithms start with the zero solution.

We use these instances in the first experiment in which we show how algorithms converge to the optimal solution for the correlated and uniform instances. We run each algorithm for all thirty instances, record the best found solution in each generation and normalise its benefit value to the interval  $[0, 1]$  with respect to the optimal benefit of the instance. We plot the average of normalised values as the success rate for each algorithm. Hence, this value is one when an algorithm has found the optimal solution for all thirty different instances.

In the last experiments in which we estimate the optimisation time for algorithms, we run each algorithm on instances with seven different number of items,  $n = 100$ ,  $n = 200$ ,  $n = 500$ ,  $n = 1000$ ,  $n = 2000$ ,  $n = 5000$  and  $n = 10000$ . For each  $n$ , We have thirty different instances which are created with the same constants as previous ones. For each algorithm, we consider the average time to find the optimal solution for the thirty instances with the same  $n$  as the optimisation time. Hence, for each algorithm, we have seven points to estimate

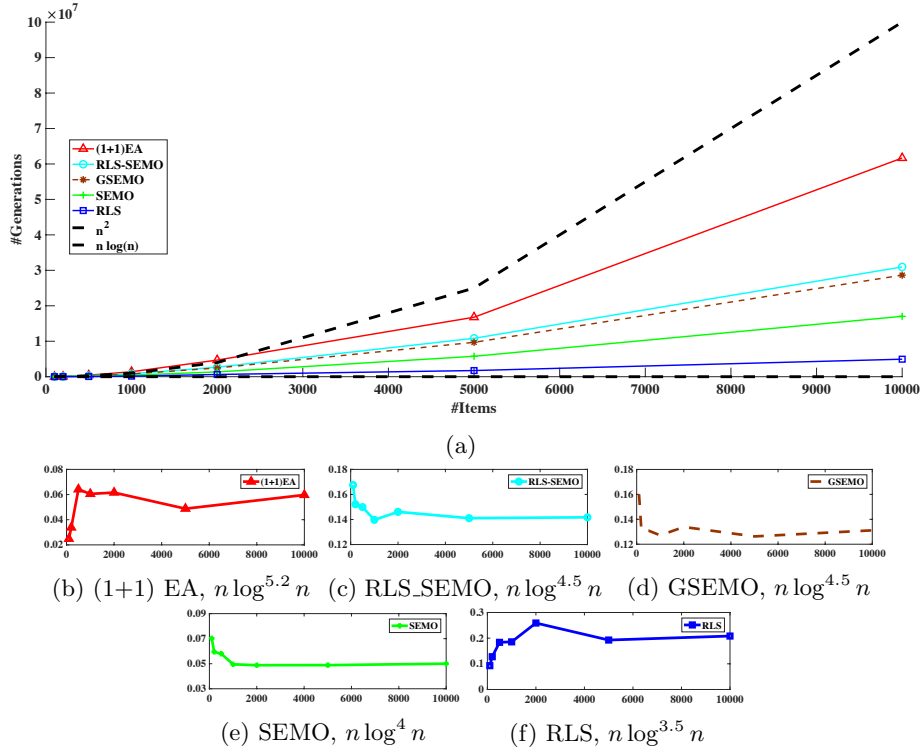


Figure 2: The average and the estimated order of running times for different algorithms to find the optimal solutions for instances with 7 different sizes.

the performance. We only do this analysis with correlated items too.

In our results for (1+1) EA and RLS, we only count the generations that an actual bit flip has happened in mutations steps. This causes the analysis to be fair since RLS guarantees to create a new solution in each mutation step but (1+1) EA and GSEMO do not flip any of the bits with the probability of  $1/e$  (36.7% of times).

## 4.2 Analysis

In this part, we analyse the performance of the algorithms based on the experimental results. Figure 1 illustrates how the algorithms converge to the optimal solutions for all the instances in average. In both types of the instances, it can be observed that (1+1) EA and RLS get close to the optimal solutions much faster than the multi-objective algorithms, which shows that the population slows down the convergence rate significantly. (1+1) EA, however, has major problems in finding the optimal solutions when few bit flips are needed and it takes longer for this algorithm to find the exact optimal solution than the other algorithms. One of the reasons is that (1+1) EA is able to flip more than one

bits. Hence, it is probable to improve the benefit value while increasing the hamming distance between the current solution and the optimal solution. In other words, (1+1) EA is able to improve the benefit function while the improved solution needs more bit flip to become the optimal solution than before.

Looking into the multi-objective algorithms, Figure 1 declares that GSEMO and RLS-SEMO behave almost similar while SEMO outperforms both. This shows that although SEMO can flip only one bit in each generation, the population enhances the algorithm to pass through the local optima and find the optimal solution. Generally speaking, RLS and SEMO which flip the least possible bits, perform better for this Packing while travelling problem.

Now we analyse the results of another experiment presented in Figure 2. We run each algorithm for instances with seven different sizes, and For each specific size, we have thirty different instances. Figure 2a shows the average of running times for each size. It validates the results of Figure 1 in which RLS has a better performance than the other algorithms while multi-objective algorithms outperform (1+1) EA. This data can also be used to estimate the expected optimisation time for each algorithm. For a general estimation, Figure 2a presents the data for  $n^2$  and  $n \log n$  expressions which illustrate that the order of expected running times for these algorithms are bounded by  $n^2$  and  $n \log n$ . For a more precise estimation, we divided the average running times by different polynomial expressions to find the best one that results in a constant value. The best polynomial estimation order found for each of the algorithms are as follow, (1+1) EA:  $n \log^{5.2} n$ , RLS-SEMO:  $n \log^{4.5} n$ , GSEMO:  $n \log^{4.5} n$ , SEMO:  $n \log^4 n$  and RLS :  $n \log^{3.5} n$ . The result of divisions are presented in Figures 2b,2c,2d,2e and 2f in which we are interested to see a straight line to show that estimated expected time is different from the real expected time by a constant factor.

## 5 Conclusion

Evolutionary algorithms are known as general problem solvers which can provide with good qualities in many cases. While these algorithms have been thoroughly studied in solving the linear function, their performance on the non-linear problems is not clear. In this paper, we theoretically study the performance of three base-line EAs on the packing while travelling problem also known as a non-linear knapsack problem. We prove that RLS and GSEMO find the optimal solution in  $O(n^3)$  expected time for instances with correlated weights and profits. Moreover, we show that (1+1)EA finds the optimal solution for instances with uniform weights in  $O(n^2 \log(\max\{n, p_{max}\}))$  where  $p_{max}$  is the highest profit of the given items. We also empirically investigate these algorithms and show that the  $O(n^2)$  is a better upper bound for the expected running time.



## References

- [BMB13] Mohammad Reza Bonyadi, Zbigniew Michalewicz, and Luigi Barone. The travelling thief problem: The first step in the transition from theoretical problems to realistic problems. In *Proceedings of the IEEE Congress on Evolutionary Computation, CEC 2013, Cancun, Mexico, June 20-23, 2013*, pages 1037–1044. IEEE, 2013.
- [BMWN19] Mohammad Reza Bonyadi, Zbigniew Michalewicz, Markus Wagner, and Frank Neumann. *Evolutionary Computation for Multicomponent Problems: Opportunities and Future Directions*, pages 13–30. Springer International Publishing, Cham, 2019.
- [DJW02] Stefan Droste, Thomas Jansen, and Ingo Wegener. On the analysis of the  $(1+1)$  evolutionary algorithm. *Theoretical Computer Science*, 276(1-2):51–81, 2002.
- [FKL<sup>+</sup>18] Tobias Friedrich, Timo Kötzing, JA Gregor Lagodzinski, Frank Neumann, and Martin Schirneck. Analysis of the  $(1+1)$  ea on subclasses of linear functions under uniform and linear constraints. *Theoretical Computer Science*, 2018.
- [Jan13] Thomas Jansen. *Analyzing Evolutionary Algorithms - The Computer Science Perspective*. Natural Computing Series. Springer, 2013.
- [NS18] Frank Neumann and Andrew M. Sutton. Runtime analysis of evolutionary algorithms for the knapsack problem with favorably correlated weights. In Anne Auger, Carlos M. Fonseca, Nuno Lourenço, Penousal Machado, Luís Paquete, and Darrell Whitley, editors, *Parallel Problem Solving from Nature - PPSN XV - 15th International Conference, Coimbra, Portugal, September 8-12, 2018, Proceedings, Part II*, volume 11102 of *Lecture Notes in Computer Science*, pages 141–152. Springer, 2018.
- [NW10] Frank Neumann and Carsten Witt. *Bioinspired computation in combinatorial optimization: Algorithms and their computational complexity*. Springer Science & Business Media, 2010.
- [PN15] Sergey Polyakovskiy and Frank Neumann. Packing while traveling: Mixed integer programming for a class of nonlinear knapsack problems. In Laurent Michel, editor, *Integration of AI and OR Techniques in Constraint Programming - 12th International Conference, CPAIOR 2015, Barcelona, Spain, May 18-22, 2015, Proceedings*, volume 9075 of *Lecture Notes in Computer Science*, pages 332–346. Springer, 2015.
- [PN17] Sergey Polyakovskiy and Frank Neumann. The packing while traveling problem. *European Journal of Operational Research*, 258(2):424–439, 2017.

- [Pop14] Elena Popovici. Anne auger and benjamin doerr (eds): Theory of randomized search heuristics: foundations and recent developments - world scientific (2011), 359 pp. *Genetic Programming and Evolvable Machines*, 15(1):111–112, 2014.
- [WPN16] Junhua Wu, Sergey Polyakovskiy, and Frank Neumann. On the impact of the renting rate for the unconstrained nonlinear knapsack problem. In Tobias Friedrich, Frank Neumann, and Andrew M. Sutton, editors, *Proceedings of the 2016 on Genetic and Evolutionary Computation Conference, Denver, CO, USA, July 20 - 24, 2016*, pages 413–419. ACM, 2016.