On Proportions of Fit Individuals in Population of Genetic Algorithm with Tournament Selection

Anton Eremeev

Abstract
In this paper, we consider a fitness-level model of a non-elitist mutation-only genetic algorithm (GA) with tournament selection. The model provides upper and lower bounds for the expected proportion of the individuals with fitness above given thresholds. In the case of GA with bitwise mutation and OneMax fitness function, the lower bounds are tight when population size equals one, while the upper bounds are asymptotically tight when population size tends to infinity.

The lower bounds on expected proportions of sufficiently fit individuals may be obtained from the probability distribution of an appropriate Markov chain. This approach yields polynomial upper bounds on the runtime of an Iterated version of the GA on 2-SAT problem and on a family of Set Cover problems proposed by E. Balas.

Keywords: Genetic algorithms, fitness levels, combinatorial optimization, Markov chains

1. Introduction
In this paper, we propose a mathematical model of a simplified version of the non-elitist genetic algorithm (GA) based on mutation and selection operators and evaluate the probability distribution of its population at a given iteration. We study the GA with s-tournament selection operator which randomly chooses s individuals from the existing population and selects the best one of them (see e.g. [9, 17]).
The estimates of GA behavior are based on a priori known parameters of the mutation operator. Using the proposed model we obtain upper and lower bounds on expected proportion of the individuals with fitness above certain thresholds. In this framework, we analyze the ONEMAX function, the vertex cover and set cover problems on graphs with special structure and the 2-satisfiability problem (2-SAT). The results presented here extend those from [4].

Let the optimization problem consist in maximization of an objective function $f$ on the set of feasible solutions $\text{Sol} \subseteq X = \{0, 1\}^n$, where $X$ is the search space of all binary strings of length $n$. The GA searches for the optimal or sub-optimal solutions using a population of individuals, where each individual (genotype) $g$ is a bitstring $(g_1, g_2, ..., g_n)$, and its components $g^i \in \{0, 1\}$, $i = 1, 2, ..., n$ are called genes.

In each iteration the GA constructs a new population on the basis of the previous one. The search process is guided by the values of fitness function

$$\phi(g) = \begin{cases} f(g) & \text{if } g \in \text{Sol}; \\ r(g) & \text{otherwise}, \end{cases}$$

where $r(\cdot)$ is a penalty function.

The individuals of the population may be ordered according to the sequence in which they are generated, thus the population may be considered as a vector of genotypes $X^t = (g_1^{(t)}), g_2^{(t)}, ..., g_\lambda^{(t)})$, where $\lambda$ is the size of population, which is constant during the run of the GA, and $t$ is the number of the current iteration. In this paper, we consider the non-elitist version of GA, where all individuals of a new population are generated independently from each other with identical probability distribution depending on the existing population only.

Each individual is generated through selection of a parent genotype by means of a selection operator, and modification of this genotype in mutation operator. During the mutation, a subset of genes in the genotype string $g$ is randomly altered. In general the mutation operator may be viewed as a random variable $\text{Mut}(g) \in X$ with the probability distribution depending on $g$.

The genotypes of the initial population $X^0$ are generated with some a priori chosen probability distribution. The stopping criterion is usually the limit of maximum iterations $t_{\text{max}}$. The result is the best solution generated during the run. The GA has the following scheme.
1. Generate the initial population $X^0$.
2. For $t := 0$ to $t_{max} - 1$ do
   2.1. For $k := 1$ to $\lambda$ do
      Choose a parent genotype $g$ from $X^t$ by $s$-tournament selection.
      Add $g_k^{(t+1)} = \text{Mut}(g)$ to the population $X^{t+1}$.

In the theoretical analysis the GAs are usually treated without a stopping
criterion. Unless the otherwise is stated, we will also assume that $t_{max} = \infty$.

The GA may be considered as a Markov chain in a number of ways. For
example, the states of the chain may correspond to different vectors of $\lambda$
genotypes that constitute the population $X^t$ (see [14]). In this case the num-
ber of states in the Markov chain is $2^{n\lambda}$. Another model representing the
GA as a Markov chain is proposed in [13], where all populations which differ
only in the ordering of individuals are considered to be equivalent. Each
state of this Markov chain may be represented by a vector of $2^n$ components,
where the proportion of each genotype in the population is indicated by
the corresponding coordinate. In the framework of this model, M.Vose and
collaborates have obtained a number of general results concerning the emer-
gent behavior of GAs by linking these algorithms to the infinite-population
GAs [18, 19].

The major difficulties in application of the above mentioned models to
the analysis of GAs for combinatorial optimization problems are connected
with the necessity to use the high-grained information about fitness value of
each genotype. In the present paper, we consider one of the ways to avoid
these difficulties by means of grouping the genotypes into larger classes on
the basis of their fitness.

The rest of the paper is organized as follows. In Section 2, we introduce an
approximating model of GA population and define some required parameters
of probability distribution of the mutation operator in terms of fitness levels.
In Section 3, using the model from Section 2, we obtain lower and upper
bounds on expected proportion of genotypes with fitness above some given
thresholds. Section 4 is devoted to analysis of an important special case
of monotone mutation operator, where the bounds obtained in the previous
section become tight or asymptotically tight. In Section 5, we consider some
illustrative examples of monotone mutation operators and demonstrate two
applications of the lower bounds from Section 3 that yield new upper bounds
on the GA runtime for a family of set cover problems proposed by E. Balas
and for the 2-SAT problem. Section 6 contains concluding remarks.
2. The Approximating Model

Assume that there are $d$ level lines of the fitness function fixed such that $\phi_1 < \phi_2 \ldots < \phi_d$ and $\phi_0 := \min_{g \in \{0,1\}^n} \phi$. The number of levels and the fitness values corresponding to them may be chosen arbitrarily, but they should be relevant to the given problem and the mutation operator to yield a meaningful model. Let us introduce the sequence of Lebesgue subsets of $\mathcal{X}$

$$H_i := \{ g : \phi(g) \geq \phi_i \}, \quad i = 0, \ldots, d.$$ 

Obviously, $H_0 = \mathcal{X}$. For the sake of convenience, we define $H_{d+1} := \emptyset$. Also, we denote the level sets $A_i := H_i \setminus H_{i+1}, \quad i = 0, \ldots, d$.

Now suppose that for all $i = 0, \ldots, d$ and $j = 1, \ldots, d$, the a priori lower bounds $\alpha_{ij}$ and upper bounds $\beta_{ij}$ on mutation transition probabilities from subset $A_i$ to $H_j$ are known, i.e.

$$\alpha_{ij} \leq \Pr\{\text{Mut}(g) \in H_j\} \leq \beta_{ij} \text{ for any } g \in A_i.$$ 

Let $A$ denote the matrix with the elements $\alpha_{ij}$ where $i = 0, \ldots, d$, and $j = 1, \ldots, d$. The similar matrix of upper bounds $\beta_{ij}$ is denoted by $B$. Let the population on iteration $t$ be represented by the population vector $z(t) = (z_1(t), z_2(t), \ldots, z_d(t))$

where $z_i(t) \in [0,1]$ is the proportion of genotypes from $H_i$ within the population $X^t$. The population vector $z(t)$ is a random vector, where $z_i(t) \geq z_{i+1}(t)$ for $i = 1, \ldots, d - 1$ since $H_{i+1} \subseteq H_i$. Let $\Pr\{g(t) \in H_j\}$ be the probability that an individual, which is added after selection and mutation into $X^t$, has a genotype from $H_j$ for $j = 0, \ldots, d$, and $t > 0$ (according to the scheme of the algorithm this probability is identical for all genotypes of $X^t$, i.e. $\Pr\{g(t) \in H_j\} = \Pr\{g_1(t) \in H_j\} = \ldots = \Pr\{g_\lambda(t) \in H_j\}$).

**Proposition 1.** $\mathbb{E}[z_i(t)] = \Pr\{g(t) \in H_i\}$ for all $t > 0, i = 1, \ldots, d$.

**Proof.** Consider the sequence of identically distributed random variables $\xi_1, \xi_2, \ldots, \xi_\lambda$, where $\xi_l = 1$ if the $l$-th individual in the population $X^t$ belongs to $H_i$, otherwise $\xi_l = 0$. By the definition, $z_i(t) = \sum_{l=1}^\lambda \xi_l / \lambda$, consequently $\mathbb{E}[z_i(t)] = \sum_{l=1}^\lambda \mathbb{E}[\xi_l] / \lambda = \sum_{l=1}^\lambda \Pr\{g(t) \in H_i\} / \lambda = \Pr\{g(t) \in H_i\}$. □
Level-Based Mutation. Note that if the probability \( \Pr\{ \text{Mut}(g) \in H_j \} \) for all \( i = 0, \ldots, d \), and \( j = 1, \ldots, d \) does not depend on the choice of \( g \in A_i \), then we may assume that \( \alpha_{ij} = \beta_{ij} = \Pr\{ \text{Mut}(g) \in H_j \} \) for all \( i \) and \( j \). In this case the mutation operator will be called level-based. In the case of level-based mutation, the probabilities \( \gamma_{ij} = \Pr\{ \text{Mut}(g) \in H_j \mid g \in A_i \} \) are well-defined. In what follows, we call \( \gamma_{ij} \) a cumulative transition probability. The symbol \( \Gamma \) will denote the matrix of cumulative transition probabilities: \( \Gamma = A = B \).

If the GA with tournament selection uses a level-based mutation operator, then the probability distribution of population \( X_{t+1} \) is completely determined by the vector \( z(t) \). In this case the GA may be viewed as a Markov chain with the states corresponding to the elements of

\[
Z_\lambda := \{ z \in \{0,1/\lambda,2/\lambda,\ldots,1\}^d : z_i \geq z_{i+1}, i = 1,\ldots,d-1 \},
\]

which is the set of all possible vectors of population of size \( \lambda \).

3. Bounds on Expected Proportions of Fit Individuals

In this section, our aim is to obtain lower and upper bounds on \( \mathbb{E}[z(t)] \) for arbitrary \( s \) and \( t \) if the distribution of the initial population is known.

Let \( P_{ch}(S,z) \) denote the probability that the genotype, chosen by the tournament selection from a population with vector \( z \), belongs to a subset \( S \subseteq \mathcal{X} \). Note that if the current population is represented by the vector \( z(t) = z \), then a genotype obtained by selection and mutation would belong to \( H_j \) with conditional probability

\[
\Pr\{ g(t+1) \in H_j \mid z(t) = z \} = \sum_{i=0}^{d} \sum_{g \in A_i} \Pr\{ \text{Mut}(g) \in H_j \mid g \} P_{ch}(\{g\}, z).
\]

3.1. Lower Bounds

Expression (1) and the definitions of bounds \( \alpha_{ij} \) yield for all \( j = 1, \ldots, d \):

\[
\Pr\{ g(t+1) \in H_j \mid z(t) = z \} \geq \sum_{i=0}^{d} \alpha_{ij} \sum_{g \in A_i} P_{ch}(\{g\}, z) = \sum_{i=0}^{d} \alpha_{ij} P_{ch}(A_i, z),
\]

which turns into an equality in the case of level-based mutation.
Given a tournament size \( s \) we obtain the following selection probabilities:

\[
P_{ch}(H_i, z^{(t)}) = 1 - (1 - z_i^{(t)})^s, \ i = 1, \ldots, d,
\]

and, consequently, \( P_{ch}(A_i, z) = (1 - z_{i+1})^s - (1 - z_i)^s \). This leads to the inequality:

\[
Pr\{g^{(t+1)} \in H_j | z^{(t)} = z\} \geq \sum_{i=0}^{d} \alpha_{ij}((1 - z_{i+1})^s - (1 - z_i)^s).
\]

By the total probability formula,

\[
Pr\{g^{(t+1)} \in H_j\} = \sum_{z \in Z_\lambda} Pr\{g^{(t+1)} \in H_j | z^{(t)} = z\} Pr\{z^{(t)} = z\} \geq \sum_{z \in Z_\lambda} \sum_{i=0}^{d} \alpha_{ij}((1 - z_{i+1})^s - (1 - z_i)^s) Pr\{z^{(t)} = z\} = \\
\sum_{i=0}^{d} \sum_{j=1}^{d} \alpha_{ij}E[(1 - z_{i+1}^{(t)})^s - (1 - z_i^{(t)})^s].
\]

Proposition 1 implies that \( E[z_j^{(t+1)}] = Pr\{g^{(t+1)} \in H_j\} \). Consequently, since \((1 - z_{d+1}^{(t)})^s = 1\) and \((1 - z_0^{(t)})^s = 0\), expression (4) gives a lower bound

\[
E[z_j^{(t+1)}] \geq \alpha_{dj} - \sum_{i=1}^{d} \alpha_{i,j}E[(1 - z_i^{(t)})^s] - \sum_{i=1}^{d} \alpha_{i,j}E[(1 - z_i^{(t)})^s].
\]

Note that (5) turns into equality in the case of level-based mutation. We would like to use (5) recursively \( t \) times in order to estimate \( E[z^{(t)}] \) for any \( t \), given the initial vector \( E[z^{(0)}] \). It will be shown in the sequel that such a recursion is possible under monotonicity assumptions defined below.

**Monotone Matrices and Mutation Operators.** In what follows, a \(((d+1) \times d)\)-matrix \( M \) with elements \( m_{ij}, i = 0, \ldots, d, \) and \( j = 1, \ldots, d \) will be called **monotone** if \( m_{i-1,j} \leq m_{i,j} \) for all \( i, j \) from 1 to \( d \). Informally, a matrix of bounds on transition probabilities is monotone if for any \( j = 1, \ldots, d \) the genotypes from any subset \( A_i \) have the bound on transition probability to \( H_j \) not less than the bounds of the genotypes from the subsets \( A_{i'} \) for all \( i' < i \). Obviously, for any mutation operator the monotone bounds exist. For example \( A = 0 \) where \( 0 \) is a zero matrix and \( B = U \) where \( U \) is the matrix with all elements equal 1.
If \( A = B \) and this matrix is monotone then operator Mut is called \textit{monotone} as well. Note that any monotone mutation operator is level-based.

Let us impose the assumption of monotonicity on matrix \( A \). Then the simple estimate \((1 - z_i^{(t)})^s \leq 1 - z_i^{(t)}\) may be applied to all terms of the sum in (5), leading to the following proposition.

**Proposition 2.** If matrix \( A \) is monotone, then for any tournament size \( s \geq 1 \) and \( j = 1, \ldots, d \) holds

\[
E[z_j^{(t+1)}] \geq \alpha_{0j} + \sum_{i=1}^{d} (\alpha_{i,j} - \alpha_{i-1,j})E[z_i^{(t)}],
\]

besides that (6) is an equality if the mutation operator is monotone and \( s = 1 \).

Bound (6) is tight for any \( s \) if the population \( X' \) is distributed so that \( E[z_i^{(t)}] = \Pr\{z_i^{(t)} = 1\} \) for all \( i = 1, \ldots, d \) and the mutation operator is monotone.

**Straightforward Lower Bounds.** Let \( W \) be a \((d \times d)\)-matrix with the elements \( w_{ij} = \alpha_{ij} - \alpha_{i-1,j} \); let \( I \) be the identity matrix of the same size, and \( \alpha = (\alpha_{01}, \ldots, \alpha_{0d}) \). Suppose that \( || \cdot || \) is some matrix norm.

**Theorem 1.** If matrix \( A \) is monotone and \( \lim_{t \to \infty} ||W^t|| = 0 \) for \( j = 1, \ldots, d \), then for all \( t \geq 1 \)

\[
E[z^{(t)}] \geq E[z^{(0)}]W^t + \alpha(I - W)^{-1}(I - W^t).
\]

**Proof.** Let us consider a sequence of \( d \)-dimensional vectors \( u^{(0)}, u^{(1)}, \ldots, u^{(t)}, \ldots \), where \( u^{(0)} = E[z^{(0)}], u^{(t+1)} = u^{(t)}W + \alpha \). Note that the right-hand side of (6) will not increase if the components of \( E[z^{(t)}] \) are substituted with their lower bounds. Thus by induction on \( t \) we get: \( E[z^{(t)}] \geq u^{(t)} \) for any \( t \).

By properties of the linear operators, due to the assumption that \( \lim_{t \to \infty} ||W^t|| = 0 \), we conclude that matrix \((I - W)^{-1}\) exists. Using the induction on \( t \), for any \( t \geq 1 \) we obtain the identity:

\[
 u^{(t)} = u^{(0)}W^t + \alpha(I - W)^{-1}(I - W^t),
\]

which leads to inequality (7). \( \square \)
In conditions of Theorem 1, the right-hand side of (7) approaches \( \alpha(I - W)^{-1} \) when \( t \) tends to infinity, thus the limit of this bound does not depend on distribution of the initial population. Also let us note that the inequality (7) turns into equation in the case of a monotone mutation operator and the tournament size \( s = 1 \).

In many GAs an arbitrary given genotype may be produced with a non-zero probability as a result of mutation. For such GAs, the condition \( \alpha_{d_j} - \alpha_{0j} < 1 \) is obviously satisfied for all \( j \), if the matrix \( A \) is properly chosen. In this case one can consider the matrix norm \( \|W\|_\infty = \max_j \sum_{i=1}^d |w_{ij}| \). If \( A \) is monotone then \( w_{ij} = \alpha_{ij} - \alpha_{i-1,j} \geq 0 \), so \( \|W\|_\infty = \max_j (\alpha_{dj} - \alpha_{0j}) < 1 \), and the condition \( \lim_{t \to \infty} \|W^t\|_\infty = 0 \) of Theorem 1 is satisfied.

**Lower Bounds from Associated Markov Chain.** Let \( T \) denote a \((d+1)\times(d+1)\)-matrix, with components

\[
t_{ij} = \alpha_{ij} - \alpha_{i,j+1}, \quad i = 0, \ldots, d, \quad j = 0, \ldots, d - 1,
\]

\[
t_{id} = \alpha_{id}, \quad i = 0, \ldots, d.
\]

Note that \( T \) is a stochastic matrix so it may be viewed as a transition matrix of a Markov chain, associated to the set of lower bounds \( \alpha_{ij} \). This chain is a model of the (1,1)-EA with an artificial monotone mutation operator \( Mut'(\xi) \) where the cumulative transition probabilities are defined by the bounds \( \alpha_{ij}, \quad i = 0, \ldots, d, \quad j = 1, \ldots, d \). The initial search point \( \xi(0) \) of the (1,1)-EA is generated at random with probability distribution defined by the probabilities \( p_{i}^{(0)} := \Pr\{\xi(0) \in A_i\} = \mathbb{E}[z_{i}^{(0)}] - \mathbb{E}[z_{i+1}^{(0)}], \quad i = 0, \ldots, d \). At every iteration \( t \) the next search point \( \xi(t+1) \) is generated as \( \xi(t+1) = Mut'(\xi(t)) \). Denoting \( p^{(t)} := (p_{0}^{(t)}, \ldots, p_{d}^{(t)}) \), where \( p_{i}^{(t)} = \Pr\{\xi(t) \in A_i\} \), by properties of Markov chains we get \( p^{(t)} = p^{(0)} T^t \). The following theorem is based a comparison of \( \mathbb{E}[z(t)] \) to the distribution of the Markov chain \( p^{(t)} \).

**Theorem 2.** Let \( A \) be monotone. Then for any \( t = 1, 2, \ldots \) holds

\[
\mathbb{E}[z_{i}^{(t)}] \geq p_{i}^{(0)} T^t L \tag{8}
\]

where \( L \) is a triangular \((d + 1) \times (d + 1)\)-matrix with components \( \ell_{ij} = 1 \) if \( i \geq j \) and \( \ell_{ij} = 0 \) otherwise.
Proof. The (1,1)-EA described above is identical to a GA with \( \lambda = 1 \), \( s = 1 \) and the artificial monotone mutation operator \( Mut' \) defined by cumulative transition probabilities matrix \( A \). Let us denote the population vector of this GA by \( \hat{z}\). Obviously,

\[
\hat{z}_i(t) = \sum_{k=i}^{d} p_k(t), \quad i = 1, \ldots, d. \tag{9}
\]

Proposition 2 implies that in the original GA with population size \( \lambda \) and tournament size \( s \), the expectation \( E[z(t)] \) is lower bounded by the expectation \( E[\hat{z}(t)] \) since (6) holds as an equality for the whole sequence of \( E[\hat{z}(t)] \) and the right-hand side of (6) is non-decreasing on \( E[z(t)] \). Equality \( p(t) = p(0) T^t \) together with (9) imply the required bound (8). □

3.2. Upper Bounds

In this subsection we obtain upper bounds on \( E[z_j(t+1)] \) using a reasoning similar to the proof of Proposition 2. Expression (1) for all \( j = 1, \ldots, d \) yields:

\[
\Pr\{g(t+1) \in H_j|z(t) = z\} \leq \sum_{i=0}^{d} \beta_{ij} P_{ch}(A_i, z) = \sum_{i=0}^{d} \beta_{ij}((1 - z_{i+1})^s - (1 - z_i)^s). \tag{10}
\]

which turns into equality in the case of level-based mutation. By the total probability formula we have:

\[
E[z_j(t+1)] = \sum_{z \in Z_{\lambda}} \Pr\{g(t+1) \in H_j|z(t) = z\} \Pr\{z(t) = z\} \tag{11}
\]

\[
\leq \sum_{i=0}^{d} \beta_{ij} E[(1 - z_{i+1})^s - (1 - z_i)^s],
\]

so

\[
E[z_j(t+1)] \leq \beta_{dj} - \sum_{i=1}^{d} (\beta_{ij} - \beta_{i-1,j}) E[(1 - z_i(t))^s]. \tag{12}
\]

Under the expectation in the right-hand side we have a convex function on \( z_i(t) \). Therefore, in the case of monotone matrix \( B \), using the Jensen’s inequality we obtain the following proposition.
Proposition 3. If $B$ is monotone then

$$E[z_j^{(t+1)}] \leq \beta_{dj} - \sum_{i=1}^{d}(\beta_{ij} - \beta_{i-1,j})(1 - E[z_i^{(t)}])^s.$$  \hspace{1cm} (13)

Note that unlike (6), this bound is not tight when $E[z_i^{(t)}] = \Pr\{z_i^{(t)} = 1\}$, $i = 1, \ldots, d$. The bound (13) is tight, however, when $z_i^{(t)} = E[z_i^{(t)}]$ for all $i = 1, \ldots, d$ with probability 1.

By means of iterative application of inequality (13) the components of the expected population vectors $E[z^{(t)}]$ may be bounded up to arbitrary $t$, starting from the initial vector $E[z^{(0)}]$. The nonlinearity in the right-hand side of (13), however, creates an obstacle for obtaining an analytical bound similar to the bounds of Theorems 1 and 2.

Note that all of the estimates obtained up to this point are independent of the population size and valid for arbitrary $\lambda$. In the following section we will see that the right-hand side of (13) reflects the asymptotic behavior of population under monotone mutation operator as $\lambda \to \infty$.

4. GA with Monotone Mutation Operator

First of all note that in the case of monotone mutation operator, two equal monotone matrices of lower and upper bounds $A = B$ exist, so the bounds (5) and (12) give equal results, and assuming $\Gamma = A = B$ we get

$$E[z_j^{(t+1)}] = \gamma_{dj} - \sum_{i=1}^{d}(\gamma_{ij} - \gamma_{i-1,j})E[(1 - z_i^{(t)})^s], \hspace{1cm} j = 1, \ldots, d, \hspace{1cm} t = 0, 1, \ldots.$$  \hspace{1cm} (14)

This equality will be used several times in what follows.

Note that in general, the population vectors are random values whose distributions depend on $\lambda$. To express this in the notation let us denote the proportion of genotypes from $H_i$ in population $X^t$ by $z_i^{(t)}(\lambda)$.

Lemma 1. Let the GA use a monotone mutation operator with cumulative transition matrix $\Gamma$, and let the genotypes of the initial population be identically distributed. Then

(i) for all $t = 0, 1, \ldots$ and $i = 1, \ldots, d$ holds

$$\lim_{\lambda \to \infty} \left( E \left[ \left( 1 - z_i^{(t)}(\lambda) \right)^s \right] - \left( 1 - E[z_i^{(t)}(\lambda)]\right)^s \right) = 0;$$  \hspace{1cm} (15)
(ii) if the sequence of $d$-dimensional vectors $u^{(0)}, u^{(1)}, ..., u^{(t)}, ...$ is defined as

$$u^{(0)} = \mathbb{E}[z^{(0)}(\lambda)],$$

$$u^{(t+1)}_j = \gamma_{dj} - \sum_{i=1}^{d} (\gamma_{ij} - \gamma_{i-1,j})(1 - u^{(t)}_i)^s$$

for $j = 1, ..., d$ and $t \geq 0$. Then $\lim_{\lambda \to \infty} \mathbb{E}[z^{(t)}(\lambda)] = u^{(t)}$ for all $j = 1, ..., d$ at any iteration $t$.

**Proof.** From (14), we conclude that if statement (i) holds, then with $\lambda \to \infty$, the convergence of $\mathbb{E}[z^{(t)}(\lambda)]$ to $u^{(t)}$ will imply that $\mathbb{E}[z^{(t+1)}(\lambda)] \to u^{(t+1)}$. Thus, statement (ii) follows by induction on $t$.

Given some $t$, to prove (15) we recall the sequence of i.i.d. random variables $I_1^i, I_2^i, ..., I_{\lambda}^i$, where $I_k^i = 1$, if the $k$-th individual of population $X^t$ belongs to $H_i$, otherwise $I_k^i = 0$. By the law of large numbers, for any $i = 1, ..., d$ and $\varepsilon > 0$, we have

$$\lim_{\lambda \to \infty} \Pr \left\{ \left| \frac{\sum_{k=1}^{\lambda} I_k^i}{\lambda} - \mathbb{E}[I_i] \right| < \varepsilon \right\} = 1.$$

Note that $\sum_{k=1}^{\lambda} I_k^i / \lambda = z_i^{(t)}(\lambda)$. Besides that, due to Proposition 1, $\mathbb{E}[I_i] = \Pr\{I_i = 1\} = \mathbb{E}[z_i^{(t)}(\lambda)]$. (In the case of $t = 0$ this equality holds as well, since all individuals of the initial population are distributed identically.) Therefore, for any $\varepsilon > 0$ the convergence $\Pr \left\{ \left| z_i^{(t)}(\lambda) - \mathbb{E}[z_i^{(t)}(\lambda)] \right| < \varepsilon \right\} \longrightarrow 1$ holds. Now by continuity of the function $(1 - x)^s$, it follows that

$$\lim_{\lambda \to \infty} \Pr \left\{ \left| (1 - z_i^{(t)}(\lambda))^s - (1 - \mathbb{E}[z_i^{(t)}(\lambda)])^s \right| \geq \varepsilon \right\} = 0.$$

Let us denote $F_\lambda(x) := \Pr \left\{ (1 - z_i^{(t)}(\lambda))^s - (1 - \mathbb{E}[z_i^{(t)}(\lambda)])^s < x \right\}$. Then

$$\lim_{\lambda \to \infty} \left( \mathbb{E} \left[ (1 - z_i^{(t)}(\lambda))^s \right] - (1 - \mathbb{E}[z_i^{(t)}(\lambda)])^s \right) \leq \lim_{\lambda \to \infty} \int_{-\infty}^{\infty} x \, dF_\lambda(x) \leq$$

$$\leq \lim_{\lambda \to \infty} \Pr \left\{ \left| (1 - z_i^{(t)}(\lambda))^s - (1 - \mathbb{E}[z_i^{(t)}(\lambda)])^s \right| \geq \varepsilon \right\} + \lim_{\lambda \to \infty} \int_{|x|<\varepsilon} \varepsilon \, dF_\lambda(x) = \varepsilon.$$
for arbitrary $\varepsilon > 0$, hence (15) holds. $\square$

Due to equalities (14) and (15), Lemma 1 implies

**Theorem 3.** If the mutation operator is monotone and individuals of the initial population are distributed identically, then

$$\lim_{\lambda \to \infty} \mathbb{E}[z_{j}^{(t+1)}(\lambda)] = \gamma_{dj} - \sum_{i=1}^{d} (\gamma_{ij} - \gamma_{i-1,j})(1 - \mathbb{E}[z_{i}^{(t)}(\lambda)])^s \quad (18)$$

for all $j = 1, \ldots, d$, $t \geq 0$.

For any $i, j$ and $t > 0$, the term $u_{j}^{(t)}$ of the sequence defined by (17) is nondecreasing in $u_{i}^{(t-1)}$ and in $s$ as well. Using this fact we can expect that the components of population vector of the infinite-population GA will typically increase with the tournament size. To make it precise, we first prove a technical lemma.

**Lemma 2.** Suppose the mutation operator is monotone, the sequence of vectors $\{u_{i}^{(t)}\}$ is defined according to (16) and (17), besides that $u_{j}^{(t)} \in \{0, 1\}$ for some $j \in \{1, \ldots, d\}$ and $t \geq 0$. Then $\mathbb{E}[z_{j}^{(t)}] = u_{j}^{(t)}$.

**Proof.** Let us consider the case of $u_{j}^{(t)} = 0$, using induction on $t$. The statement of the lemma holds for $t = 0$ by the definition of $u_{i}^{(0)}$. Suppose the statement holds for $t$ and prove it for $t + 1$.

For notational convenience we assume that $u_{d+1}^{(t)} = 0$ for all $t$ and define the number $k$ as follows:

$$k := \max\{i : 1 \leq i \leq d + 1, \ u_{i}^{(t)} \neq u_{i-1}^{(t)}\}.$$ 

Then $\gamma_{k-1,j} = 0$, since (17) implies

$$0 = u_{j}^{(t+1)} = \sum_{i=1}^{d+1} \gamma_{i-1,j}((1 - u_{i}^{(t)})^s - (1 - u_{i-1}^{(t)})^s),$$

where $u_{i}^{(t)} \leq u_{i-1}^{(t)}$. The latter inequality follows by induction, since

$$u_{j}^{(t+1)} - u_{j-1}^{(t+1)} = \sum_{i=1}^{d+1} (\gamma_{i-1,j} - \gamma_{i-1,j-1})((1 - u_{i}^{(t)})^s - (1 - u_{i-1}^{(t)})^s).$$
Using the monotonicity assumption we conclude that
\[ \gamma_0 = \cdots = \gamma_{k-1} = 0. \quad (19) \]
The choice of \( k \) implies that \( u^{(t)} = \cdots = u^{(t)}_{d+1} = 0 \), and by inductive assumption we also have \( E[z^{(t)}_k] = \cdots = E[z^{(t)}_{d+1}] = 0 \). Hence, in view of (14) and (19), we have
\[ E[z^{(t+1)}_i(\lambda)] = \sum_{i=1}^{d+1} \gamma_{i-1,j}((1 - E[z^{(t)}_i])^s - (1 - E[z^{(t)}_{i-1}])^s) = 0. \]
The proof for the case of \( u^{(t)}_j = 1 \) is done analogously. \( \Box \)

**Theorem 4.** Let \( z^{(t)} \) and \( \hat{z}^{(t)} \) correspond to GAs with tournament sizes \( s \) and \( \hat{s} \), where \( s \leq \hat{s} \). Besides that, suppose that Mut is monotone and the individuals of initial populations are identically distributed. Then given a sufficiently large \( \lambda \),
\[ E[\hat{z}^{(t)}_i(\lambda)] \geq E[z^{(t)}_i(\lambda)], \quad t = 0, 1, \ldots, i = 1, \ldots, d. \]

**Proof.** For any \( i, j \in \{1, \ldots, d\} \) and \( t > 0 \), the element \( u^{(t)}_j \) of the sequence \( \{u^{(t)}_j\} \) defined in Lemma 1 is a non-decreasing function of \( u^{(t-1)}_i \) and \( s \). Moreover, all of the elements \( 0 < u^{(t)}_i < 1 \) are strictly increasing with growth of \( s \). For all of such \( i \) and \( t \), given sufficiently large population size, the inequality \( E[\hat{z}^{(t)}_i(\lambda)] > E[z^{(t)}_i(\lambda)] \) is satisfied, as it follows from Lemma 1. Otherwise, i.e. when \( u^{(t)}_i \in \{0, 1\} \), the theorem follows from Lemma 2. \( \Box \)

5. Applications and Illustrative Examples

5.1. Examples of Monotone Mutation Operators

Let us consider two optimization problems for which the matrix \( \Gamma \) of cumulative transition probabilities is easily determined.

First we consider the simple fitness function \( ONEMAX(g) \equiv ||g||_1 \), where \( g \in \{0, 1\}^n \) and \( ||g||_1 = \sum_{i=1}^{n} |g_i| \). Suppose that the GA uses the bitwise mutation operator, changing every gene with a given probability \( p_m \), independently of other genes. Let the subsets \( H_0, \ldots, H_d \) be defined by the
level lines $\phi_0 = 0, \phi_1 = 1, \ldots, \phi_d = d$ and $d = n$. The matrix $\Gamma$ for this operator could be obtained using the result from [1], but here we shall consider this example as a special case of more general setting.

Let the representation of the problem admit a decomposition of the genotype string into $d$ non-overlapping substrings (called blocks here) in such a way that the fitness function equals to the number of blocks for which a certain property $K$ holds. The functions of this type belong to the class of additively decomposed functions, where the elementary functions are Boolean and substrings are non-overlapping (see e.g. [12]). Let $m$ be the number of blocks and let $K(g, \ell) = 1$ if $K$ holds for the block $\ell$ of genotype $g$, and $K(g, \ell) = 0$ otherwise (here $\ell = 1, \ldots, m$).

Suppose that during mutation, any block for which $K$ did not hold gets the property $K$ with probability $\tilde{r}$, i.e.

$$\Pr\{K(\text{Mut}(g), \ell) = 1|K(g, \ell) = 0\} = \tilde{r}, \ \ell = 1, \ldots, m.$$ 

On the other hand, assume that a block with the property $K$ keeps this property during mutation with probability $r$, i.e.

$$\Pr\{K(\text{Mut}(g), \ell) = 1|K(g, \ell) = 1\} = r; \ell = 1, \ldots, m.$$ 

Let the subsets $H_0, \ldots, H_d$ correspond to the level lines $\phi_0 = 0, \phi_1 = 1, \ldots, \phi_d = d$ again.

The probability that during mutation $d - i$ blocks without property $K$ would produce $k = 0, \ldots, d - i$ blocks with this property is $(\binom{d-i}{k})\tilde{r}^k(1-\tilde{r})^{d-i-k}$. On the other hand, under mutation in a set of $i$ blocks with property $K$ this property will be lost by not more than $l$ blocks with probability

$$\sum_{\nu=0}^{\min\{i,l\}} \binom{i}{\nu} (1-r)^\nu r^{i-\nu}, \ l = 0, \ldots, d.$$ 

Thus,

$$\gamma_{ij} = \sum_{k=0}^{d-i} \binom{d-i}{k} \tilde{r}^k(1-\tilde{r})^{d-i-k} \sum_{\nu=0}^{\min\{i,i-j+k\}} \binom{i}{\nu} (1-r)^\nu r^{i-\nu}. \quad (20)$$ 

It is shown in [3] that if $r \geq \tilde{r}$ then matrix $\Gamma$ defined by (20) is monotone.

Now matrix $\Gamma$ for the ONEMAX function is obtained assuming that $\tilde{r} = (1-r) = p_m$ and $d = n$. This operator is monotone if $p_m \leq 0.5$. 

14
Expression (20) may be also used for finding the cumulative transition matrices of some other optimization problems with a "regular" structure. As an example, below we consider the vertex cover problem (VCP) on graphs of a special structure.

In general, the vertex cover problem is formulated as follows. Let $G = (V, E)$ be a graph with a set of vertices $V$ and the edge set $E$. A subset $C \subseteq V$ is called a vertex cover of $G$ if every edge has at least one endpoint in $C$. The vertex cover problem is to find a vertex cover $C^*$ of minimal cardinality.

Following [20] we denote by $G(d)$ the graph consisting of $d$ disconnected triangle subgraphs. Each triangle is covered optimally by two vertices and the redundant cover consists of three vertices. In spite of simplicity of this problem, it is proven in [20] that some well-known algorithms of branch and bound type require exponential in $d$ number of iterations if applied to the VCP on graph $G(d)$.

Suppose that the VCP is handled by the GA with the following representation: each gene $g^i \in \{0, 1\}, i = 1, \ldots, |E|$ corresponds to an edge of $G(d)$, assigning one of its endpoints which has to be included in the cover. Let $C(g)$ denote the cover, containing all vertices which are assigned by at least one of the genes and consider the bitwise mutation operator where each gene is flipped with probability $p_m$. A natural way to choose the fitness function in this case is to assume $\phi(g) = |V| - |C(g)|$. Then for $G(d)$, the fitness $\phi(g)$ coincides with the number of optimally covered triangles in $C(g)$. Let the genes representing the same triangle constitute a single block, and let the property $K$ imply that a block is optimally covered. Then $\tilde{r} = 1 - p_m^3 - (1 - p_m)^3$ and $r = 1 - p_m(1 - p_m)^2 - p_m^2(1 - p_m)$. Using (20) we obtain the cumulative transition matrix for this mutation operator. It is easy to verify that in this case the inequality $r \geq \tilde{r}$ holds for any mutation probability $p_m$, and therefore the operator is always monotone.

Computational Experiments. Below we present some experimental results in comparison with the theoretical estimates obtained in Section 3. To this end we consider an application of GA to the VCP on graphs $G_d$. The proportion of the optimal genotypes in the population for different population size is presented in Figure 1. Here $d = 8, p_m = 0.01, s = 2$ and $z^{(0)} = 0$.

The experimental results are shown in dotted lines. The solid lines correspond to the lower and upper bounds given by the expressions (7) and (13). The plot shows that the upper bound (13) gives a good approximation to the
value of $z_d^{(t)}$ even if the population size is not very large. The lower bound (7) coincides with the experimental results when $\lambda = 1$, up to a minor sampling error.

\begin{figure}[h]
\centering
\includegraphics[width=0.5\textwidth]{figure1.png}
\caption{Average proportion of optimal VCP solutions as a function of the iteration number, $d = 8$, $p_m = 0.01$, $s = 2$.}
\end{figure}

Another series of experiments was carried out to compare the behavior of GAs with different tournament sizes. Figure 2 presents the experimental results for the GA with $p_m = 0.1$, $\lambda = 100$ and $z^{(0)} = 0$ solving the VCP on $G_6$. This plot demonstrates the increase in proportion of the optimal genotypes as a function of the tournament size, which is consistent with Theorem 4.

In what follows, we demonstrate the usage of lower bounds $\alpha_{ij}$ when mutation operator is not level-based.

5.2. Lower Bounds and Runtime Analysis for 2-SAT Problem

The Satisfiability problem (SAT) in general is known to be NP-complete [8], but it is polynomially solvable in the special case denoted by 2-SAT: given a Boolean formula with CNF where each clause contains at most two literals, find out whether a satisfying assignment of variables exists.

Let $n$ be the number of logical variables and let $m$ be the number of clauses in the CNF. A natural encoding of solutions is a binary string $g$ where $g_i = 1$ if the $i$-th logical variable has the value “true” and otherwise $g_i = 0$. 

16
We consider a GA with the tournament size \( s = 1 \) and the following mutation operator: choose a clause which is not satisfied and modify one of the variables involved in it, otherwise keep the solution unchanged. This method of random perturbation was proposed in the randomized algorithm of C. Papadimitriou [15] which has the runtime \( O(n^2) \), if the CNF is satisfiable.

The fitness function does not influence the GA execution when \( s = 1 \) but it will be useful here for the sake of theoretical analysis: let us assume that \( \phi(g) \) equals the Hamming distance to a satisfying assignment. Here and below, we assume that a satisfying assignment exists.

For any non-satisfying truth assignment the improvement probability is \( 1/2 \), so we can apply the following monotone bounds: \( \alpha_{ij} = 1 \) for all \( i = 1, \ldots, d, j = 0, \ldots, i - 1 \); \( \alpha_{i,i+1} = 1/2 \) for \( i = 0, \ldots, d - 1 \);

\[
\alpha_{ii} = \begin{cases} 
1/2 & \text{if } i = 1, \ldots, d - 1; \\
1 & \text{if } i = d;
\end{cases}
\]

\( \alpha_{ij} = 0, i = 0, \ldots, d - 2, j = i + 2, \ldots, d \). These bounds define the same matrix \( T \) as the symmetric "gambler’s ruin" random walk with one reflecting barrier (state 0) and one absorbing barrier (state \( d \)). The result from [15] implies that, regardless of the initial state, there exists a constant \( c > 0 \), such that after \( t_{\max} = \lceil cn^2 \rceil \) steps the absorbing probability \( p_d^t \) is at least \( 1/2 \). Theorem 2 yields the following

**Corollary 1.** If the GA for 2-SAT has the tournament size \( s = 1 \) then the
probability of generating a satisfying assignment in population $X[cn^2]$ is at least $1/2$ for some constant $c > 0$.

If the GA is restarted every $t_{max}$ iterations, then the overall runtime of this iterated GA is $O(\lambda n^2)$. In a similar way an Iterated GA can simulate the randomized algorithm of Schöning [16] for $k$-SAT with runtime $O(2 - 2/k)^k$.

5.3. Lower Bounds and Runtime Analysis for Balas Set Cover Problems

In general the set cover problem (SCP) is formulated as follows. Given: $M = \{1, ..., m\}$ and a set of subsets $M_j \subseteq M$, $j \in U := \{1, \ldots, n\}$. A subset $J \subseteq U$ is called a cover if $\cup_{j \in J} M_j = M$. The goal is to find a cover of minimum cardinality. In what follows we denote by $N_i$ the set of numbers of the subsets that cover an element $i$, i.e. $N_i = \{j : i \in M_j\}$ for any $i$.

Consider a family $B(n, k)$ of SCPs introduced by E. Balas in [2]. Here it is assumed that $m = C^{k-1}_n$ and the set $\{N_1, N_2, ..., N_m\}$ consists of all $n - k + 1$-element subsets of $U$. Thus $J \subseteq U$ is an optimal cover iff $|J| = k$.

Family $B(n, k)$ is known to have large fractional cover [2] which implies that these SCPs are hard for integer programming algorithms. In particular, it was shown in [20] that problems from this class are hard to solve using the $L$-class enumeration method [10]. When $n$ is even and $k = n/2$, the $L$-class enumeration method needs an exponential number of iterations in $n$. In what follows we analyze GA in this special case.

Note that in the case of tournament selection, equivalently to studying the GA for $B(n, n/2)$ SCPs with even $n$ we may study the GA where the fitness is given by a function of unitation, such that

\[
\phi(g) = \begin{cases} 
R(||g||_1) & \text{if } ||g||_1 \geq n/2; \\
L(||g||_1) & \text{otherwise},
\end{cases}
\]

where function $R$ is decreasing, function $L$ is increasing and $L(n/2 - 1) < R(n)$.

Consider a point mutation operator: with probability $q > 0$ keep the given genotype unchanged; otherwise (with probability $1 - q$) choose $i$ from $\{1, \ldots, n\}$ and change gene $i$. Let $d = n/2$ and let the thresholds $\phi_0, \phi_1, ..., \phi_d$
be equal to fitness of genotypes that contain 0, 1, ..., d genes "1" accordingly. Note that $J(g)$ is a cover iff $\phi(g) \geq \phi_d$.

We have the following bounds: $\alpha_{ij} = 1$ for all $i = 1, \ldots, d$, $j = 0, \ldots, i - 1$; $\alpha_{i,i+1} = (1 - q)(n - i)/n$ for $i = 0, \ldots, d - 1$;

$$
\alpha_{ii} = \begin{cases} 
q + \alpha_{i,i+1} & \text{if } i = 1, \ldots, d - 1; \\
q & \text{if } i = d;
\end{cases}
$$

$\alpha_{i,j} = 0$, $i = 0, \ldots, d - 2$, $j = i + 2, \ldots, d$. These lower bounds $\alpha_{ij}$ coincide with the corresponding cumulative transition probabilities except for level $d$, where we pessimistically assume $\alpha_{dd} = q$. It is easy to verify that $A$ satisfies the monotonicity condition when $q \geq 1/(n + 1)$.

The Markov chain associated with these lower bounds has the following nonzero transition probabilities

$$
t_{ii} = q, \ t_{i,i-1} = (1 - q)i/n, \ t_{i,i+1} = (1 - q)(1 - i/n), \ i = 1, \ldots, d - 1,
$$

$$
t_{0,1} = 1 - q, \ t_{d,d-1} = 1 - q.
$$

In what follows, first we will obtain a lower bound on $E[z^{(t)}_d]$ for $t \to \infty$, using Theorem 2 and the stationary distribution of the associated Markov chain $P^{(t)} = P^{(0)} T^t$. After that, using Theorem 1 we will compute a simple lower bound on $E[z^{(t)}_d]$ for finite $t$.

The stationary distribution of the associated Markov chain may be found from the well-known model for diffusion of P. Ehrenfest and T. Ehrenfest. Consider $n$ molecules in a rectangular container divided into two equal parts A and B. At any time $t$, one randomly chosen molecule moves to another part. The state of the system is defined by the number of molecules $j$, $j = 0, \ldots, n$, in container A. The corresponding random walk has transition probabilities

$$
\tau_{j,j-1} = j/n, \ \tau_{j,j+1} = 1 - j/n, \ j = 1, \ldots, n - 1,
$$

$$
\tau_{0,1} = 1, \ \tau_{n,n-1} = 1.
$$

The stationary distribution in Ehrenfests model (see e.g [7], chapter. 15, § 6) is given by $\pi_j := \binom{n}{j}/2^n$, $j = 0, \ldots, n$. Grouping each couple of symmetric states (i.e. the state where A contains $j$ molecules, B contains $n - j$ molecules and the state where A contains $n - j$ molecules and B contains $j$ molecules, $j = 0, \ldots, n/2$) into one state we conclude that the Markov chain with transition matrix $T$ has the stationary distribution $u = (2\pi_1, \ldots, 2\pi_d)$ for any $q < 1$. So by Theorem 2, vector $uL$ is the limiting lower bound for $E[z^{(t)}_d]$.  

19
If we are interested in transient behavior of this GA, we can obtain a lower bound for the expected population vector $\mathbb{E}[z(t)]$ using Theorem 1. Consider the matrix norm $||W||_{\infty} = \max \sum_{j=1}^{d} |w_{ij}|$ which is associated to the vector norm $||\cdot||_1$ in the case of right-hand side multiplication of vectors. For the matrix $W$, corresponding to the set of lower bounds $\alpha_{ij}$, defined above, we have $||W||_{\infty} = 1 - 2(1 - q)/n$, i.e. the condition $\lim_{t \to \infty} ||W^t||_{\infty} = 0$ is satisfied for $q < 1$.

Let us find the limiting population vector $v = a(I - W)^{-1}$ as $t \to \infty$. To this end, it suffices to solve the system of equations

$$
-\frac{n - i + 1}{n} + v_{i-1} - \frac{i}{n} = 0, \quad i = 2, \ldots, \frac{n}{2} - 1, \quad (21)
$$

$$
v_{1} - \frac{2}{n} = 1, \quad -v_{n/2 - 1} + \frac{n + 2}{2n} = 0. \quad (22)
$$

Theorem 2 suggests to put $v = uL$, i.e.

$$
v_i = \frac{1}{2^{n-1}} \sum_{\ell=i}^{\frac{n}{2}} \binom{n}{\ell}, \quad i = 1, \ldots, \frac{n}{2}. \quad (23)
$$

Indeed, (21) and (22) for this vector $v$ are verified straightforwardly.

Let $e = (1, \ldots, 1)$. By properties of the norms under consideration, $vW^t \leq e||vW^t||_1 \leq e||v||_1 \cdot ||W||_t^\infty \leq d||W||_\infty^t e$, so by Theorem 1

$$
\mathbb{E}[z(t)] \geq \mathbb{E}[z^{(0)}]W^t + a(I - W)^{-1}(I - W^t) \geq v - d||W||_\infty^t e
$$

for any $t$. With $q = 1/(n + 1)$, the average proportion of feasible genotypes is lower-bounded by $v_d = d\left(\frac{n-1}{n+1}\right)^t$. Using (23) and the Stirling inequality we conclude that $v_d = \left(\frac{n}{n/2}\right)^{t} = \Omega(n^{-\frac{1}{2}})$. Note that given a sufficiently large $t \geq \frac{n+1}{2} \ln \frac{n}{v_d}$,

$$
\frac{v_d}{n} \geq \left(\frac{1}{e}\right)^{\frac{2t}{n+1}} \geq \left(\left(1 - \frac{2}{n+1}\right)^{\frac{n+1}{2}}\right)^{\frac{2t}{n+1}} = \left(\frac{n - 1}{n + 1}\right)^t,
$$

so $\frac{n}{2} \left(\frac{n-1}{n+1}\right)^t \leq \frac{v_d}{2}$ and $\mathbb{E}[z_d^{(t)}] \geq \frac{v_d}{2}$.

Suppose that the initial population consists of all-zero strings. Then presence of at least one individual from $H_d$ in the current population implies
that an optimal solution to a problem $B(n, n/2)$ was already found at least once. This way, in view of Proposition 1, after $\Omega(n \ln n)$ iterations of the GA with $q = 1/(n+1)$, the probability of finding an optimum is at least $\Omega(n^{-0.5})$ and we have proved the following

**Corollary 2.** Given $X^0 = (0, \ldots, 0)$ and a tournament size $s \geq 1$, there exists a constant $c$, such that the probability to reach an optimum of problem $B(n, n/2)$ until iteration $\lceil cn \ln n \rceil$ is $\Omega(n^{-0.5})$.

If the GA is restarted every $t_{\text{max}} = \lceil cn \ln n \rceil$ iterations, then the overall runtime of this iterated GA is $O(\lambda n^{1.5} \log n)$ for any $\lambda$. Note that a more general result from [5] gives a weaker runtime bound $O(n^2 \log n)$ for this iterated GA with $\lambda = \Omega(n \log n)$.

The probability that an optimal solution with $k = n/2$ ones will not downgrade to a lower fitness level is $O(1/n)$. In such a case the GA analysis from [11] is not applicable because it requires that the downgrade probability is bounded by a constant less than 1. Besides that, the approach from [11] requires a non-zero selection pressure, i.e. $s > 1$.

**6. Conclusions**

In this paper, we presented an approximating model of non-elitist mutation-based GA with tournament selection and obtained upper and lower bounds on proportion of ”sufficiently good” genotypes in population using this model. In the special case of monotone mutation operator, the obtained bounds become tight in different situations. The analysis of infinite population GA with monotone mutation suggests an optimal selection mechanism converting the GA into $(1, \lambda)$-EA. Application of the lower bounds to a family of set covering problems proposed by E. Balas and to the 2-satisfiability problem yields new runtime bounds for the iterated version of the GA.

It is expected that the further research will involve applications of the proposed approach to other combinatorial optimization problems, in particular, the problems with regular structure.

The obtained lower and upper bounds do not take the tournament size into account. It remains an open research question of how to construct the tighter bounds w.r.t. the tournament size. The bounds presented in this paper are already tight for some ”exotic” distributions of the current population, which implies that subsequent research might benefit from joining the analysis of expectation of population vector with some variance analysis.
Another open question is how to incorporate the crossover operator into the approximating model. For some types of crossover operators, such as those based on solving the optimal recombination problem [6], the lower bounds from this paper may be easily extended, ignoring the improving capacity of crossover. It is important, however, to take the improving capacity of crossover into account and it is not clear how the monotonicity conditions could be meaningfully extended for this purpose.

Acknowledgements

The research was supported by Russian Foundation for Basic Research grant 15-01-00785. The author is grateful to Sergey A. Klokov and Boris A. Rogozin for helpful comments on earlier versions of this work.

References


