Parameterized Analysis of Multi-objective Evolutionary Algorithms and the Weighted Vertex Cover Problem

Mojgan Pourhassan[1], Feng Shi[1,2], and Frank Neumann[1]

[1] Optimisation and Logistics, School of Computer Science, The University of Adelaide, Australia
[2] School of Information Science and Engineering, Central South University, Changsha, P.R.China

Abstract. A rigorous runtime analysis of evolutionary multi-objective optimization for the classical vertex cover problem in the context of parameterized complexity analysis has been presented by Kratsch and Neumann [1]. In this paper, we extend the analysis to the weighted vertex cover problem and provide a fixed parameter evolutionary algorithm with respect to OPT, the cost of the optimal solution for the problem. Moreover, using a diversity mechanism, we present a multi-objective evolutionary algorithm that finds a $2-$approximation in expected polynomial time.

1 Introduction

The area of runtime analysis has provided many rigorous new insights into the working behaviour of bio-inspired computing methods such as evolutionary algorithms and ant colony optimization [2–4]. In recent years, the parameterized analysis of bio-inspired computing has gained additional interest [1, 5, 6]. Here the runtime of bio-inspired computing is studied in dependence of the input size and additional parameters such as the solution size and/or other structural parameters of the given input.

One of the classical problems that has been studied extensively in the area of runtime analysis is the classical NP-hard vertex cover problem. Here, an undirected graph is given and the goal is to find a minimum set of nodes V' such that each edge has at least one endpoint in V'. Friedrich et al. [7] have shown that the single-objective evolutionary algorithm (1+1) EA can not achieve a better than trivial approximation ratio in expected polynomial time. Furthermore, they have shown that a multi-objective approach using Global SEMO gives a factor $O(\log n)$ approximation for the wider classes of set cover problems in expected polynomial time. Further investigations regarding the approximation behaviour of evolutionary algorithms for the vertex cover problem have been carried out in [8, 9]. Edge-based representations in connection with different fitness functions have been investigated in [10, 11] according to their approximation behaviour in the static and dynamic setting. Kratsch and Neumann [1] have studied evolutionary algorithms and the vertex cover problem in the context of parameterized complexity. They have shown that Global SEMO, with a problem specific mutation operator is a fixed parameter evolutionary algorithm for this problem and finds $2-$approximations in expected polynomial time. Kratsch and Neumann [1] have also introduced an alternative mutation operator and have proved that Global SEMO using this mutation operator finds a $(1+\varepsilon)-$approximation in expected time $O(n^2 \log n + OPT \cdot n^2 + n \cdot 4^{(1-\varepsilon)OPT})$. Jansen et al. [10] have shown that a 2-approximation can also be obtained by using an edge-based representation in the (1+1) EA combined with a fitness function formulation based on matchings.

To our knowledge all investigations so far in the area of runtime analysis consider the (unweighted) vertex cover problem. In this paper, we consider the weighted vertex cover problem where in addition weights on the nodes are given and the goal is to find a vertex cover of minimum weight. We extend the investigations carried out in [1] to the weighted minimum vertex cover problem. In [1], multi-objective models in combination with a simple multi-objective evolutionary algorithm called Global SEMO are investigated. One key argument for the results presented for the (unweighted) vertex cover problem is that the population size is always upper bounded by $n+1$. This argument does not hold in the weighted case. Therefore, we study how a variant of Global SEMO using an appropriate diversity mechanism is able to deal with the weighted case.

Our focus is on finding good approximations of an optimal solution. We analyse the time complexity with respect to n, W_{max}, and OPT, which denote the number of vertices, the maximum weight in the input graph, and the cost of the optimal solution respectively. We first study the expected time of finding a solution with expected approximation ratio $(1+\varepsilon)$ for this problem by Global SEMO with alternative mutation operator. Afterwards, we consider DEMO, a variant of Global SEMO, which incorporates ε-dominance [12] as diversity mechanism. We show that DEMO using standard mutation finds a 2-approximation in expected polynomial time.

The outline of the paper is as follows. In Section 2, the problem definition is presented as well as the classical Global SEMO algorithm and DEMO algorithm. Runtime analysis for finding a $(1+\varepsilon)-$approximation by Global SEMO is presented in Section 3. Section 4 includes the analysis that shows DEMO can find $2-$approximations of the optimum in expected polynomial time. At the end, in Section 5 we summarize and conclude.

2 Preliminaries

We consider the weighted vertex cover problem defined as follows. Given a graph $G = (V, E)$ with vertex set $V = \{v_1, \dots, v_n\}$ and edge set $E = \{e_1, \dots, e_m\}$, and a positive weight function $w : V \to \mathbb{N}^+$ on the vertices, the goal is to find a subset of nodes, $V_C \subseteq V$, that covers all edges and has minimum weight, i.e. $\forall e \in E, e \cap V_C \neq \emptyset$ and $\sum_{v \in V_C} w(v)$ is minimized. We consider the standard node-based approach, i.e. the search space is $\{0,1\}^n$ and for a solution $x = (x_1, \dots, x_n)$ the node v_i is chosen iff $x_i = 1$. The Integer Linear Programming (ILP) formulation of this problem is:

$$\begin{aligned} & \min \sum_{i=1}^{n} w(v_i) \cdot x_i \\ s.t. \quad & x_i + x_j \geq 1 \quad \forall (i,j) \in E \\ & x_i \in \{0,1\} \end{aligned}$$

By relaxing the constraint $x_i \in \{0,1\}$ to $x_i \in [0,1]$, the linear program formulation of Fractional Weighted Vertex Cover is obtained.

1. Choose $x \in \{0,1\}^n$ uniformly at random and set $P = \{x\}$;
2. while (not termination condition)
   - Choose $x \in P$ uniformly at random and set $x' = x$;
   - Let $E(x) \subseteq E$ denote the set of edges that are not covered by x and $S(x) \subseteq \{1, \ldots, n\}$ the vertices being incident on the edges in $E(x)$.
   - Choose $b \in \{0,1\}$ uniform at random.
   - If $b = 0$ flip each bit of x' independently with probability $1/n$.
   - Otherwise flip each bit of $S(x')$ independently with probability $1/2$ and each other bit independently with probability $1/n$.
   - If there is no $y \in P$ with $f(y) \leq f(x')$ then delete all $z \in P$ with $f(x') \leq f(z)$ from P and add x' to P.

Algorithm 1: Global SEMO

We consider primarily multi-objective approaches for the weighted vertex cover problem. Given a multi-objective fitness function $f = (f_1, \ldots, f_d)\colon S \to \mathbb{R}$ where all d objectives should be minimized, we have $f(x) \leq f(y)$ iff $f_i(x) \leq f_i(y)$, $1 \leq i \leq d$. We say that x (weakly) dominates y iff $f(x) \leq f(y)$.

Let $G(x)$ be the graph obtained from G by removing all nodes chosen by x and the corresponding covered edges. Formally, we have $G(x) = (V(x), E(x))$ where $V(x) = V \setminus \{v_i \mid x_i = 1\}$ and $E(x) = E \setminus \{e \mid e \cap (V \setminus V(x)) \neq \emptyset\}$. Kratsch and Neumann [1] investigated a multi-objective baseline algorithm called Global SEMO using the LP-value for $G(x)$ as one of the fitness values for the (unweighted) minimum vertex cover problem. In order to expand the analysis on behaviour of multi-objective evolutionary algorithms to the Weighted Vertex Cover problem, we modify the fitness function that was used in Global SEMO in [1], to match the weighted version of the problem. We investigate the multi-objective fitness function $f(x) = (Cost(x), LP(x))$, where

- $Cost(x) = \sum_{i=1}^{n} w(v_i)x_i$ is the sum of weights of selected vertices
- $LP(x)$ is the value of an optimal solution of the LP for $G(x)$.

We investigate Global SEMO with alternative mutation operator (Algorithm 1) introduced in [1]. Here, the nodes that are adjacent to uncovered edges are mutated with probability $1/2$ in some steps. In the fitness function used in Global SEMO, both $Cost(x)$ and $LP(x)$ can be exponential with respect to the input size; therefore, we need to deal with exponentially large number of solutions, even if we only keep the Pareto front.

One approach for dealing with this problem is using the concept of $\varepsilon-$dominance [12]. The concept of $\varepsilon-$dominance has previously been proved to be useful for coping with exponentially large Pareto fronts in some problems [13, 14]. Having two objective vectors $u = (u_1, \cdots, u_m)$ and $v = (v_1, \cdots, v_m)$, u $\varepsilon-$dominates v, denoted by $u \preceq_\varepsilon v$, if for all $i \in \{1, \cdots, m\}$ we have $(1+\varepsilon)u_i \leq v_i$. In this approach, the objective space is partitioned into a polynomial number of boxes in which all solutions $\varepsilon-$dominate each other, and at most one solution from each box is kept in the population.

Motivated by this approach, DEMO (Diversity Evolutionary Multi-objective Optimizer) has been investigated in [14, 15]. In Section 4, we analyze DEMO (see Al-

1. Choose $x \in \{0,1\}^n$ uniformly at random and set $P = \{x\}$;
2. while (not termination condition)
   – Choose $x \in P$ uniformly at random and set $x' = x$;
   – Flip each bit of x' independently with probability $1/n$.
   – If there is a $y \in P$ where $(f(y) \leq f(x') \wedge f(y) \neq f(x'))$ or $(b(y) = b(x') \wedge Cost(y) + 2 \cdot LP(y) \leq Cost(x') + 2 \cdot LP(x'))$ then keep P unchanged and go to 4;
   – Otherwise delete all $z \in P$ with $f(x') \leq f(z) \vee b(z) = b(x')$ from P and add x' to P.

Algorithm 2: DEMO

gorithm 2) in which only one non-dominated solution can be kept in the population for each box based on a predefined criteria. In our setting, among two solutions x and y from one box, y is kept in P and x is discarded if $Cost(y) + 2 \cdot LP(y) \leq Cost(x) + 2 \cdot LP(x)$. To implement the concept of $\varepsilon-$dominance in DEMO, we use the parameter $\delta = \frac{1}{2n}$ and define the boxing function $b : \{0,1\}^n \to \mathbb{N}^2$ as $b_1(x) = \lceil \log_{1+\delta}(1 + Cost(x)) \rceil$ and $b_2(x) = \lceil \log_{1+\delta}(1 + LP(x)) \rceil$.

Analysing the runtime of our evolutionary algorithms, we are interested in the expected number of rounds of the while loop until a solution of desired quality has been obtained. We call this the expected time until the considered algorithm has achieved its desired goal.

3 Analysis of Global SEMO

In this section, we analyse the expected time of Global SEMO (Algorithm 1) to find a (1+ε)-approximation. Before we present our analysis for Global SEMO, we state some basic properties of the solutions in our multi-objective model. The following theorem shown by Balinski [16] states that all basic feasible solutions of the fractional vertex cover, which are the extremal points or the corner solutions of the polyhedron that forms the feasible space, are half-integral.

Theorem 1. *Each basic feasible solution x of the relaxed Vertex Cover ILP is half-integral, i.e., $x \in \{0, 1/2, 1\}^n$. [16]*

As a result, there always exists a half integral optimal LP solution for a vertex cover problem. This result and the following lemmata are used in the analysis of Theorem 2 which presents the main approximation result for Global SEMO. The proof of Lemma 3 can be found in [17].

Lemma 1. *For any $x \in \{0,1\}^n$, $LP(x) \leq LP(0^n) \leq OPT$.*

Proof. Let y be the LP solution of $LP(0^n)$. Also, for any solution x, let $G(x)$ be the graph obtained from G by removing all vertices chosen by x and their edges. The solution 0^n contains no vertices; therefore, y is the optimal fractional vertex cover for all edges of the input graph. Thus, for any solution x, y is a (possibly non-optimal) fractional cover for $G(x)$; therefore, $LP(x) \leq LP(0^n)$. Moreover, we have $LP(0^n) \leq OPT$ as $LP(0^n)$ is the optimal value of the LP relaxation. □

Lemma 2. *Let $x = \{x_1, \cdots, x_n\}, x_i \in \{0, 1\}$ be a solution and $y = \{y_1, \cdots, y_n\}$, $y_i \in [0, 1]$ be a fractional solution for $G(x)$. If there is a vertex v_i where $y_i \geq \frac{1}{2}$, mutating x_i from 0 to 1 results in a solution x' for which $LP(x') \leq LP(x) - y_i \cdot w(v_i) \leq LP(x) - \frac{1}{2}w(v_i)$.*

Proof. The graph $G(x')$ is the same as $G(x)$ excluding the edges connected to v_i. Therefore, the solution $y' = \{y_1, \cdots, y_{i-1}, 0, y_{i+1}, y_n\}$ is a fractional vertex cover for $G(x')$ and has a cost of $LP(x) - y_i w(v_i)$. The cost of the optimal fractional vertex cover of $G(x')$ is at most as great as the cost of y'; thus $LP(x') \leq LP(x) - y_i \cdot w(v_i) \leq LP(x) - \frac{1}{2}w(v_i)$. □

Lemma 3. *The population size of Global SEMO (Algorithm 1) is upper bounded by $2 \cdot OPT + 1$ and the search point 0^n is included in the population of Global SEMO, in expected time $O\left(OPT \cdot n(\log W_{max} + \log n)\right)$.*

Lemma 4. *A solution x fulfilling the two properties*
1. *$LP(x) = LP(0^n) - Cost(x)$ and*
2. *there is an optimal solution of the LP for G(x) which assigns 1/2 to each non-isolated vertex of $G(x)$*

is included in the population of Global SEMO in expected time $O(OPT \cdot n(\log W_{max} + \log n + OPT))$.

Proof. The search point 0^n which satisfies property 1 is included in the population in expected time of $O(OPT \cdot n(\log W_{max} + \log n))$, due to Lemma 3. Let $P' \subseteq P$ be a set of solutions such that for each solution $x \in P'$, $LP(x) + Cost(x) = LP(0^n)$. Let $x_{min} \in P'$ be a solution such that $LP(x_{min}) = min_{x \in P'} LP(x)$.

If the optimal fractional vertex cover for $G(x_{min})$ assigns 1/2 to each non-isolated vertex of $G(x_{min})$, then the conditions of the lemma hold. Otherwise, it assigns 1 to some non-isolated vertex, say v. The probability that the algorithm selects x_{min} and flips the bit corresponding to v, is $\Omega(\frac{1}{OPT \cdot n})$, because the population size is $O(OPT)$ (Lemma 3). Let x_{new} be the new solution. We have $Cost(x_{new}) = Cost(x_{min}) + w(v)$, and by Lemma 2, $LP_w(x_{new}) \leq LP_w(x_{min}) - w(v)$. This implies that $LP(x_{new}) + Cost(x_{new}) = LP(0^n)$; hence, x_{new} is a Pareto Optimal solution and is added to the population P.

Since $LP_w(x_{min}) \leq OPT$ (Lemma 1) and the weights are at least 1, assuming that we already have the solution 0^n in the population, by means of the method of fitness based partitions, we find the expected time of finding a solution that fulfils the properties given above as $O(OPT^2 \cdot n)$. Since the search point 0^n is included in expected time $O(OPT \cdot n(\log W_{max} + \log n))$, the expected time that a solution fulfilling the properties given above is included in P is $O(OPT \cdot n(\log W_{max} + \log n + OPT))$. □

Theorem 2. *The expected time until Global SEMO has obtained a solution that has expected approximation ratio $(1+\varepsilon)$ is $O(OPT \cdot 2^{\min\{n, 2(1-\varepsilon)OPT\}} + OPT \cdot n(\log W_{max} + \log n + OPT))$.*

Proof. By Lemma 4, a solution x that satisfies the two properties given in Lemma 4 is included in the population in expected time of $O(OPT \cdot n(\log W_{max} + \log n + OPT))$. For a set of nodes, X', we define $Cost(X') = \sum_{v \in X'} w(v)$. Let X be the vertex set of graph $G(x)$. Also, let $S \subseteq X$ be a vertex cover of $G(x)$ with the minimum weight over all vertex covers of $G(x)$, and T be the set containing all non-isolated vertices in $X \setminus S$. Note that all vertices in $X \setminus (S \cup T)$ are isolated vertices in $G(x)$. Due to property 2 of Lemma 4, $\frac{1}{2}Cost(S) + \frac{1}{2}Cost(T) = LP(x) \leq Cost(S)$; therefore, $Cost(T) \leq Cost(S)$. Let $OPT' = OPT - Cost(x)$. Observe that $OPT' = Cost(S)$.

Let $s_1, \ldots, s_{|S|}$ be a numbering of the vertices in S such that $w(s_i) \leq w(s_{i+1})$, for all $1 \leq i \leq |S| - 1$. And let $t_1, \ldots, t_{|T|}$ be a numbering of the vertices in T such that $w(t_i) \geq w(t_{i+1})$, for all $1 \leq i \leq |T| - 1$. Let $S_1 = \{s_1, s_2, \ldots, s_\rho\}$, where $\rho = min\{|S|, (1-\varepsilon) \cdot OPT'\}$, and $T_1 = \{t_1, t_2, \ldots, t_\eta\}$, where $\eta = min\{|T|, (1 - \varepsilon) \cdot OPT'\}$.

With probability $\Omega(\frac{1}{OPT})$, the algorithm Global SEMO selects the solution x, and sets $b = 1$. With $b = 1$, the probability that the bits corresponding to all vertices of S_1 are flipped, is $\Omega((\frac{1}{2})^\rho)$, and the probability that none of the bits corresponding to the vertices of T_1 are flipped is $\Omega((\frac{1}{2})^\eta)$. Also, the bits corresponding to the isolated vertices of $G(x)$ are flipped with probability $\frac{1}{n}$; hence, the probability that none of them flips is $\Omega(1)$. As a result, with probability $\Omega(\frac{1}{OPT} \cdot (\frac{1}{2})^{\rho+\eta})$, solution x is selected, the vertices of S_1 are included, and the vertices of T_1 and isolated vertices are not included in the new solution x'. Since $\rho + \eta \leq 2(1-\varepsilon) \cdot OPT' \leq 2(1-\varepsilon) \cdot OPT$, and also $\rho + \eta \leq n$; the expected time until solution x' is found after reaching solution x, is $O(OPT \cdot 2^{\min\{n, 2(1-\varepsilon)OPT\}})$.

Note that the bits corresponding to vertices of $S_2 = S \setminus S_1$ and $T_2 = T \setminus T_1$, are arbitrarily flipped in solution x' with probability $1/2$ by the Alternative Mutation Operator. Here we show that for the expected cost and the LP value of x', the following constraint holds: $E[Cost(x')] + 2 \cdot LP(x') \leq (1 + \varepsilon) \cdot OPT$.

Let $S' \subseteq S$ and $T' \subseteq T$ denote the subset of vertices of S and T that are actually included in the new solution x' respectively. In the following, we show that for the expected values of $Cost(S')$ and $Cost(T')$, we have:

$$E\left[Cost(S')\right] \geq (1-\varepsilon) \cdot OPT' + E\left[Cost(T')\right] \tag{1}$$

Since the bits corresponding to the vertices of S_2 and T_2 are flipped with probability 1/2, for the expected values of $Cost(S')$ and $Cost(T')$ we have:

$$\begin{aligned} E\left[Cost(S')\right] &= Cost(S_1) + \frac{Cost(S_2)}{2} = Cost(S_1) + \frac{Cost(S) - Cost(S_1)}{2} \\ &= 1/2Cost(S) + 1/2Cost(S_1) \end{aligned}$$

and $E\left[Cost(T')\right] = 1/2Cost(T_2)$.

If $\rho = |S|$, then $S_1 = S$ and $Cost(S_1) = Cost(S) = OPT'$. If $\rho = (1-\varepsilon) \cdot OPT'$, we have $Cost(S_1) \geq (1-\varepsilon) \cdot OPT'$, since each vertex has a weight of at least 1. Using $Cost(S) = OPT'$ and the inequality above, we have

$$E\left[Cost(S')\right] \geq (1-\varepsilon) \cdot OPT' + \frac{\varepsilon \cdot OPT'}{2}$$

We divide the analysis into two cases based on the relation between η and $|T|$.

Case (I). $\eta = |T|$. Then $T_2 = T' = \emptyset$. Thus, $E\left[Cost(T')\right] = 0$ and Inequality (1) holds true.

Case (II). $\eta = (1-\varepsilon)\cdot OPT' < |T|$. Since $w(t_i) \geq w(t_{i+1})$ for $1 \leq i \leq |T|-1$ and $Cost(T) \leq Cost(S) = OPT'$, we have

$$Cost(T_2) \leq \frac{|T|-\eta}{|T|}Cost(T) \leq \frac{OPT'-(1-\varepsilon)\cdot OPT'}{OPT'}Cost(T)$$
$$\leq \varepsilon Cost(S) = \varepsilon \cdot OPT'$$

Thus for the expected value of $Cost(T')$, we have $E\left[Cost(T')\right] = \frac{1}{2}Cost(T_2) \leq \frac{\varepsilon \cdot OPT'}{2}$.

Summarizing above analysis, we can get that the Inequality (1) holds. Using this inequality, we prove that in expectation, the new solution x' satisfies the inequality $Cost(x') + 2\cdot LP(x') \leq (1+\varepsilon)\cdot OPT$:

$$\begin{aligned}
&E\left[Cost(x')\right] + 2\cdot LP(x') = Cost(x) + E\left[Cost(S')\right] + E\left[Cost(T')\right] + 2\cdot LP(x')\\
&\leq Cost(x) + E\left[Cost(S')\right] + E\left[Cost(S')\right] - (1-\varepsilon)\cdot OPT' + 2\cdot LP(x')\\
&\leq Cost(x) + 2E\left[Cost(S')\right] - (1-\varepsilon)\cdot OPT' + 2\cdot\left(OPT' - E\left[Cost(S')\right]\right)\\
&= Cost(x) + (1+\varepsilon)\cdot OPT' = Cost(x) + (1+\varepsilon)\cdot(OPT - Cost(x))\\
&\leq (1+\varepsilon)\cdot OPT.
\end{aligned}$$

Now we analyze whether the new solution x' could be included in the population P. If x' could not be included in P, then there is a solution x'' dominating x, i.e., $LP(x'') \leq LP(x')$ and $Cost(x'') \leq Cost(x')$. This implies $Cost(x'') + 2\cdot LP(x'') < Cost(x') + 2\cdot LP(x') \leq (1+\varepsilon)\cdot OPT$. Therefore, after having a solution that fulfils the properties of Lemma 4 in P, in expected time $O(OPT \cdot 2^{\min\{n, 2(1-\varepsilon)OPT\}})$, the population would contain a solution y such that $Cost(y) + 2\cdot LP(y) \leq (1+\varepsilon)\cdot OPT$.

Let P' contain all solutions $x \in P$ such that $Cost(x) + 2\cdot LP(x) \leq (1+\varepsilon)\cdot OPT$, and let $x_{\min}$ be the one that minimizes LP. Let $y = \{y_1, \cdots, y_n\}$ be a basic LP solution for $G(x_{\min})$. According to Theorem 1, y is a half-integral solution.

Let Δ^t be the improvement that happens on the minimum LP value in p' at time step t. Also let k be the number of nodes that are assigned at least $\frac{1}{2}$ by y. Flipping only one of these nodes by the algorithm happens with probability at least $\frac{k}{e\cdot n}$. According to Lemma 2, flipping one of these nodes, v_i, results in a solution x' with $LP(x') \leq LP(x_{\min}) - y_i \cdot w(v_i) \leq LP(x_{\min}) - \frac{1}{2}\cdot w(v_i)$. Observe that the constraint of $Cost(x') + 2\cdot LP(x') \leq 2\cdot OPT$ holds for solution x'. Therefore, $\Delta^t \geq y_i \cdot w(v_i)$, which is in expectation at least $\frac{LP(x_{\min})}{k}$ due to definition of $LP(x_{\min})$.

Moreover, at each step, the probability that $x_{\min}$ is selected and only one of the k bits defined above flips is $\frac{k}{(2\cdot OPT+1)\cdot e\cdot n}$, As a result we have:

$$E[\Delta^t \mid x_{\min}] \geq \frac{k}{(2\cdot OPT+1)\cdot e\cdot n}\cdot\frac{LP(x_{\min})}{k} = \frac{LP(x_{\min})}{e\cdot n\cdot(2\cdot OPT+1)}$$

According to Lemma 1 for any solution x, we have $LP(x) \leq OPT$. We also know that for any solution x which is not a complete cover, $LP(x) \geq 1$, because the weights are positive integers. Using the method of Multiplicative Drift Analysis [18] with $s_0 \leq OPT$ and $s_{min} \geq 1$, we get the expected time $O(OPT \cdot n \log OPT)$ to find a solution z for which $LP(z) = 0$ and $Cost(z) + 2 \cdot LP(z) \leq (1 + \varepsilon) \cdot OPT$.

Overall, the expected number of iterations of Global SEMO, for finding a $(1 + \varepsilon)$-approximate weighted vertex cover, is bounded by $O(OPT \cdot 2^{\min\{n, 2(1-\varepsilon)OPT\}} + OPT \cdot n(\log W_{max} + \log n + OPT))$. □

4 Analysis of DEMO

In this section, we analyse the other evolutionary algorithm, DEMO (Algorithm 2), that uses some diversity handling mechanisms for dealing with exponentially large population sizes. We are making use of the following lemma whose proof can be found in [17].

Lemma 5. *The population size of DEMO is upper bounded by* $O\left(n \cdot (\log n + \log W_{max})\right)$ *and the search point* 0^n *is included in the population in expected time of* $O(n^3(\log n + \log W_{max})^2)$.

Lemma 6. *Let* $x \in P$ *be a search point such that* $Cost(x) + 2 \cdot LP(x) \leq 2 \cdot OPT$ *and* $b_2(x) > 0$*. There exists a 1-bit flip leading to a search point* x' *with* $Cost(x') + 2 \cdot LP(x') \leq 2 \cdot OPT$ *and* $b_2(x') < b_2(x)$.

Proof. Let $y = \{y_1 \cdots y_n\}$ be a basic half integral LP solution for $G(x)$. Since $b_2(x) = LP(x) \neq 0$, there must be at least one uncovered edge; hence, at least one vertex v_i has a $y_i \geq \frac{1}{2}$ in LP solution y. Consider v_j the vertex that maximizes $y_i w(v_i)$ among vertices v_i, $1 \leq i \leq n$. Also, let x' be a solution obtained by adding v_j to x. Since solutions x and x' are only different in one vertex, v_j, we have $Cost(x') = Cost(x) + w(v_j)$. Moreover, according to Lemma 2, $LP(x') \leq LP(x) - \frac{1}{2} \cdot w(v_j)$. Therefore,

$$Cost(x') + 2 \cdot LP(x') \leq Cost(x) + w(v_j) + 2\left(LP(x) - \frac{w(v_j)}{2}\right)$$

$$\leq Cost(x) + 2 \cdot LP(x) \leq 2 \cdot OPT$$

which means solution x' fulfils the mentioned constraint. If $LP(x) = W$, then $y_j w(v_j) \geq \frac{W}{n}$, because n is an upper bound on the number of vertices selected by the LP solution. As a result, using Lemma 2, we get $LP(x') \leq W \cdot (1 - \frac{1}{n})$. Therefore, we have:

$$\begin{aligned}(1+\delta)\left(1 + LP(x')\right) &\leq 1 + \delta + W\left(1 - \frac{1}{n}\right)(1+\delta) \\ &\leq 1 + \delta + W + W(\delta - \frac{1}{n} - \frac{\delta}{n}) \\ &\leq 1 + W + W(2\delta - \frac{1}{n} - \frac{\delta}{n}) \leq 1 + W\end{aligned}$$

which implies $1 + \log_{1+\delta}(1 + LP(x')) \leq \log_{1+\delta}(1 + W)$. As a result, $b_2(x') < b_2(x)$ holds for x', which is obtained by performing a 1-bit flip on x, and the lemma is proved. □

Theorem 3. *The expected time until DEMO constructs a 2-approximate vertex cover is* $O\left(n^3 \cdot (\log n + \log W_{max})^2\right)$.

Proof. Consider solution $x \in P$ that minimizes $b_2(x)$ under the constraint that $Cost(x) + 2 \cdot LP(x) \leq 2 \cdot OPT$. Note that 0^n fulfils this constraint and according to Lemma 5, the solution 0^n will be included in P in time $O\left(n^3(\log n + \log W_{max})^2\right)$.

If $b_2(x) = 0$ then x covers all edges and by selection of x we have $Cost(x) \leq 2 \cdot OPT$, which means that x is a $2-$approximation.

In case $b_2(x) \neq 0$, according to Lemma 6 there is a one-bit flip on x that results in a new solution x' for which $b_2(x') < b_2(x)$, while the mentioned constraint also holds for it. Since the population size is $O\left(n \cdot (\log n + \log W_{max})\right)$ (Lemma 5), this 1-bit flip happens with a probability of $\Omega\left(n^{-2} \cdot (\log n + \log W_{max})^{-1}\right)$ and x' is obtained in expected time of $O(n^3 \cdot (\log n + \log W_{max})^2)$. This new solution will be added to P because a solution y with $Cost(y) + 2 \cdot LP(y) > 2 \cdot OPT$ can not dominate x' with $Cost(x') + 2 \cdot LP(x') \leq 2 \cdot OPT$, and x' has the minimum value of b_2 among solution that fulfil the constraint. Moreover, if there already is a solution, x_{prev}, in the same box as x', it will be replaced by x' because $Cost(x_{prev}) + 2 \cdot LP(x_{prev}) > 2 \cdot OPT$; otherwise, it would have been selected as x.

There are at most $A = 1 + \lceil \frac{\log n + \log W_{max}}{\log(1+\delta)} \rceil$ different values for b_2 in the objective space, and since $\delta = \frac{1}{2n}$, $A = O(n \cdot (\log n + \log W_{max}))$. Therefore, the expected time until a solution x'' is found so that $b_2(x'') = 0$ and $Cost(x'') + 2 \cdot LP(x'') \leq 2 \cdot OPT$, is at most $O(n^3 \cdot (\log n + \log W_{max})^2)$. □

5 Conclusion

The minimum vertex cover problem is one of the classical NP-hard combinatorial optimization problems. In this paper, we have generalized previous results of Kratsch and Neumann [1] for the unweighted minimum vertex cover problem to the weighted case where in addition weights on the nodes are given. We have studied the expected time required by Global SEMO to find a $(1 + \varepsilon)$-approximation. Furthermore, our investigations show that the algorithm DEMO using the ε-dominance approach reaches a 2-approximation in expected polynomial time.

Acknowledgements

This research has been supported by Australian Research Council grants DP140103400 and DP160102401.

Bibliography

[1] Kratsch, S., Neumann, F.: Fixed-parameter evolutionary algorithms and the vertex cover problem. Algorithmica **65**(4) (2013) 754–771

[2] Neumann, F., Witt, C.: Bioinspired Computation in Combinatorial Optimization:Algorithms and Their Computational Complexity. 1st edn. Springer-Verlag New York, Inc., New York, NY, USA (2010)

[3] Auger, A., Doerr, B.: Theory of Randomized Search Heuristics: Foundations and Recent Developments. World Scientific Publishing Co., Inc. (2011)
[4] Jansen, T.: Analyzing Evolutionary Algorithms - The Computer Science Perspective. Natural Computing Series. Springer (2013)
[5] Sutton, A.M., Neumann, F.: A parameterized runtime analysis of simple evolutionary algorithms for makespan scheduling. In: Proceedings of the Twelfth Conference on Parallel Problem Solving from Nature (PPSN 2012), Springer (2012) 52–61
[6] Sutton, A.M., Neumann, F., Nallaperuma, S.: Parameterized runtime analyses of evolutionary algorithms for the planar euclidean traveling salesperson problem. Evolutionary Computation **22**(4) (2014) 595–628
[7] Friedrich, T., Hebbinghaus, N., Neumann, F., He, J., Witt, C.: Approximating covering problems by randomized search heuristics using multi-objective models. In: Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation. GECCO '07, New York, NY, USA, ACM (2007) 797–804
[8] Friedrich, T., He, J., Hebbinghaus, N., Neumann, F., Witt, C.: Analyses of simple hybrid algorithms for the vertex cover problem. Evolutionary Computation **17**(1) (2009) 3–19
[9] Oliveto, P.S., He, J., Yao, X.: Analysis of the (1+1) -ea for finding approximate solutions to vertex cover problems. IEEE Trans. Evolutionary Computation **13**(5) (2009) 1006–1029
[10] Jansen, T., Oliveto, P.S., Zarges, C.: Approximating vertex cover using edge-based representations. In Neumann, F., Jong, K.A.D., eds.: Foundations of Genetic Algorithms XII, FOGA '13, Adelaide, SA, Australia, January 16-20, 2013, ACM (2013) 87–96
[11] Pourhassan, M., Gao, W., Neumann, F.: Maintaining 2-approximations for the dynamic vertex cover problem using evolutionary algorithms. In Silva, S., Esparcia-Alcázar, A.I., eds.: Proceedings of the Genetic and Evolutionary Computation Conference, GECCO 2015, Madrid, Spain, July 11-15, 2015, ACM (2015) 903–910
[12] Laumanns, M., Thiele, L., Deb, K., Zitzler, E.: Combining convergence and diversity in evolutionary multiobjective optimization. Evol. Comput. **10**(3) (September 2002) 263–282
[13] Horoba, C., Neumann, F.: Benefits and drawbacks for the use of ϵ-dominance in evolutionary multi-objective optimization. In: In Proc. of GECCO 2008. (2008)
[14] Neumann, F., Reichel, J., Skutella, M.: Computing minimum cuts by randomized search heuristics. Algorithmica **59**(3) (2011) 323–342
[15] Neumann, F., Reichel, J.: Approximating minimum multicuts by evolutionary multi-objective algorithms. In Rudolph, G., Jansen, T., Lucas, S.M., Poloni, C., Beume, N., eds.: Parallel Problem Solving from Nature - PPSN X, 10th International Conference Dortmund, Germany, September 13-17, 2008, Proceedings. Volume 5199 of Lecture Notes in Computer Science., Springer (2008) 72–81
[16] Balinski, M.: On the maximum matching, minimum covering. In: Proc. Symp. Math. Programming, Princeton University Press (1970) 434–445
[17] Pourhassan, M., Shi, F., Neumann, F.: Parameterized analysis of multi-objective evolutionary algorithms and the weighted vertex cover problem. CoRR **http://arxiv.org/abs/1604.01495** (2016)

[18] Doerr, B., Johannsen, D., Winzen, C.: Multiplicative drift analysis. Algorithmica **64**(4) (2012) 673–697