This is the supplementary material for the paper "Initial Results on Runtime Analysis of $(1+1)$ Evolutionary Algorithm Controlled with Q-learning using Greedy Exploration Strategy on OneMax+ZeroMax Problem", submitted to the PPSN conference. It contains a proof of the $O(N \log N)$ bound on the number of fitness evaluations of the algorithm described in the paper on the OneMax+ZeroMax problem. More precisely, it is shown that the running time is at most $\frac{16e^2}{7} N \ln N + o(N \ln N)$.

1 Excerpts from the Paper

For the sake of convenience, we outline here the necessary results from the paper.

The algorithm considered in the paper is the $(1+1)$ evolutionary algorithm solving the OneMax+ZeroMax problem. The objective selection is done using the EA+RL approach, which, in this paper, uses the greedy Q-learning reinforcement learning algorithm.

A Markov chain that is used to model this algorithm has $2^N \cdot N + 1$ states of the form $\langle B, V \rangle$, where B is the number of one-bits in the current bit vector and V is the set of reinforcement learning states, which have been left at least once (i.e. the reinforcement learning states where OneMax will always be selected). For convenience, we merge all the states with $B = N$, as they all are terminal states for the optimization process.

The mutation operator is used which flips each bit with the probability of $1/N$. The probability $P^{i,j}$ of going from the state that has i one-bits to the state that has j one-bits is:

$$P^{i,j} = \begin{cases} \sum\limits_{k=0}^{\min(N-j,i)} \binom{N-i}{j-i+k} \binom{i}{k} \left(\frac{1}{N}\right)^{j-i+2k} \left(1 - \frac{1}{N}\right)^{N-(j-i+2k)} & \text{if } i < j, \\ \sum\limits_{k=0}^{\min(N-i,j)} \binom{i}{i-j+k} \binom{N-i}{k} \left(\frac{1}{N}\right)^{i-j+2k} \left(1 - \frac{1}{N}\right)^{N-(i-j+2k)} & \text{if } i > j, \\ \sum\limits_{k=0}^{\min(N-i,i)} \binom{N-i}{k} \binom{i}{k} \left(\frac{1}{N}\right)^{2k} \left(1 - \frac{1}{N}\right)^{N-2k} & \text{if } i = j. \end{cases} \quad (1)$$

In a state with the known optimal action ("gray states"), OneMax is selected with the probability of 1.0, and the mutation is accepted only if the number of ones does not decrease. This sums up to the following probabilities:

$$P_G^{i,j} = \begin{cases} P^{i,j} & \text{if } i < j, \\ 1 - \sum_{k=i+1}^{j} P^{i,j} & \text{if } i = j, \\ 0 & \text{if } i > j. \end{cases} \quad (2)$$

In a state with no known optimal action ("white states"), OneMax is selected with the probability of 0.5, as well as ZeroMax. In the case of OneMax, the mutations that do not decrease the number of ones are accepted, and in the case of ZeroMax, the opposite mutations are accepted. This results in the following:

$$P_W^{i,j} = \begin{cases} \frac{1}{2}P^{i,j} & \text{if } i < j, \\ \frac{1}{2}P^{i,j} & \text{if } i > j, \\ 1 - \frac{1}{2}\sum_{k \neq i} P^{i,k} = \frac{1}{2}\left(1 + P^{i,i}\right) & \text{if } i = j. \end{cases} \quad (3)$$

We define $E(B, V)$ as the expectation of the number of fitness evaluations needed to get to the terminal state from the state $\langle B, V \rangle$. It is assumed in this part of the paper that for any set V and any two $B_1 < B_2$ the inequality $E(B_1, V) > E(B_2, V)$ holds (the *non-increasing assumption*).

2 Proof of $O(N \log N)$ Bound Under Non-Increasing Assumption

This section is dedicated to the proof that the expectation of the number of fitness evaluations of the algorithm is $O(N \log N)$ provided the non-increasing assumption is true.

First, we will give lower and upper bounds on $P^{i,j}$, the components of transition probabilities in the Markov chain. Next, we show that if the algorithm never accepts mutations that increase the number of ones by more than one, the required expectation does not decrease. Finally, we simplify the Markov chain and prove the result.

2.1 Lower and Upper Bound on $P^{i,j}$

The following theorem gives a lower and an upper bound on $P^{i,j}$.

Theorem 1. *Assume that $i \neq j$. Let $S^{i,j}$ be the following:*

$$S^{i,j} = \begin{cases} \binom{N-i}{j-i} \left(\frac{1}{N}\right)^{j-i} \left(1 - \frac{1}{N}\right)^{N-(j-i)} & \text{if } i < j, \\ \binom{i}{i-j} \left(\frac{1}{N}\right)^{i-j} \left(1 - \frac{1}{N}\right)^{N-(i-j)} & \text{if } i > j. \end{cases} \quad (4)$$

Then $S^{i,j} \leq P^{i,j} \leq \frac{8}{7}S^{i,j}$.

Proof. The lower bounds are proven easily, since $S^{i,j}$ are the addends for $k = 0$ in (1), and all these addends are positive.

Now we are proving the upper bound. We denote as $S_k^{i,j}$ the k-th addend of the sum in (1) corresponding to $P^{i,j}$. Specifically, $S^{i,j} = S_0^{i,j}$.

Consider the case of $i < j$. The ratio of the k-th addend to the $(k+1)$-th addend is:

$$\frac{S_k^{i,j}}{S_{k+1}^{i,j}} = \frac{\binom{N-i}{j-i+k}\binom{i}{k}\left(\frac{1}{N}\right)^{j-i+2k}\left(1 - \frac{1}{N}\right)^{N-(j-i+2k)}}{\binom{N-i}{j-i+k+1}\binom{i}{k+1}\left(\frac{1}{N}\right)^{j-i+2k+2}\left(1 - \frac{1}{N}\right)^{N-(j-i+2k)-2}} =$$

$$= \frac{(j-i+k+1)(k+1)}{(N-j-k)(i-k)} N^2 \left(1 - \frac{1}{N}\right)^2 = \frac{(j-i+k+1)(k+1)}{(N-j-k)(i-k)}(N-1)^2.$$

When i and j are fixed, this ratio grows as k grows, so

$$\frac{S_k^{i,j}}{S_{k+1}^{i,j}} \ge \frac{S_0^{i,j}}{S_1^{i,j}} = \frac{j-i+1}{(N-j)i}(N-1)^2.$$

When i is fixed, this ratio grows as j grows, so we replace j with its minimum possible value $i+1$ and then minimize the result with $i = \frac{N-1}{2}$:

$$\frac{S_k^{i,j}}{S_{k+1}^{i,j}} \ge \frac{2(N-1)^2}{(N-i-1)i} \ge \frac{2(N-1)^2}{\frac{N-1}{2}\frac{N-1}{2}} = \frac{8(N-1)^2}{(N-1)^2} = 8.$$

This means that $P^{i,j}$ can be bounded by a sum of geometric progression:

$$P^{i,j} = \sum_{k=0}^{\min(N-j,i)} S_k^{i,j} \le \sum_{k=0}^{\min(N-j,i)} \left(\frac{1}{8}\right)^k S_0^{i,j} \le \sum_{k=0}^{\infty} \left(\frac{1}{8}\right)^k S_0^{i,j} = \frac{8}{7} S_0^{i,j} = \frac{8}{7} S^{i,j}.$$

The case of $i > j$ is proven in the same way. □

2.2 Model Simplification

In this section, we remove some of the transitions without decreasing of expectations $E(B,V)$ and simplify the Markov chain.

Gray States. In a gray state $\langle B, V\rangle$, the expectation of the number of fitness evaluations needed to reach an optimum can be expressed, using the transition probabilities from (2), as follows:

$$E_G(B,V) = \frac{1 + \sum_{i=B+1}^{N} E(i,V) P_G^{B,i}}{\sum_{i=B+1}^{N} P_G^{B,i}}.$$

As by the non-increasing assumption $E(i,V) < E(i+1,V)$, and every probability $P_G^{B,i} > 0$ for $i > B$, the expression above can be bounded by:

$$\begin{aligned} E_G(B,V) &< \frac{1 + \sum_{i=B+1}^{N} E(B+1,V) P_G^{B,i}}{\sum_{i=B+1}^{N} P_G^{B,i}} = \frac{1}{\sum_{i=B+1}^{N} P_G^{B,i}} + E(B+1,V) \le \\ &< \frac{1}{P_G^{B,B+1}} + E(B+1,V) = \frac{1}{P^{B,B+1}} + E(B+1,V). \end{aligned} \tag{5}$$

This effectively means that for any gray state all transitions, except for the transition that increases the number of one-bits exactly by one, can be prohibited without decrease of $E(B,V)$.

White States. The expectation $E(B,V)$ for a white state can be expressed using the transition probabilities from (3):

$$E_W(B,V) = \frac{1 + \sum\limits_{i \neq B} E(i, V \cup \{B\}) P_W^{B,i}}{\sum\limits_{i \neq B} P_W^{B,i}}.$$

If for $i > B+1$ we write $E(B+1, V \cup \{B\})$ as an upper bound for $E(i, V \cup \{B\})$, the following is true ($V' = V \cup \{B\}$ for convenience):

$$E_W(B,V) \leq \frac{1 + \sum\limits_{i=0}^{B-1} E(i,V')P_W^{B,i} + \sum\limits_{i=B+1}^{N} E(B+1,V')P_W^{B,i}}{\sum\limits_{i=0}^{B-1} P_W^{B,i} + \sum\limits_{i=B+1}^{N} P_W^{B,i}} <$$

$$< \frac{1}{\sum\limits_{i=0}^{B-1} P_W^{B,i} + P_W^{B,B+1}} + \frac{\sum\limits_{i=0}^{B-1} E(i,V')P_W^{B,i} + E(B+1,V') \sum\limits_{i=B+1}^{N} P_W^{B,i}}{\sum\limits_{i=0}^{B-1} P_W^{B,i} + \sum\limits_{i=B+1}^{N} P_W^{B,i}}.$$

The right fraction is effectively the linear combination of $E(B+1, V')$ and all $E(i,V')$ for $0 \leq i < B$, and the quotients are normalized (their sum is equal to 1). By the non-increasing assumption, we know that $E(i,V') > E(B+1,V')$. So if we decrease the quotient at $E(B+1,V')$ and increase appropriately the other quotients, the sum will increase. We replace $\sum_{i=B+1}^{N} P_W^{B,i}$ to $P_W^{B,B+1}$, which yields:

$$E_W(B,V) < \frac{1 + \sum\limits_{i=0}^{B-1} E(i,V')P_W^{B,i} + E(B+1,V')P_W^{B,B+1}}{\sum\limits_{i=0}^{B-1} P_W^{B,i} + P_W^{B,B+1}}. \tag{6}$$

So if transitions from a white state, which increase the number of one-bits by more than one, are prohibited, the expectation does not decrease. Together with the same statement for gray states, this helps constructing a simplified Markov chain. As for all $i \in [0; N]$ it holds that $E(0, \{\}) \geq E(i, \{\})$ by the non-increasing assumption, the optimization starts from zero number of one-bits in the worst case, which yields the Markov chain shown on Fig. 1.

2.3 Total Running Time

In this section, we finally prove the $O(N \log N)$ bound, given the non-increasing assumption holds.

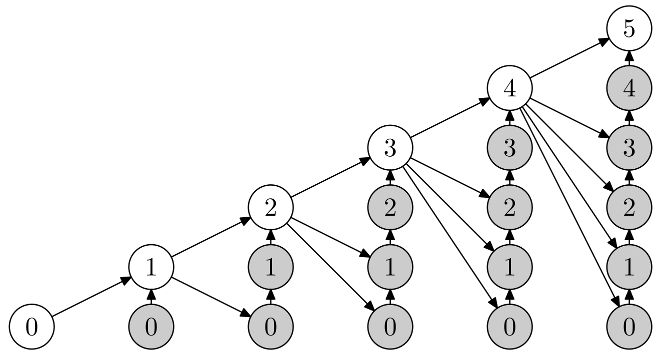

Fig. 1. An example of the simplified Markov chain (with all transitions which increase the number of one-bits by more than one are removed) for $N = 5$. Labels of the states are equal to the number of one-bits. States in which the optimal actions are known are marked gray. Loops in each state (except for one with the label 5, which has no loop) are not shown. Probabilities for each non-loop transition are the same as in the non-simplified Markov chain.

Expression for the Total Running Time. Denote as $E_U(i,j)$ the expectation on the number of steps necessary to get from a gray vertex with i one-bits to a white vertex with j one-bits. Using (5), we can write it as:

$$E_U(i,j) = \frac{1}{P^{i,i+1}} + E_U(i+1,j) = \sum_{k=i}^{j-1} \frac{1}{P^{k,k+1}}.$$

Denote as $E_V(i)$ the expectation on the number of steps necessary to get from a white vertex with i one-bits to a white vertex with $i+1$ one-bits. This can be written using (6) and the equation above as:

$$E_V(i) = \frac{1 + \sum_{j=0}^{i-1} E_U(j, i+1) P_W^{i,j}}{P_W^{i,i+1} + \sum_{j=0}^{i-1} P_W^{i,j}} = \frac{2 + \sum_{j=0}^{i-1} \sum_{k=j}^{i} \frac{P^{i,j}}{P^{k,k+1}}}{P^{i,i+1} + \sum_{j=0}^{i-1} P^{i,j}}.$$

The total running time is just the sums of all $E_V(i)$ for $i \in [0; N)$:

$$E = \sum_{i=0}^{N-1} E_V(i) = \sum_{i=0}^{N-1} \frac{2 + \sum_{j=0}^{i-1} \sum_{k=j}^{i} \frac{P^{i,j}}{P^{k,k+1}}}{P^{i,i+1} + \sum_{j=0}^{i-1} P^{i,j}}. \qquad (7)$$

Lower Bound on Denominators. The following lower bound on denominators of the fractions from (7) is true:

Theorem 2. *$P^{i,i+1} + \sum\limits_{j=0}^{i-1} P^{i,j} > \left(1 - \frac{1}{N}\right)^N$ for sufficiently large N.*

Proof. By Theorem 1:

$$\begin{aligned}
\sum_{j=0}^{i-1} P^{i,j} &\geq \sum_{j=0}^{i-1} \binom{i}{i-j} \left(\frac{1}{N}\right)^{i-j} \left(1 - \frac{1}{N}\right)^{N-i+j} = \\
&= \left(1 - \frac{1}{N}\right)^{N-i} \left(\sum_{j=0}^{i-1} \binom{i}{i-j} \left(\frac{1}{N}\right)^{i-j} \left(1 - \frac{1}{N}\right)^{j}\right) = \\
&= \left(1 - \frac{1}{N}\right)^{N-i} \left(1 - \left(1 - \frac{1}{N}\right)^{i}\right) = \left(1 - \frac{1}{N}\right)^{N-i} - \left(1 - \frac{1}{N}\right)^{N}.
\end{aligned}$$

Using the same theorem, $P^{i,i+1} \geq \frac{N-i}{N}\left(1 - \frac{1}{N}\right)^{N-1}$. Their sum $S(i)$ is equal to:

$$\begin{aligned}
S(i) &= \frac{N-i}{N}\left(1 - \frac{1}{N}\right)^{N-1} + \left(1 - \frac{1}{N}\right)^{N-i} - \left(1 - \frac{1}{N}\right)^{N} \\
&= \left(1 - \frac{1}{N}\right)^{N-1} \left(\left(\frac{N}{N-1}\right)^{i-1} - \frac{i-1}{N}\right).
\end{aligned} \tag{8}$$

Let us find the minimum of $f(i) = \left(\frac{N}{N-1}\right)^{i-1} - \frac{i-1}{N}$ for $i \in [0; N-1]$. This function is continuous at the given interval, so, by the extreme value theorem, the minimum is reachable at either an endpoint of the interval or a zero of the derivative. The values at the endpoints are:

- $f(0) = \frac{N-1}{N} + \frac{1}{N} = 1$;
- $f(N-1) = \left(1 + \frac{1}{N-1}\right)^{N-2} - \frac{N-2}{N} \approx e - 1$.

The derivative is:

$$\left(\left(\frac{N}{N-1}\right)^{i-1} - \frac{i-1}{N}\right)' = \left(\frac{N}{N-1}\right)^{i-1} \ln\frac{N}{N-1} - \frac{1}{N}.$$

Its zero is at the point:

$$i - 1 = -\log_{\frac{N}{N-1}} N \ln\frac{N}{N-1}.$$

The value of $f(i)$ at this point is:

$$f\left(1 - \log_{\frac{N}{N-1}} N \ln\frac{N}{N-1}\right) = \frac{1}{N \ln\frac{N}{N-1}} + \frac{\log_{\frac{N}{N-1}} N \ln\frac{N}{N-1}}{N} \geq \frac{1}{N \ln\frac{N}{N-1}}. \tag{9}$$

To estimate this expression, we use the expression for partial sums of harmonic series:

$$\sum_{k=1}^{N} \frac{1}{k} = \ln(N) + \gamma + \frac{1}{2N} - \frac{1}{12N^2} + O\left(\frac{1}{N^4}\right).$$

From the previous expression we get:

$$\begin{aligned}
\ln \frac{N}{N-1} &= \ln(N) - \ln(N-1) = \\
&= \frac{1}{N} + \frac{1}{2(N-1)} - \frac{1}{2N} + \frac{1}{12N^2} - \frac{1}{12(N-1)^2} + O\left(\frac{1}{N^4}\right) = \\
&= \frac{1}{N} + \frac{1}{2N(N-1)} + O\left(\frac{1}{N^3}\right).
\end{aligned}$$

We can estimate the value in (9) for sufficiently large N as follows:

$$f(i) \geq \frac{1}{N \ln \frac{N}{N-1}} = \frac{1}{N\left(\frac{1}{N} + \frac{1}{2N(N-1)} + O\left(\frac{1}{N^3}\right)\right)} \geq \frac{1}{1 + \frac{1}{N-1}} = \frac{N-1}{N}.$$

This, together with (8), immediately proves the theorem. □

Upper Bound on Numerator. The following upper bound on parts of numerators of the fractions from (7) is true:

Theorem 3. *The following inequality is true:*

$$\sum_{j=0}^{i-1} P^{i,j} \sum_{k=j}^{i} \frac{1}{P^{k,k+1}} \leq \frac{8e}{7}\left(\frac{i}{N-i} + \frac{i}{N-i+1}\right).$$

Proof. First, we rewrite the probabilities using the bounds from Theorem 1: the upper bounds for numerators and the lower bounds for denominators.

$$\begin{aligned}
\sum_{j=0}^{i-1} \sum_{k=j}^{i} \frac{P^{i,j}}{P^{k,k+1}} &= \sum_{j=1}^{i} \sum_{k=i-j}^{i} \frac{P^{i,i-j}}{P^{k,k+1}} \leq \frac{8}{7} \sum_{j=1}^{i} \sum_{k=i-j}^{i} \frac{\binom{i}{j}\left(\frac{1}{N}\right)^j \left(1-\frac{1}{N}\right)^{N-j}}{(N-k)\left(\frac{1}{N}\right)\left(1-\frac{1}{N}\right)^{N-1}} = \\
&= \frac{8}{7} \sum_{j=1}^{i} \binom{i}{j} \left(\frac{1}{N}\right)^{j-1} \left(1-\frac{1}{N}\right)^{1-j} \sum_{k=i-j}^{i} \frac{1}{N-k} = \\
&= \frac{8}{7} \sum_{j=1}^{i} \binom{i}{j} \left(\frac{1}{N-1}\right)^{j-1} \left(\sum_{k=N-i}^{N-i+j} \frac{1}{k}\right).
\end{aligned}$$

We estimate the ratio of two consecutive addends of the sum above:

$$\begin{aligned}
\frac{s_j}{s_{j+1}} &= (N-1) \frac{\binom{i}{j} \sum_{k=N-i}^{N-i+j} \frac{1}{k}}{\binom{i}{j+1}\left(\sum_{k=N-i}^{N-i+j} \frac{1}{k} + \frac{1}{N-i+j+1}\right)} = \\
&= \frac{(N-1)(j+1)}{i-j} \frac{\sum_{k=N-i}^{N-i+j} \frac{1}{k}}{\sum_{k=N-i}^{N-i+j} \frac{1}{k} + \frac{1}{N-i+j+1}}.
\end{aligned}$$

Denote as $X(i,j)$ the value of $\sum_{k=N-i}^{N-i+j} \frac{1}{k}$. Note that $X(i,j) \geq \frac{j}{N-i+j}$. We estimate the right multiple of the expression above as follows:

$$\frac{\sum_{k=N-i}^{N-i+j} \frac{1}{k}}{\sum_{k=N-i}^{N-i+j} \frac{1}{k} + \frac{1}{N-i+j+1}} = \frac{X(i,j)}{X(i,j) + \frac{1}{N-i+j+1}} = 1 - \frac{1}{(N-i+j+1)X(i,j)+1} \geq$$
$$\geq 1 - \frac{1}{(N-i+j)\min_{i',j'} X(i',j')+1} \geq$$
$$\geq 1 - \frac{1}{(N-i+j)\frac{j}{N-i+j}+1} = \frac{j}{j+1}.$$

Using this, we estimate s_j/s_{j+1} as:

$$\frac{s_j}{s_{j+1}} \geq \frac{(N-1)(j+1)}{i-j} \frac{j}{j+1} = \frac{N-1}{i-j} j \geq j.$$

By using this estimation, we prove that the required sum is at most:

$$\sum_{j=0}^{i-1} P^{i,j} \sum_{k=j}^{i} \frac{1}{P^{k,k+1}} \leq \frac{8}{7} \sum_{j=1}^{i} s_j \leq \frac{8}{7} s_1 \sum_{j=1}^{i} \frac{1}{(j-1)!} \leq \frac{8e}{7} s_1 =$$
$$= \frac{8e}{7} \left(\frac{i}{N-i} + \frac{i}{N-i+1} \right). \quad \square$$

Final Result We conclude by the proof of the $O(N \log N)$ bound.

Theorem 4. *The running time of the algorithm, given that the non-increasing assumption is true, is at most* $\frac{16e^2}{7} N \ln N + o(N \ln N)$.

Proof. We substitute the bounds from Theorems 2 and 3 into (7):

$$E = \sum_{i=0}^{N-1} \frac{2 + \sum\limits_{j=0}^{i-1} \sum\limits_{k=j}^{i} \frac{P^{i,j}}{P^{k,k+1}}}{P^{i,i+1} + \sum\limits_{j=0}^{i-1} P^{i,j}} \leq \frac{8}{7} \frac{e}{\left(1-\frac{1}{N}\right)^N} \sum_{i=0}^{N-1} \left(2 + \frac{i}{N-i} + \frac{i}{N-i+1} \right) \leq$$
$$\leq \frac{8}{7} \frac{e}{\left(1-\frac{1}{N}\right)^N} \sum_{i=0}^{N-1} \left(2 + \frac{i}{N-i} + \frac{i}{N-i} \right) \leq \frac{16}{7} \frac{e}{\left(1-\frac{1}{N}\right)^N} \sum_{i=0}^{N=1} \frac{N}{N-i} =$$
$$= \frac{16}{7} \frac{e}{\left(1-\frac{1}{N}\right)^N} (N \ln N + o(N \ln N)) = \frac{16e^2}{7} N \ln N + o(N \ln N). \quad \square$$