

Approximating the Genetic Diversity of Populations in the Quasi-Equilibrium State

Jarosław Arabas

Abstract—This paper analyzes an evolutionary algorithm in the quasi-equilibrium state, i.e., when the population of chromosomes fluctuates around a single peak of the fitness function. The analysis is aimed at approximating the genetic variance of the population when chromosomes are real-valued. The infinite population model is considered which allows the quasi-equilibrium state to be defined as the state when the density of chromosomes contained by the population remains unchanged over consecutive generations. This paper provides formulas for genetic diversity in the quasi-equilibrium state for fitness proportionate, tournament, and truncation selection types, with and without elitism, with Gaussian mutation, and with and without arithmetic crossover. The formulas are experimentally validated.

Index Terms—Evolutionary computing, infinite population, population diversity.

I. INTRODUCTION

POPLATION diversity is believed to have a significant impact on the effectiveness of evolutionary algorithms (EAs) in searching for the global maximum [1], [2]. Therefore, theoretical analysis of EAs includes models that define the effects of applying certain selection schemes and genetic operators, with special focus on population diversity. The concept of population diversity can be defined in different ways, see [3] for an overview for binary coded and real coded EAs, and [4] for diversity analysis in genetic programming. It seems that most popular approaches are based on certain statistics (such as mean, variance, entropy, and so on) of features that characterize individuals contained in the population. Those features include individuals' fitness, the distance between chromosomes, and the distribution of chromosomes in the search space [1]. Characterization of the last two features can be regarded as a form of expressing the population's genetic diversity. This contribution concentrates on the third feature and discusses mean and variance of chromosomes contained in populations.

When the genetic diversity of a population is known, it can be helpful to characterize the globality of the search realized by EA. Consider the example of chromosomes being real vectors. In that case, if the population distribution can be well approximated by a normal distribution with a certain mean vector \mathbf{m} and covariance matrix C , then a natural consequence

Manuscript received May 10, 2010; revised October 26, 2010, February 23, 2011, and July 15, 2011; accepted August 7, 2011. Date of publication February 10, 2012; date of current version September 27, 2012.

The author is with the Faculty of Electronics and Computer Engineering, Warsaw University of Technology, Warsaw 00-661, Poland (e-mail: jarabas@elka.pw.edu.pl).

Digital Object Identifier 10.1109/TEVC.2011.2166157

of that fact is that most of the generated chromosomes will be contained in a hyperellipsoid. Their position, size, and shape in that case will be defined by \mathbf{m} and C . Then the EA can advance into areas that contain chromosomes that have high fitness function values provided that they overlap with the aforementioned hyperellipsoid. In other words, knowledge of population diversity allows one to recognize to what degree the solution yielded by the EA can be changed with no significant chance to obtain any better chromosome. Note that understanding principles that rule the dynamics of the population distribution, and the genetic diversity in particular, is a crucial issue in the estimation of distribution algorithms (EDA) family of optimization methods [5], [6] as well as in the covariance matrix adaptation evolution strategy (CMA-ES) [7].

Analysis of diversity which concentrates on the distribution of distances between chromosomes or their fitness function values has been quite well developed—see a brief overview in Section II. Much less attention has been paid to the analysis of the dynamics of population distribution over consecutive generations. Except for the fitness proportionate selection without crossover and with Gaussian mutation, the author was unable to find appropriate references to definitions of population variance values for real-coded EAs. This contribution attempts to fill that gap in the diversity analysis. A quasi-equilibrium state is considered, when the population fluctuates over many generations in a certain area of the search space. Formulas are presented which approximate the genetic diversity of the population when the quasi-equilibrium state is reached. These formulas characterize the way in which the population diversity depends on the selection type, arithmetic crossover probability, Gaussian mutation variance, and elitism ratio.

To facilitate the analysis it is assumed that the fitness function is Gaussian, whose variance indicates its “flatness.” The formulas are not exact, therefore they are verified by comparing the predicted diversity level with the average population diversity obtained in a single EA run. The results of experiments indicate a good match between the theory presented in this paper and practice.

This paper is composed as follows. Following the short introduction in Section I, a brief overview of research on EA population diversity is provided in Section II. Section III defines the EA under consideration and gives an example simulation of the EA searching for the maximum of the Rastrigin function to motivate the work. In addition, selected issues relating to the EA model that uses an infinite population size are discussed. In Section IV, an infinite population model

is used to define the distribution of points that result from selection, assuming that the distribution prior to the selection was Gaussian with a given diversity. Then, the diversity of points after selection is used to compare selection methods. In Section V, equations approximating the diversity of points in the quasi-equilibrium state are given. These equations are introduced for each selection method separately and take into account the mutation variance, the crossover probability, and the elitist replacement. Section VI contains the results of experiments when a realistic EA was used to optimize the Gaussian function. Genetic diversity of chromosomes generated by the EA in the quasi-equilibrium state is compared to the results of the theoretical analysis and reveals a good fit of theory and practice. Section VII concludes this paper and outlines possible directions for further research. This paper includes an appendix which contains details of the proofs.

II. SELECTED EARLIER WORK ON POPULATION DIVERSITY

The issue of population diversity was raised prior to the development of evolutionary computation. Population genetics [8] elaborated several concepts that have since been adopted in EA analysis. The takeover time (see [9], [10]) is the time needed to observe the homogeneity of the population when only selection is applied and no diversity is introduced by genetic operators. Other concepts, selection intensity and selection variance [11], [12] evaluate the change in the mean and the variance of the fitness function values observed in the population before and after application of the selection. When analyzing selection intensity it is usually assumed that the distribution of fitness values is normal [13]. This assumption seems reasonable for an EA which represents chromosomes as vectors of real number, and where the fitness function is continuous. An analysis of a situation when this assumption fails is provided in [14] for tournament selection. The authors discussed the distribution of selection probabilities and reproduction rates for genetic programming. They also analyzed the case when the fitness values in the population are uniformly or quadratically distributed.

Schema analysis [15]–[17], although not focused on population diversity, is an approach that gives a feeling for the dynamics of population contents. This approach focuses on schemata which are primitives that are maintained in populations and that can be used to combine solutions. The main results are asymptotic in time and population size, i.e., it is shown that schemata contained in over-average solutions tend to grow at the expense of those that are contained in poor solutions. Diversity of a population can be characterized by proportions of different schemata that can be found in that population. Therefore, equations which describe dynamics of the number of schema instances may give an indication of diversity.

A significant progress in theoretical analysis of binary coded EAs has been made due to the Markov models. The EA is modeled as a Markov process whose state is represented by the population contents and the transition matrix is defined by the way in which selection, crossover, and mutation are defined. Vose [18] presented a deep analysis of the population contents

dynamics assuming the EA model where the population size is infinite. In works of Baake [19], [20], a more detailed model has been analyzed which treats the EA as a kind of a branching process. Genealogy of individuals is used to predict the genetic diversity of finite populations of binary chromosomes.

Another approach to diversity analysis focuses on the sampling distribution, i.e., the probability distribution which is used to generate chromosomes [21], rather than on predicting the population contents. This kind of analysis gave an inspiration for the family of EDA optimization methods where, in each step of the algorithm, the random variable defined by the sampling distribution is repeatedly realized yielding a population which is used to define the sampling distribution for the next iteration. A similar methodology is also used in CMA-ES where the sampling distribution is assumed to be Gaussian.

Qi and Palmieri [22] performed the sampling distribution analysis for the infinite population model of the EA in real space. When an infinite population size is assumed, variance of the sampling distribution becomes identical with the genetic diversity, measured as a variance of chromosomes from that population. Infinite population size allows for a formula to be defined which gives the sampling distribution in the next population when the sampling distribution in the current population is known. Nomura [23] generalized their results for arbitrary linear crossover scheme.

Not much attention has been paid to the results from Qi and Palmieri since their publication. A possible cause is the fact that they considered evolution of moments defining the infinite population distribution for a general class of fitness functions, including multimodal ones. When looking at a realistic EA with a finite population size, values of the population mean and variance may be significantly different from the values obtained with an infinite population model. In particular, when considering an example fitness function with two peaks of equal height and shape, for a range of the mutation variance, the population will occupy the neighborhood of a single peak (and its mean will be close to the top of one peak), whereas the population mean obtained for an infinite population model will be located somewhere between the two peaks. Similarly, actual population variance will be much smaller than the variance predicted for an infinite population. Detailed criticism on that issue can be found, e.g., in a work by Prügel-Bennett [24], who showed a similar example for a binary coded genetic algorithm.

Karcz-Dulęba [25] overcame these difficulties by concentrating on a unimodal fitness function, which seems to be a typical method for EA convergence analysis [26]. She considered the dynamics of an EA with fitness proportionate selection, nonelitist generational replacement, Gaussian mutation and without crossover. She used a Gaussian fitness function and derived the limiting equilibrium sampling distribution when the number of generations grows to infinity. She then generalized the results for bimodal fitness being a sum of two Gaussian functions. She proved that the stationary distribution which defines position of chromosomes after an infinite number of generations is either unimodal or bimodal, depending on the mutation variance. In the former case, population mean and variance can be well approximated by equations derived for a

single Gaussian term of the analyzed sum. In the latter case, density of points is the sum of two clusters, each connected with one Gaussian term of the fitness function. Within each cluster, mean and variance can be obtained by applying the analysis for a single Gaussian term of the fitness function.

In this paper, a Gaussian fitness function in the space of real numbers, \mathbf{R}^1 , is considered. Karcz-Dulęba's results for the fitness proportionate selection are generalized by considering the effects of the arithmetic crossover and the elitist replacement. In addition, the article covers truncation and tournament selection. In contrast to the analysis by Karcz-Dulęba, the distribution of points in the analyzed cases is not normal. Nevertheless, it is approximated by the normal distribution which allows the equations that define the dynamics of only two first central moments to be considered. These results are only approximate and therefore an experimental validation has been performed. The experimental validation demonstrated that the population diversity can be predicted by the presented formulas with an error of not more than 5% with a wide range of parameter values.

III. SUBJECT OF ANALYSIS

A. Outline of the Evolutionary Algorithm

Outline of the EA under consideration is depicted in Fig. 1. Chromosomes are real numbers and $q : \mathbf{R} \rightarrow \mathbf{R}$ is the fitness function to be maximized. In each iteration, the algorithm processes the base population P^t which contains μ individuals; we denote the i th individual by P_i^t . Individuals are reproduced which yields the population R^t .

Individuals generated by crossing over chromosomes from R^t are contained in the population C^t . The offspring population O^t is composed of mutated version of chromosomes taken from C^t and R^t , which is controlled by the crossover probability p_c . The base population for the next generation P^{t+1} contains best points selected from P^t and O^t , and the proportion between them is controlled by the elitism ratio η . In this paper, we focus on the arithmetic crossover and Gaussian mutation.

In this contribution, three selection schemes are considered which differ in their method for defining the selection probability:

- 1) fitness proportionate selection

$$P_{\text{sel}}(i|P^t) = \frac{q(P_i^t)}{\sum_{j=1,\dots,\mu} q(P_j^t)}; \quad (1)$$

- 2) tournament selection which assumes a tournament size s

$$P_{\text{sel}}(i|P^t) = \frac{1}{|S(i|P^t)|} \sum_{j \in S(i|P^t)} P_{\text{div}}(i|P^t) \quad (2)$$

where $S(i|P^t)$ is the set of individuals with the fitness function value equal to $q(P_i^t)$

$$S(i|P^t) = \{j \in \{1, \dots, \mu\} : q(P_j^t) = q(P_i^t)\} \quad (3)$$

and P_{div} is defined as

$$P_{\text{div}}(i|P^t) = \frac{1}{\mu^s} ((\mu - i + 1)^s - (\mu - i)^s); \quad (4)$$

```

 $t \leftarrow 0$ 
 $P^0 \leftarrow \text{initialization}()$ 
repeat
  for all  $i = 1, \dots, \mu$  do
    {selection with/without arithmetic crossover:}
     $j \leftarrow 1$ 
    if  $U(0, 1) < p_c$  then
       $k, l \leftarrow \text{select from } (1, \dots, \mu) \text{ with distribution } P_{\text{sel}}(\cdot|P^t)$ 
       $R_j^t \leftarrow P_k^t, R_{j+1}^t \leftarrow P_l^t$ 
      {crossover:}  $C_i^t \leftarrow (R_j^t + R_{j+1}^t)/2$ 
       $j \leftarrow j + 2$ 
    else
       $k \leftarrow \text{select from } (1, \dots, \mu) \text{ with distribution } P_{\text{sel}}(\cdot|P^t)$ 
       $C_i^t \leftarrow R_j^t \leftarrow P_k^t$ 
       $j \leftarrow j + 1$ 
    end if
    {mutation:}  $O_i^t \leftarrow C_i^t + \xi$ 
     $\{\xi \text{ is a zero mean Gaussian variate with variance } v_m\}$ 
  end for
  {elitist/nonelitist replacement:}
  if  $\eta > 0$  then
     $e(P^t) \leftarrow \text{select } \eta\mu \text{ best chromosomes from } P^t$ 
     $e(O^t) \leftarrow \text{select } (1 - \eta)\mu \text{ best chromosomes from } O^t$ 
     $P^{t+1} \leftarrow e(P^t) \cup e(O^t)$ 
  else
     $P^{t+1} \leftarrow O^t$ 
  end if
   $t \leftarrow t + 1$ 
until stop condition satisfied

```

Fig. 1. Outline of the evolutionary algorithm under consideration.

- 3) truncation selection which takes a parameter $\theta \in (0, 1)$

$$P_{\text{sel}}(i|P^t) = \frac{1}{\lceil \theta \mu \rceil} \chi_{[1, \dots, \lceil \theta \mu \rceil]}(i) \quad (5)$$

where $\chi_S(x)$ is the characteristic function of the set S

$$\chi_S(x) = \begin{cases} 1 & x \in S \\ 0 & x \notin S \end{cases} \quad (6)$$

and $\lceil x \rceil$ is the smallest integer number no smaller than x . In the case of the truncation selection and tournament selection a population is assumed to be sorted such that for each pair of numbers $i, j = 1, \dots, \mu$, $i < j$ it holds $q(P_i^t) \geq q(P_j^t)$. In this paper, a simple tournament method is considered when s individuals are sampled with replacement from the population with uniform distribution, and the best chromosome from that sample reproduces. Note that for the tournament selection the aforementioned sorting is only a convention to simplify the analysis, since in the selection process one compares fitness of s individuals and no sorting is needed.

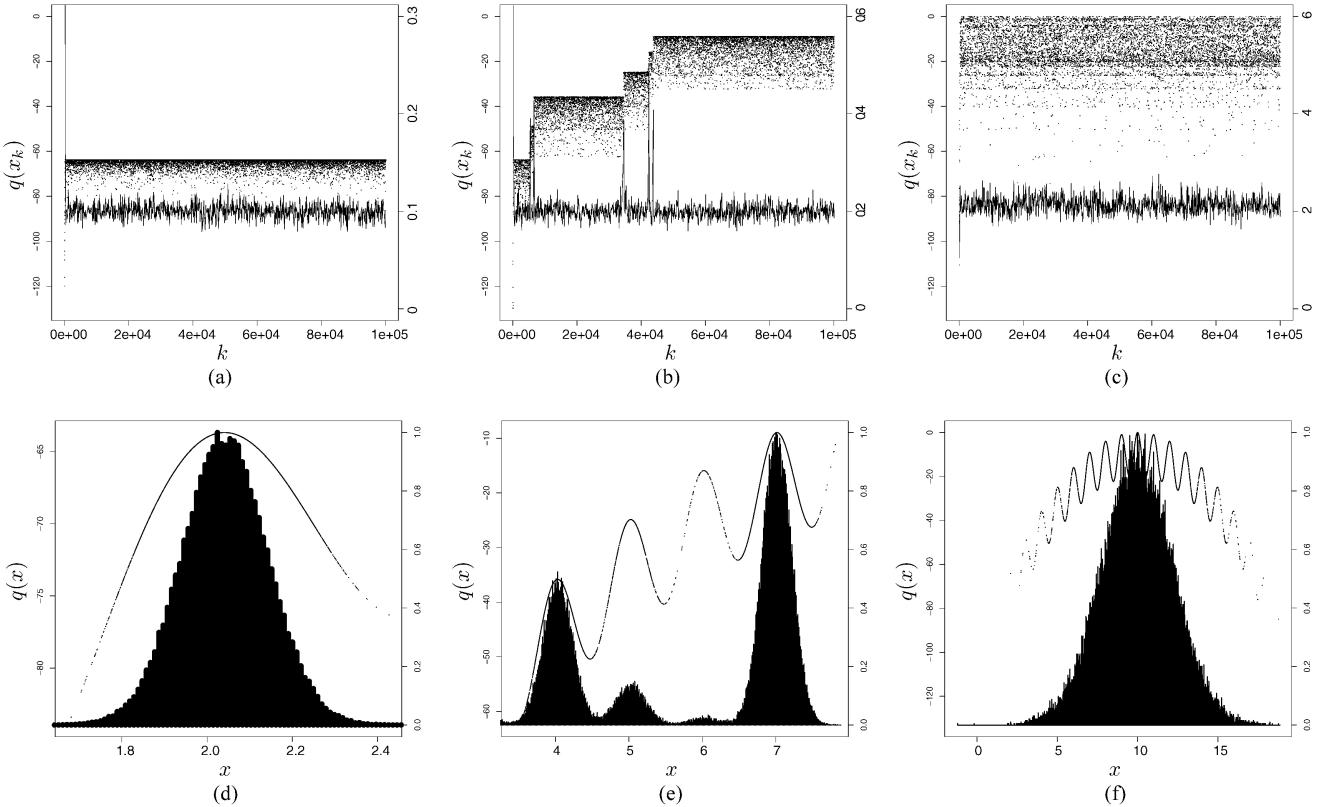


Fig. 2. Upper row of plots: fitness function values of subsequent chromosomes generated by the AE and the population variance in each generation; lower row of plots: fitness function values of chromosomes from generations 500–1000 as a function of the chromosome value, and the histogram of all chromosomes generated in generations 500–1000; results for the mutation standard deviation: (a), (d) $v_m = 0.01$, (b), (e) $v_m = 0.04$, (c), (f) $v_m = 4$.

In continuous optimization, when the domain contains no subset with nonzero measure containing points with same fitness function value, all points in populations will usually have different fitness function value. In this case, (4) can be directly used to define the selection probability.

Probability density function (pdf) and cumulated density function (cdf) of the normal distribution are given by equations

$$g_{m,v}(x) = \frac{1}{\sqrt{2\pi v}} \exp\left(-\frac{(x-m)^2}{2v}\right) \quad (7)$$

and

$$G_{m,v}(x) = \frac{1}{2} \left(1 + \operatorname{erf}\left(-\frac{x-m}{\sqrt{2v}}\right) \right) \quad (8)$$

where m is the expectation and v is the variance. For the standardized normal distribution, a simplified notation is used: $g(x) = g_{0,1}(x)$, $G(x) = G_{0,1}(x)$.

B. Example Dynamics of the EA

To motivate the theoretical analysis, consider a simple illustrative experiment where an EA from Fig. 1 was used to maximize a shifted inverted Rastrigin function in \mathbf{R}^1

$$q_{\text{Rast}}(x) = -x^2 + 10 \cos(2\pi(x-10)) - 10. \quad (9)$$

The initial population was generated randomly with the standardized normal distribution, population size was $\mu = 100$, truncation selection was used with $\theta = 1/7$, elite factor equaled $\eta = 0$ (generational replacement) and crossover probability

was $p_c = 0.7$. Three EA runs were performed which differed in the value of the mutation variance, which came from the set $v_m = 0.01, 0.04, 4$. For each setting of v_m , EA was run for 1000 generations, and each run is characterized with two plots which are depicted in Fig. 2. The upper plot shows the fitness function values of chromosomes generated by the AE in the sequence in which they are evaluated, and the population variance in each generation. The lower plot shows the fitness function values of chromosomes from generations 500–1000 as a function of the chromosome value, and the histogram of all chromosomes generated in generations 500–1000.

When looking at Fig. 2 one can conclude that after a short period of time, diversity of population starts fluctuating around a certain characteristic level, which seems to depend, *inter alia*, on the mutation variance. When v_m is small, the population will settle down in the attraction basin of a single local maximum. After v_m exceeds a certain level, it is possible to observe the dynamics which resembles a *punctuated equilibria* model [27]: long periods of stabilization, when the population stays within the attraction basin of a single local maximum, and the variance does not change much, are interleaved with very short periods of change, when the population migrates to another attraction basin. When v_m increases further the dynamics becomes similar to the case of small v_m : each population occupies attraction basins of many local maxima and their variance fluctuates around a specific level. When looking at histograms of chromosomes one can observe that when $v_m = 0.01$ and $v_m = 4$, chromosomes

generated by the EA are distributed in a way that resembles a normal distribution. When $v_m = 0.04$, the histogram resembles a mixture of normal distributions.

In this contribution, we analytically investigate the variance of the distribution of points when the population stayed in the attraction basin of a single local maximum. The analysis is performed for a Gaussian fitness function. The results of the analysis are compared with the variance of chromosomes obtained in the run of a realistic EA for the Gaussian function.

C. Infinite Population EA Model

EA generates points in a randomized way and the EA action can be described with its sampling distribution which defines the possible position of each chromosome in the population P^t ; the pdf of this distribution is denoted by f_P^t . In other words, each chromosome from P^t is a random variate whose pdf is f_P^t . Pdf of the distribution of chromosomes that have been reproduced is denoted by f_R^t , pdf of the distribution of arithmetic crossover results C^t is denoted by f_C^t , and the pdf of the distribution of chromosomes contained by the offspring population O^t is denoted by f_O^t .

Glivenko-Cantelli theorem [28] states that if the number of independent realizations of a random variable increases to infinity, the empirical cdf of the set of realizations uniformly converges to the cdf of that random variable. From that theorem it follows that if the population size goes to infinity, probability density functions f_P^t , f_R^t , f_C^t , and f_O^t can be treated as limiting densities of points observed at different phases of the EA. When the infinite population model is assumed [22], the EA transforms distributions of chromosomes in the following way.

Pdf of chromosomes that resulted from selection and arithmetic crossover is the self-convolution of distribution of chromosomes resulting from selection, since points under arithmetic crossover are independent and identically distributed (iid) random variates which share the same distribution, $f_R^t(x)$. Since mutation is independent of the mutated chromosome yet another convolution takes place and one gets finally

$$f_O^t(x) = ((1 - p_c)f_R^t(x) + p_c (f_R^t(2x) * f_R^t(2x))) * g_{0, v_m}(x) \quad (10)$$

where “**” is the symbol of convolution. Elitism in the analyzed form combines the fraction of η best parents with the fraction of $1 - \eta$ best offspring, so

$$f_P^{t+1}(x) = \chi_{L(a)} f_P^t(x) + \chi_{L(b)} f_O^t(x) \quad (11)$$

where χ is the characteristic function of a set, $L(y) = \{x | q(x) > y\}$ is a level set of the fitness function q and the values a, b can be obtained by solving

$$\int_{L(a)} f_P^t(x) dx = \eta \quad \int_{L(b)} f_O^t(x) dx = 1 - \eta. \quad (12)$$

Dependence of $f_R^t(x)$ on $f_P^t(x)$ is defined by the selection scheme [29]:

1) fitness proportionate selection

$$f_R^t(x) = A q(x) f_P^t(x); \quad (13)$$

2) tournament selection

$$f_R^t(x) = A \cdot \left(1 - \int_{L(q(x))} f_P^t(y) dy \right)^{s-1} f_P^t(x); \quad (14)$$

3) truncation selection

$$f_R^t(x) = A \cdot \chi_{L(a)}(x) f_P^t(x) \quad (15)$$

$$\text{where } \int_{L(a)} f_P^t(x) dx = \theta.$$

In all formulas A is a normalizing constant which provides that $\int_{-\infty}^{\infty} f_R^t(x) dx = 1$. The above formulas allow us to compute the distribution for the next generation f_P^{t+1} provided that we know f_R^t , the distribution of points resulting from selection. Therefore, in the next section, the analytical form of f_R^t is derived for three selection schemes: fitness proportionate, tournament, and truncation selection.

IV. DISTRIBUTION OF REPRODUCED INDIVIDUALS

In the analysis provided in this section, a Gaussian fitness function is considered. Although Gaussian fitness is not of great concern, in many cases, the fitness function can be locally approximated by the Gaussian function. According to the central limit theorem, sum of a sufficiently large number of independent random variables, each with finite mean and variance, will be approximately normally distributed. For this reason it is assumed that in the generation number t , the sampling distribution is Gaussian with mean m_P^t and variance v_P^t , i.e., $f_P^t(x) = g_{m_P^t, v_P^t}(x)$. It is possible then to bound the expected value of the distribution of points which have been selected for reproduction, m_R^t , using the following theorem.

Theorem 1: When $f_P^t(x) = g_{m_P^t, v_P^t}(x)$ and $q(x) = g_{m_q, v_q}(x)$ then

$$m_q < m_R^t < m_P^t \text{ when } m_q < m_P^t \quad (16)$$

$$m_P^t < m_R^t < m_q \text{ when } m_q > m_P^t \quad (17)$$

$$m_P^t = m_R^t = m_q \text{ when } m_q = m_P^t \quad (18)$$

for the fitness proportionate, tournament and truncation selection.

The proof is given in the Appendix.

Observe that from Theorem 1 it follows that the expected value of chromosomes that have been selected to reproduce will be closer to the value of m_q than the expected value m_P^t of the population P^t , from which these chromosomes have been selected. Crossover produces offspring chromosomes by averaging the parents, so the expected value of chromosomes before and after crossover will be equal. Since mutation is performed using zero mean normal distribution it will not change the expected value of generated chromosomes. Therefore, when the infinite population model is considered, it is possible to assume that the value of m_P^t will approach m_q in consecutive generations. For these reasons, and since the analysis is aimed at the quasi-equilibrium state, variance of the population R^t , which is denoted by v_R^t , is derived under the assumption that $m_P^t = m_q$.

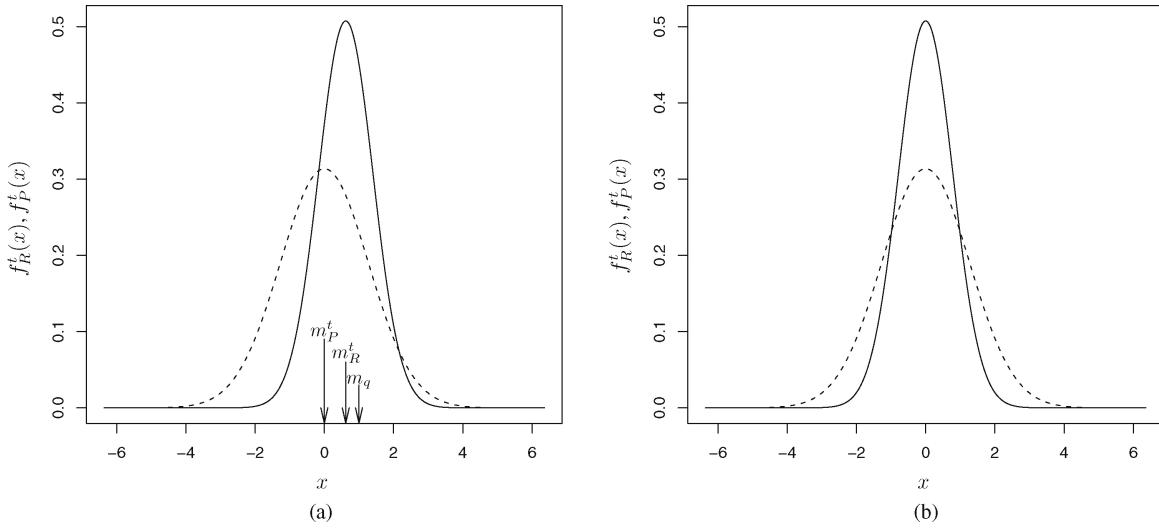


Fig. 3. Plot of distributions $f_R^t(x)$ (dashed line) and $f_P^t(x)$ (solid line) for the fitness proportionate selection, $v_q = 1, v_m = 1$. (a) $m_P^t = 0, m_q = 1$. (b) $m_P^t = m_q = 0$.

A. Fitness Proportionate Selection

Theorem 2: When $f_P^t(x) = g_{m'_P, v'_P}(x)$ and $q(x) = g_{m_q, v_q}(x)$ and the fitness proportionate selection is used then $f_R^t(x)$ is Gaussian and

$$v_R^t = v_P^t \frac{1}{1 + \frac{v'_P}{v_q}}. \quad (19)$$

Proof: From (13) it follows that

$$f_R^t(x) = A \cdot \frac{1}{2\pi\sqrt{v_q v_P^t}} \exp\left(-\frac{(x-m_q)^2}{v_q}\right) \exp\left(-\frac{(x-m_P^t)^2}{v_P^t}\right). \quad (20)$$

Further transformations lead to

$$f_R^t(x) = g_{m'_R, v'_R}(x) \quad (21)$$

where

$$m'_R = \frac{m_P^t v_q + m_q v_P^t}{v_q + v_P^t} \quad (22)$$

$$v'_R = \frac{v_q v_P^t}{v_q + v_P^t} \quad (23)$$

which proves (19). ■

Fig. 3 shows the effect of applying the fitness-proportionate selection when the population P' is normally distributed, for two cases: when $m_P^t \neq m_q$ and $m_P^t = m_q$.

B. Tournament Selection

For the tournament selection, the distribution of selected points is obtained by considering the fact that a point x which wins the tournament is selected with the distribution whose pdf is f_P^t . The other competitors should be inferior, so they must fall into the range of points with smaller fitness values. For the Gaussian fitness $q(x) = g_{m_q, v_q}(x)$, point y is worse than x provided that $y < x$ or $y > 2m_q - x$, when $x < m_q$, and $y > x$ or $y < 2m_q - x$, when $x > m_q$. Since the other competitors

are selected with the same probability distribution as x , the probability of obtaining point y worse than x can be computed by integrating f_P^t in the range where $q(y) < q(x)$. Therefore, for the Gaussian fitness $q(x) = g_{m_q, v_q}(x)$, the distribution f_R^t is defined as [29]

$$f_R^t(x) = \begin{cases} A \cdot (G_{m'_P, v'_P}(x) + 1 - G_{m'_P, v'_P}(2m_q - x))^{s-1} g_{m'_P, v'_P}(x) & x < m_q \\ A \cdot (1 - G_{m'_P, v'_P}(x) + G_{m'_P, v'_P}(2m_q - x))^{s-1} g_{m'_P, v'_P}(x) & x \geq m_q \end{cases} \quad (24)$$

where A is a normalization constant. The distribution f_R^t is not Gaussian any more, see its plots in Fig. 4 for $s = 2$.

Under the assumption that $m_P^t = m_q$, formula for f_R^t becomes much simpler and reads

$$f_R^t = \begin{cases} A \cdot (G_{m_q, v'_P}(x))^{s-1} g_{m_q, v'_P}(x), & x < m_q \\ A \cdot (1 - G_{m_q, v'_P}(x))^{s-1} g_{m_q, v'_P}(x), & x \geq m_q. \end{cases} \quad (25)$$

Formula (25) can be used to compute the variance of points after selection.

Theorem 3: When $f_P^t(x) = g_{m_q, v'_P}(x)$ and $q(x) = g_{m_q, v_q}(x)$ and the tournament selection is used with tournament size s then

$$v_R^t = \gamma(s)v_P^t \quad (26)$$

where

$$\gamma(s) = \int_{-\infty}^0 \frac{4s}{\sqrt{\pi}} y^2 (1 + \text{erf}(y))^{s-1} \exp(-y^2) dy. \quad (27)$$

Proof: Variance of the f_R^t distribution is given by

$$v_R^t = AB \quad (28)$$

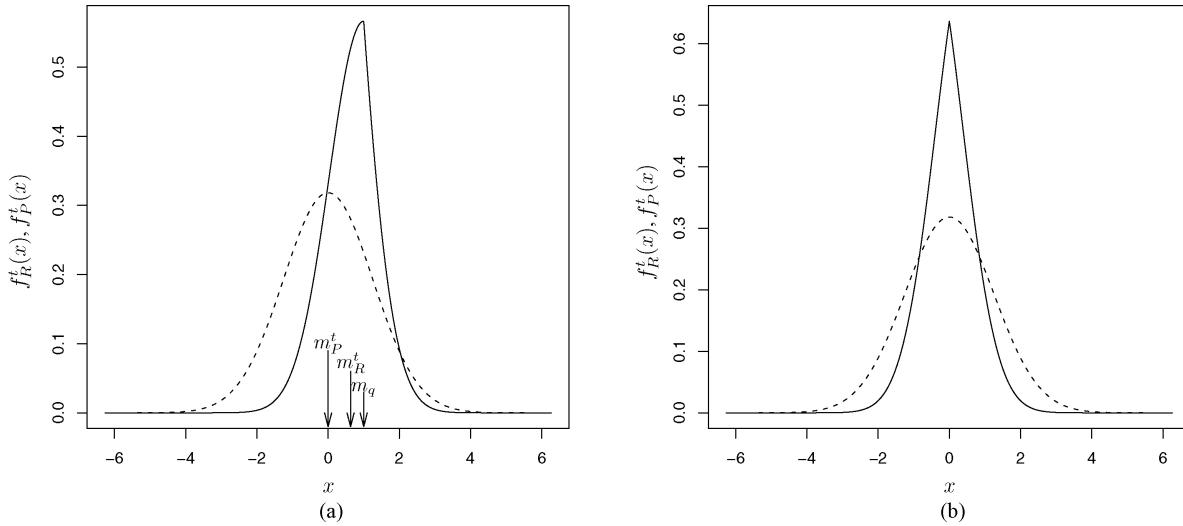


Fig. 4. Plot of distributions $f_R^t(x)$ (solid line) and $f_P^t(x)$ (dashed line) for the binary tournament selection, $v_m = 1$. (a) $m_P^t = 0, m_q = 1$. (b) $m_P^t = m_q = 0$.

where

$$\begin{aligned} 1/A &= \int_{-\infty}^{m_q} (G_{m_q, v_p^t}(x))^{s-1} g_{m_q, v_p^t}(x) dx \\ &= \frac{1}{s} (G_{m_q, v_p^t}(x))^s |_{-\infty}^{m_q} \\ &= \frac{1}{s \cdot 2^s} \end{aligned} \quad (29)$$

$$\begin{aligned} B &= \int_{-\infty}^{m_q} (x - m_q)^2 (G_{m_q, v_p^t}(x))^{s-1} g_{m_q, v_p^t}(x) dx \\ &= \frac{1}{\sqrt{v_p^t}} \int_{-\infty}^0 x^2 \left(G \left(\frac{x}{\sqrt{v_p^t}} \right) \right)^{s-1} g \left(\frac{x}{\sqrt{v_p^t}} \right) dx. \end{aligned} \quad (30)$$

Putting formulas for $g(x)$ and $G(x)$ into (30), and substituting $y = x/\sqrt{2v_p^t}$, one obtains

$$B = \frac{1}{2^{s-1}} \int_{-\infty}^0 \frac{2v_p^t y^2}{\sqrt{\pi}} (1 + \text{erf}(y))^{s-1} \exp(-y^2) dy = \frac{v_p^t \gamma(s)}{s \cdot 2^s}. \quad (31)$$

This proves (26). ■

For a tournament size $s = 2$, an exact value of $\gamma(s)$ can be given observing that

$$\int \frac{1}{\sqrt{\pi}} x^2 (1 + \text{erf}(x)) \exp(-x^2) dx = \quad (32)$$

$$\begin{aligned} &\frac{1}{8} (\text{erf}(x))^2 + \frac{1}{4} \text{erf}(x) - \frac{\exp(-2x^2)}{4\pi} \\ &- \frac{x \exp(-x^2)(1 + \text{erf}(x))}{2\sqrt{\pi}} + \text{const.} \end{aligned} \quad (33)$$

Then the value of the definite integral (27) reads

$$\int_{-\infty}^0 \frac{8}{\sqrt{\pi}} x^2 (1 + \text{erf}(x)) \exp(-x^2) dx = 1 - \frac{2}{\pi}. \quad (34)$$

For $s > 2$ we were unable to give an exact value for $\gamma(s)$, therefore we numerically estimated the value of $\gamma(s)$ for the tournament size from the range $s = 3, \dots, 100$. We

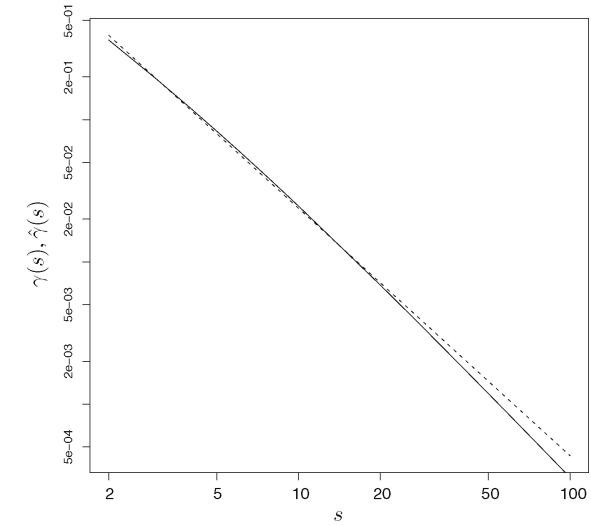


Fig. 5. Plot of dependence between the tournament size s and the values of $\gamma(s)$ (solid line) and its approximation $\hat{\gamma}(s)$ by (35)—dashed line.

observed that these estimates can be well approximated by a function

$$\hat{\gamma}(s) = 1.31s^{-1.74}. \quad (35)$$

Logarithmic scale plot of $\gamma(s)$ and $\hat{\gamma}(s)$ versus s is provided in Fig. 5. The approximation quality is acceptable: in the range $s = 3, \dots, 20$ the maximum absolute value of the approximation error equals 4.74%.

C. Truncation Selection

When the truncation selection is applied, the distribution of selected points can be obtained by truncating the distribution of points before selection to the range where a fraction of best $\theta\mu$ points is located. For the Gaussian fitness $q(x) = g_{m_q, v_q}(x)$ this range is the interval $[m_q - a, m_q + a]$ where $a > 0$ results from solving the equation

$$G_{m_p^t, v_p^t}(m_q + a) - G_{m_p^t, v_p^t}(m_q - a) = \theta. \quad (36)$$

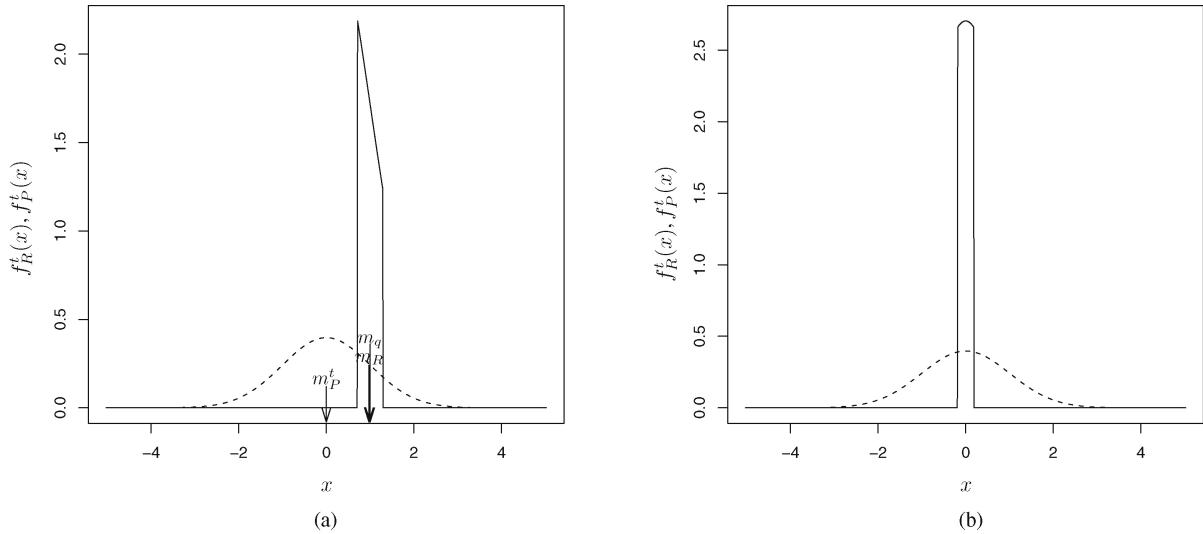


Fig. 6. Plot of distributions $f_R^t(x)$ (dashed line) and $f_P^t(x)$ (solid line) for the truncation selection, $\theta = 1/7$, $v_m = 1$. (a) $m_P^t = 0$, $m_q = 1$. (b) $m_P^t = m_q = 0$.

Thus, the distribution of selected points results reads

$$f_R^t(x) = A \cdot \chi_{[-a+m_q, a+m_q]}(x) g_{m'_P, v'_P}(x). \quad (37)$$

Observe that (37) defines the pdf of the truncated normal distribution whose example plots are depicted in Fig. 6. Formulas to compute moments of the truncated normal distribution [30] are very complex in a general case but get simpler when the truncation range is symmetrical about the expected value.

In that case, $m_q = m'_P$ and the variance of reproduced points is defined as follows.

Theorem 4: When $f_P^t(x) = g_{m_q, v'_P}(x)$ and $q(x) = g_{m_q, v_q}(x)$ and the truncation selection is used then

$$v'_R = \left(1 - \frac{2\alpha(\theta)}{\theta} \right) v'_P \quad (38)$$

where

$$\alpha(\theta) = Q\left(\frac{\theta+1}{2}\right) g\left(Q\left(\frac{\theta+1}{2}\right)\right) \quad (39)$$

and $Q(x) = G^{-1}(x)$ is the quantile function of the standardized normal distribution.

Proof: When $m_P^t = m_q$, (36) can be simplified to

$$\theta = G_{m_q, v'_P}(m_q + a) - G_{m_q, v'_P}(m_q - a) = 2G\left(\frac{a}{\sqrt{v'_P}}\right) - 1. \quad (40)$$

Then the value of a equals

$$a = Q\left(\frac{\theta+1}{2}\right) \sqrt{v'_P}. \quad (41)$$

Value of the variance v'_R can be defined using the formula for the truncated normal distribution [30]; when $m_P^t = m_q$, it equals

$$v'_R = \left(1 - \frac{\frac{a}{\sqrt{v'_P}} g\left(\frac{a}{\sqrt{v'_P}}\right) - \frac{-a}{\sqrt{v'_P}} g\left(\frac{-a}{\sqrt{v'_P}}\right)}{G\left(\frac{a}{\sqrt{v'_P}}\right) - G\left(\frac{-a}{\sqrt{v'_P}}\right)} \right) v'_P \quad (42)$$

which can be transformed to

$$v'_R = \left(1 - \frac{2\frac{a}{\sqrt{v'_P}} g\left(\frac{a}{\sqrt{v'_P}}\right)}{2G\left(\frac{a}{\sqrt{v'_P}}\right) - 1} \right) v'_P. \quad (43)$$

Putting (41) into (43) one obtains

$$v'_R = \left(1 - \frac{2Q\left(\frac{\theta+1}{2}\right) g\left(Q\left(\frac{\theta+1}{2}\right)\right)}{2G\left(Q\left(\frac{\theta+1}{2}\right)\right) - 1} \right) v'_P. \quad (44)$$

Since $G(Q(x)) = x$, (44) can be easily transformed to (38). ■

D. Comparison of Selection Methods

The selection process results in a reduction of the genetic diversity. We can interpret the coefficient $\rho = 1 - v'_R/v'_P$ as an indication of the selection pressure and characterize it using (19), (26), and (38). Note that the value of ρ is independent of the fitness function variance v_q for the tournament and truncation selection. In other words, for those methods, the only factor that drives the selection pressure is the method's parameter: truncation ratio θ or tournament size s . Fig. 7(a) shows the influence of these parameters on the value of ρ . In the case of the fitness proportionate selection, ρ is a function of the proportion v'_P/v_q of the population variance and the fitness function variance [see Fig. 7(b)].

It can be observed that the binary tournament selection reduces the diversity to a degree smaller than the truncation selection for the considered range of values of θ . The diversity reduction when the tournament size equals 10 is roughly an equivalent to the truncation parameter $\theta \approx 1/5$. For the fitness proportionate selection, when $v'_P/v_q \ll 1$, the original diversity is preserved with almost no change (at least for the infinite population), since in that case, chromosomes are roughly equally fit. When the population diversity is increasing the diversity of reproduced points gets smaller and achieves the level comparable to the truncation and tournament selection when $v'_P/v_q > 2$.

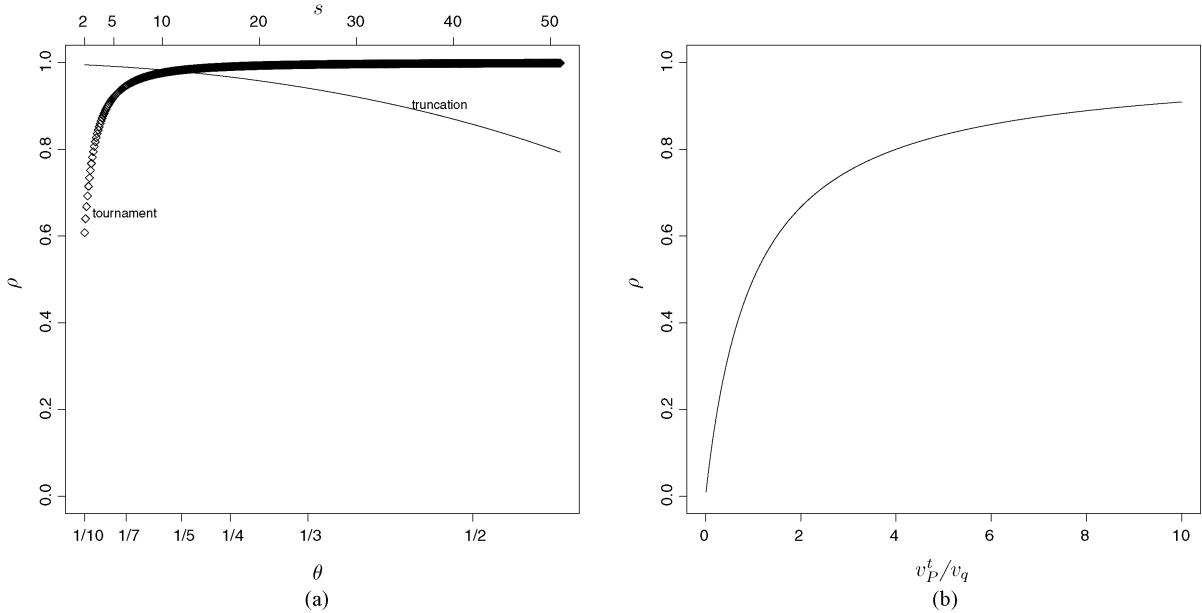


Fig. 7. Plots of the ratio of the diversity reduction due to selection. (a) Tournament and truncation selection. (b) Fitness proportionate selection.

V. LIMITING DISTRIBUTION VARIANCE

Combination of the selection process, arithmetic crossover, and Gaussian mutation yields the offspring population O^t , whose probability distribution is characterized by the pdf defined by (10). Moments of the distribution of O^t are given by

$$m_O^t = m_R^t \quad (45)$$

$$v_O^t = \left(1 - \frac{1}{2} p_c\right) v_R^t + v_m \quad (46)$$

where m_R^t, v_R^t are moments of the population of reproduced individuals which were discussed in the preceding section.

To define the distribution of individuals in the next generation, one has to take into account the replacement procedure. In this contribution it is assumed that the elitist replacement composes the population P^{t+1} of $e(P^t)$, which is the fraction of $\eta\mu$ best individuals from the population P^t , and of $e(O^t)$, which is the fraction of $(1-\eta)\mu$ best individuals from O^t . Since the analysis focuses on the limiting distribution variance it is assumed that $m_P^t = m_q$. In the analysis results obtained for the truncation selection can be easily adopted.

Best individuals from P^t span the range $(m_q - c, m_q + c)$ where

$$c = Q\left(\frac{\eta+1}{2}\right) \sqrt{v_P^t}. \quad (47)$$

Mean and variance of best individuals from P^t can be computed using results obtained for the truncation selection which yields

$$E[e(P^t)] = m_P^t \quad (48)$$

$$V[e(P^t)] = \left(1 - \frac{2\alpha(\eta)}{\eta}\right) v_P^t \quad (49)$$

where α is defined by (39). Similarly, best individuals from O^t span the range $(m_q - d, m_q + d)$ where

$$d = Q\left(\frac{(2-\eta)}{2}\right) \sqrt{v_P^{t+1}}. \quad (50)$$

Mean and variance of best individuals from O^t equals

$$E[e(O^t)] = m_O^t \quad (51)$$

$$V[e(O^t)] = \left(1 - \frac{2\alpha(1-\eta)}{1-\eta}\right) v_O^t. \quad (52)$$

Therefore, mean and variance of the mixture of best points from O^t and P^t is given by

$$m_P^{t+1} = m_P^t \quad (53)$$

$$v_P^{t+1} = (\eta - 2\alpha(\eta)) v_P^t + (1 - \eta - 2\alpha(1-\eta)) v_O^t. \quad (54)$$

Equations (46) and (54) are analyzed for each selection scheme to define the limiting variance of the sampling distribution v_∞ .

A. Formulas for the Limiting Distribution Variance

Observation 1: When $q(x) = g_{m_q, v_q}(x)$ and the fitness proportionate selection is used then

$$\begin{aligned} v_\infty \approx & \frac{v_m}{2\beta(\eta)} \left(1 - \left(\beta(\eta) - 1 + \frac{p_c}{2}\right) \frac{v_q}{v_m}\right) \\ & + \sqrt{\left(1 - \left(\beta(\eta) - 1 + \frac{p_c}{2}\right) \frac{v_q}{v_m}\right)^2 + 4 \frac{v_q}{v_m} \beta(\eta)} \end{aligned} \quad (55)$$

where

$$\beta(\eta) = \frac{1 - \eta + 2\alpha(\eta)}{1 - \eta - 2\alpha(1 - \eta)} \quad (56)$$

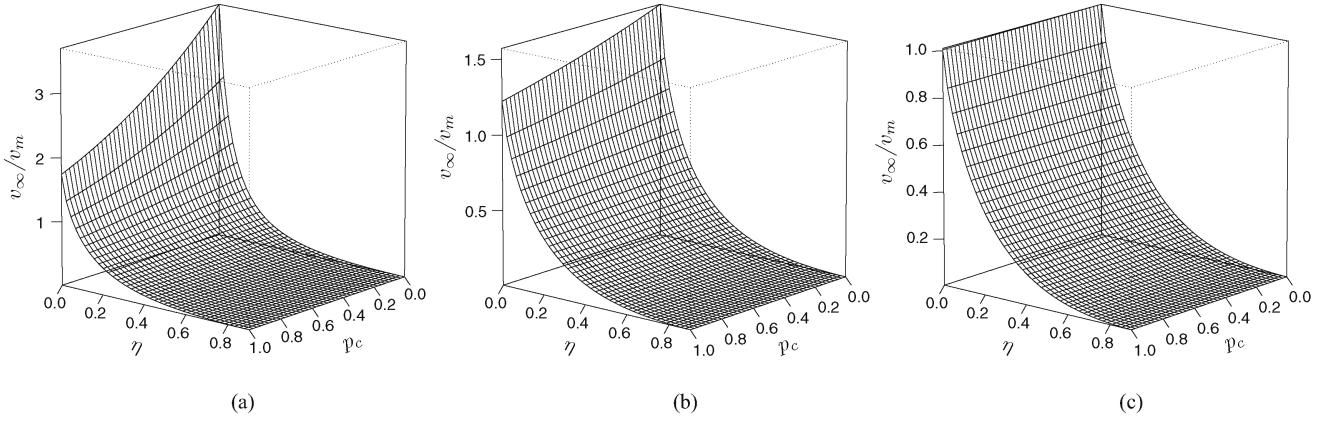


Fig. 8. Plots of the proportion v_∞/v_m as a function of the arithmetic crossover probability p_c and elite fraction η , for the following selection methods. (a) Fitness proportionate, $v_m/v_q = 0.1$. (b) Binary tournament. (c) Truncation, $\theta = 1/7$.

and α is given by (39).

Proof: Combination of (19), (46), and (54) leads to

$$\begin{aligned} v_P^{t+1} &\approx (\eta - 2\alpha(\eta)) v_P^t + (1 - \eta - 2\alpha(1 - \eta)) \\ &\quad \left(\left(1 - \frac{p_c}{2}\right) \frac{v_q v_P^t}{v_q + v_P^t} + v_m \right). \end{aligned} \quad (57)$$

Putting v_∞ into (57) instead of v_P^t and v_P^{t+1} yields

$$v_\infty \approx \frac{1}{\beta(\eta)} \left(\left(1 - \frac{p_c}{2}\right) \frac{v_q v_\infty}{v_q + v_\infty} + v_m \right). \quad (58)$$

Solution to (58) is given by (55). ■

Observation 2: When $q(x) = g_{m_q, v_q}(x)$ and the tournament selection is used then

$$v_\infty \approx \frac{1}{\beta(\eta) + (p_c/2 - 1)\gamma(s)} v_m \quad (59)$$

where $\gamma(s)$ is given by (27).

Proof: Combining (26), (46), and (54) leads to the following formula:

$$\begin{aligned} v_P^{t+1} &\approx (\eta - 2\alpha(\eta)) v_P^t + (1 - \eta - 2\alpha(1 - \eta)) \\ &\quad \left(\left(1 - \frac{p_c}{2}\right) v_P^t \gamma(s) + v_m \right). \end{aligned} \quad (60)$$

Assuming $v_P^{t+1} = v_P^t = v_\infty$ and solving for v_∞ one gets (59). ■

In the binary tournament case, a more accurate formula can be obtained from (59) by putting $(1 - 2/\pi)$ instead of $\gamma(s)$, which yields

$$v_\infty \approx \frac{2\pi}{2\pi(\beta(\eta) - 1) + (\pi - 2)p_c + 4} v_m. \quad (61)$$

Observation 3: When $q(x) = g_{m_q, v_q}(x)$ and the truncation selection is used then

$$v_\infty \approx \frac{2\theta}{2\theta(\beta(\eta) - 1) + (\theta - 2\alpha(\theta))p_c + 4\alpha(\theta)} v_m. \quad (62)$$

Proof: From (38), (46), and (54) it follows that

$$\begin{aligned} v_P^{t+1} &\approx (\eta - 2\alpha(\eta)) v_P^t + (1 - \eta - 2\alpha(1 - \eta)) \\ &\quad \left(\left(1 - \frac{p_c}{2}\right) \left(1 - \frac{2\alpha(\theta)}{\theta}\right) v_P^t + v_m \right). \end{aligned} \quad (63)$$

In the equilibrium state one obtains

$$\begin{aligned} v_\infty &\left(1 - \eta - 2\alpha(\eta) - (1 - \eta - 2\alpha(1 - \eta)) \right. \\ &\quad \left. \left(1 - \frac{2\alpha(\theta)}{\theta} + p_c \left(\frac{\alpha(\theta)}{\theta} - \frac{1}{2} \right) \right) \right) \\ &\approx v_m (1 - \eta - 2\alpha(1 - \eta)) \end{aligned} \quad (64)$$

which proves (62). ■

B. Discussion

Formulas (55), (59), and (62) define the value of the limiting sampling distribution variance v_∞ for three selection methods under consideration. In all cases, v_∞ is a function of the mutation variance v_m , arithmetic crossover probability p_c and elitism ratio η . Observe that for the tournament and truncation selection, v_∞ is a linear function of v_m , and the dependence of v_∞ on p_c and η is slightly more complicated. For the tournament and truncation selection, value of v_∞ is influenced by an appropriate selection parameter (s or θ). When the selection is fitness proportionate, the role of the tuning parameter is taken by the proportion v_m/v_q .

Better understanding of the relationship between v_∞ , p_c , and η can be achieved by analyzing Fig. 8 where the value of v_∞/v_m is plotted as a function of p_c and η , assuming $s = 2$, $\theta = 1/7$, and $v_m/v_q = 0.1$. A general impression is that both arithmetic crossover and elitist replacement result in a decrease in the population diversity.

It can be observed however that for the investigated settings, selection methods differ in their sensitivity to the arithmetic crossover with fitness proportional selection most sensitive and truncation selection only slightly influenced by the arithmetic crossover. The influence of the elitism ratio η on the limiting variance v_∞ is very strong for all selection methods under comparison. When $\eta \approx 0.01$ (which is equivalent to having a single point in the elite and $\mu = 100$ individuals in the population) one can observe that the diversity is reduced by approximately 20% in comparison to the nonelitist case.

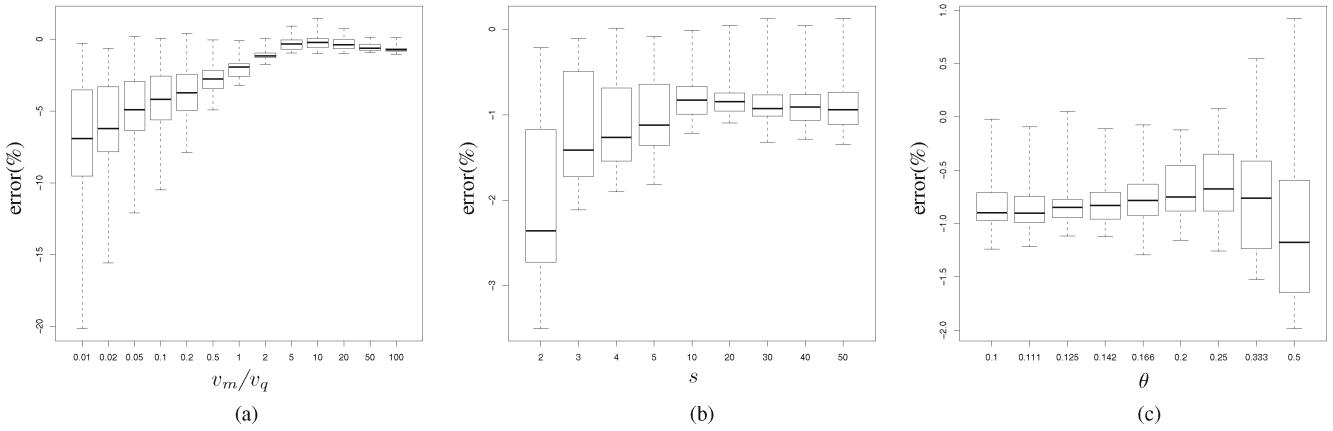


Fig. 9. Box-and-whisker plots of the error of approximating the real variance of chromosomes by the limiting variance value for the infinite population model and the selection method being (a) fitness proportionate, (b) tournament, and (c) truncation; each element of the diagram represents first quartile and third quartile (limits of the box), median (horizontal line in the middle of the box) and minimum and maximum (extreme lines of whiskers) of error values observed for a given value of a selection method parameter.

VI. EXPERIMENTAL VALIDATION

Approximate equations for the limiting population variance have been obtained under two assumptions: that the population size μ is infinite, and that the sampling distributions f_P^t is Gaussian for each t . In the realistic EA neither of these assumptions is fulfilled which raises the question about the approximation quality. This section provides results of experiments which give empirical evidence to answer this question. Each experiment was performed for the Gaussian fitness function with zero mean and variance v_q . Initial population was generated from the standardized normal distribution. During the experiment, for each combination of selection method and parameter values, 100 independent EA runs were performed. EA was run for 10 000 generations and all chromosomes contained in generations 5000–10 000 were taken as a sample whose variance was computed. During tests, population size was assumed $\mu = 100$, mutation variance was $v_m = 1$, crossover probability p_c came from the set $\{0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1\}$ and the elite fraction η from the set $\{0, 0.01, 0.02, 0.05, 0.1, 0.2\}$. For each choice of the selection method and each selection parameter value all possible combinations of p_c and η were tested which produced 66 values of the mean variance of chromosomes. Each mean value of the variance v_{emp} was compared to the limiting variance value predicted for the infinite population model v_∞ and the relative error was computed as $(v_\infty - v_{\text{emp}})/v_{\text{emp}}$. Box-and-whisker plots of the mean error are depicted in Fig. 9.

The results show that the actual variance of generated chromosomes is usually underestimated by the value obtained for the infinite population model. This effect can be interpreted by the fact that a finite population will fluctuate around the fitness function maximum which will result in denser sampling in tails of the distribution of chromosomes obtained with the infinite population model. This effect is particularly strong for the fitness proportionate selection when the mutation variance is small in comparison to the variance of the fitness function. Then, if all chromosomes from one population were located close to the peak of the fitness function and their variance was equal the variance

obtained with infinite population model, the distribution of selection probabilities would be almost uniform. This would effect in increasing diversity of the offspring population.

It can be also observed that, for the tournament and truncation selection, infinite population model yields a very good approximation to the actual variance of chromosomes generated by the EA. In the whole range of parameter settings the approximation error never exceeded $\pm 5\%$ and the error median was about -1% (with one exception for the binary tournament where it was nearly -2.5%). Such agreement of theory and practice was also observed for the fitness proportionate selection when the ratio v_m/v_q ranged between 0.5 and 100. When the ratio v_m/v_q decreases below 0.5, fitness function values “observed” by the population become very little diversified which results in a significantly higher population diversity than expected from the theory.

VII. CONCLUSION

This paper filled a gap in the genetic diversity analysis of the EA in the quasi-equilibrium state, i.e., when the population fluctuates around a single local maximum of the fitness function. Formulas are provided that define the variance of chromosomes in the quasi-equilibrium state when the fitness function is Gaussian and chromosomes are real numbers. It is shown that the formulas, although approximate, yield diversity values that are quite close to the actual ones when the Gaussian fitness function is used. In the case of the truncation and tournament selection the results can be directly applied to predict the diversity of chromosomes for any fitness function which is strictly concave and even. It is possible to generalize the results which were obtained for the fitness proportionate selection to an arbitrary Gaussian fitness in R^n when $n > 1$. In that case, instead of the variance of chromosomes, a whole covariance matrix can be predicted.

Obtained results shed some light on tuning EA parameters also for a realistic case when the fitness function is multimodal. If the mutation variance is sufficiently small then the EA will tend to stay for a number of generations in a quasi-equilibrium state when the population is contained in

the attraction basin of a single local maximum of the fitness. This quasi-equilibrium may become a trap when the mutation variance is too small. There exists a threshold variance value which should be exceeded to allow the population to leave the quasi-equilibrium. Another quasi-equilibrium can, however, be characterized with a different threshold variance. These observations support the tendency to perform a kind of adaptation of the mutation variance, which fits into the major research direction in contemporary EAs where adaptation rules of EA parameter values are developed [31]. The results presented here could be used to formulate yet another adaptation rule, which compares the population diversity values which have been predicted and which are observed in current populations, to tune not only the mutation variance, but also the arithmetic crossover probability, the elitism ratio and the selection method parameter.

Presented study is limited to a specific case of an EA which uses non-adaptive Gaussian mutation, arithmetic crossover that produces a single offspring chromosome, and a specific elitist replacement. It also covers only three basic selection schemes. For future work it would be interesting to investigate some other methods to realize EA steps, like steady-state replacement, different crossover schemes, multiple individuals resulting from crossover and multiparent crossover, different mutation method, and so on. An important issue which needs investigation is to analyze the effects introduced by the space dimensionality. Yet another promising research direction seems to study the effects that will appear when the fitness function is not even and when the mutation variance is so large that in the quasi-equilibrium state, the population covers the attraction basins of many local maxima.

ACKNOWLEDGMENT

Integration was performed with WolframAlpha (<http://integrals.wolfram.com>). Numerical experiments were performed and analyzed under the R package (<http://cran.r-project.org>).

APPENDIX PROOF OF THE THEOREM 1

Without generality loss, in further considerations it is assumed that $m_q = 0$ and $m_p^t < 0$. To prove Theorem 1 under the aforementioned assumptions, it is sufficient to show that it holds $m_p^t < m_R^t < 0$.

Proof of Theorem 1 for the Fitness Proportionate Selection: Value of v_R^t is defined by (22). Since $v_q > 0$, $v_p^t > 0$, and $m_p^t < 0$ then it holds $m_R^t > m_p^t$ and $m_R^t < 0$. ■

Proof of Theorem 1 for the Tournament Selection: Value of m_R^t can be computed using the formula for the pdf of the population R^t

$$m_R^t = \frac{\int_{-\infty}^{\infty} x f_R^t(x) dx}{\int_{-\infty}^{\infty} f_R^t(x) dx}. \quad (65)$$

Formula (65) can be rewritten as

$$m_R^t = \frac{\int_0^0 x (G_{m_p^t, v_p^t}(x) + 1 - G_{m_p^t, v_p^t}(-x)) g_{m_p^t, v_p^t}(x) dx + \int_0^{\infty} x (1 - G_{m_p^t, v_p^t}(x) + G_{m_p^t, v_p^t}(-x)) g_{m_p^t, v_p^t}(x) dx}{\int_{-\infty}^0 (G_{m_p^t, v_p^t}(x) + 1 - G_{m_p^t, v_p^t}(-x)) g_{m_p^t, v_p^t}(x) dx + \int_0^{\infty} (1 - G_{m_p^t, v_p^t}(x) + G_{m_p^t, v_p^t}(-x)) g_{m_p^t, v_p^t}(x) dx}. \quad (66)$$

Taking into account that $G_{m, v}(x) = 1 - G_{-m, v}(-x)$ one obtains

$$m_R^t = \frac{\int_{-\infty}^0 x H_{m_p^t, v_p^t}(x) g_{m_p^t, v_p^t}(x) dx + \int_0^{\infty} x H_{m_p^t, v_p^t}(-x) g_{m_p^t, v_p^t}(x) dx}{\int_{-\infty}^0 H_{m_p^t, v_p^t}(x) g_{m_p^t, v_p^t}(x) dx + \int_0^{\infty} H_{m_p^t, v_p^t}(-x) g_{m_p^t, v_p^t}(x) dx} \quad (67)$$

where $H_{m_p^t, v_p^t}(x) = G_{m_p^t, v_p^t}(x) + G_{-m_p^t, v_p^t}(x)$.

Note that when $m_p^t < 0$ then $|x - m_p^t| < |x + m_p^t|$ for $x < 0$. Therefore $g_{m_p^t, v_p^t}(x) > g_{-m_p^t, v_p^t}(x)$. Since $H_{m_p^t, v_p^t}(x) > 0$ for all x and the integration is made for $x < 0$ therefore $B > 0$. Since $C < 0$, it can be concluded that $m_R^t < 0$.

To prove that $m_R^t > m_p^t$, observe that

$$\begin{aligned} & \int_{-\infty}^0 x H_{m_p^t, v_p^t}(x) g_{m_p^t, v_p^t}(x) dx + \\ & \int_0^{\infty} x H_{m_p^t, v_p^t}(-x) g_{m_p^t, v_p^t}(x) dx = A + B + C \end{aligned} \quad (68)$$

where

$$\begin{aligned} A &= \int_{-\infty}^{m_p^t} x H_{m_p^t, v_p^t}(x) g_{m_p^t, v_p^t}(x) dx + \\ & \int_{m_p^t}^{\infty} x H_{m_p^t, v_p^t}(2m_p^t - x) g_{m_p^t, v_p^t}(x) dx \end{aligned} \quad (69)$$

$$B = \int_{m_p^t}^0 x (H_{m_p^t, v_p^t}(x) - H_{m_p^t, v_p^t}(2m_p^t - x)) g_{m_p^t, v_p^t}(x) dx \quad (70)$$

$$C = \int_0^{\infty} x (H_{m_p^t, v_p^t}(-x) - H_{m_p^t, v_p^t}(2m_p^t - x)) g_{m_p^t, v_p^t}(x) dx. \quad (71)$$

Observe also that

$$\begin{aligned} & \int_{-\infty}^0 H_{m_p^t, v_p^t}(x) g_{m_p^t, v_p^t}(x) dx + \\ & \int_0^{\infty} H_{m_p^t, v_p^t}(-x) g_{m_p^t, v_p^t}(x) dx = D + E + F \end{aligned} \quad (72)$$

where

$$\begin{aligned} D &= \int_{-\infty}^{m_p^t} H_{m_p^t, v_p^t}(x) g_{m_p^t, v_p^t}(x) dx + \\ & \int_{m_p^t}^{\infty} H_{m_p^t, v_p^t}(2m_p^t - x) g_{m_p^t, v_p^t}(x) dx \end{aligned} \quad (73)$$

$$E = \int_{m_p^t}^0 (H_{m_p^t, v_p^t}(x) - H_{m_p^t, v_p^t}(2m_p^t - x)) g_{m_p^t, v_p^t}(x) dx \quad (74)$$

$$F = \int_0^{\infty} (H_{m_p^t, v_p^t}(-x) - H_{m_p^t, v_p^t}(2m_p^t - x)) g_{m_p^t, v_p^t}(x) dx. \quad (75)$$

For all $m_p^t < x < 0$ it holds $2m_p^t - x < m_p^t < x$ and $H_{m_p^t, v_p^t}(x) > H_{m_p^t, v_p^t}(2m_p^t - x)$, therefore $B < 0$ and $E > 0$. For all $x > 0$ it holds $2m_p^t - x < -x$ and $H_{m_p^t, v_p^t}(-x) > H_{m_p^t, v_p^t}(2m_p^t - x)$, so $C > 0$ and $F > 0$.

Note that $A/D = m_p^t$ and $B > m_p^t E$. For that reason it can be concluded that $m_R^t = (Dm_p^t + B + C)/(D + E + F) > (Dm_p^t + m_p^t E + C)/(D + E + F) > m_p^t$.

This proves that $m_R^t > m_p^t$. ■

Proof of Theorem 1 for the Truncation Selection: Value of m_R^t is defined as the expected value of the truncated normal distribution

$$m_R^t = \frac{\int_{-a}^{+a} x g_{m_p^t, v_p^t}(x) dx}{\int_{-a}^{+a} g_{m_p^t, v_p^t}(x) dx} \quad (76)$$

which equals [30]

$$m_R^t = m_p^t - \frac{g_{m_p^t, v_p^t}(a) - g_{m_p^t, v_p^t}(-a)}{G_{m_p^t, v_p^t}(a) - G_{m_p^t, v_p^t}(-a)} \sqrt{v_p^t}. \quad (77)$$

From the definition of cdf it follows that $G_{m_p^t, v_p^t}(-a) < G_{m_p^t, v_p^t}(a)$ provided that $a > 0$. If $m_p^t < 0$, then $| -a - m_p^t | < | a - m_p^t |$ and $g_{m_p^t, v_p^t}(-a) > g_{m_p^t, v_p^t}(a)$, so $m_R^t > m_p^t$.

To prove that $m_R^t < 0$ observe that (76) can be expressed as

$$m_R^t = \frac{\int_{-a}^0 x g_{m_p^t, v_p^t}(x) dx + \int_0^a x g_{m_p^t, v_p^t}(x) dx}{\int_{-a}^{+a} g_{m_p^t, v_p^t}(x) dx}. \quad (78)$$

Since $g_{m_p^t, v_p^t}(-x) = g_{-m_p^t, v_p^t}(x)$, after some transformations one obtains

$$m_R^t = \frac{\int_{-a}^0 x (g_{m_p^t, v_p^t}(x) - g_{-m_p^t, v_p^t}(x)) dx}{\int_{-a}^a g_{m_p^t, v_p^t}(x) dx} = \frac{A}{B}. \quad (79)$$

If $x < 0$ and $m_p^t < 0$ then $g_{m_p^t, v_p^t}(x) > g_{-m_p^t, v_p^t}(x)$ and $A < 0$, $B > 0$. This proves that $m_R^t < 0$. ■

REFERENCES

- [1] C. Mattiussi, M. Waibel, and D. Floreano, "Measures of diversity for populations and distances between individuals with highly reorganizable genomes," *Evol. Comput.*, vol. 12, no. 4, pp. 495–515, 2004.
- [2] S. W. Mahfoud, "A comparison of parallel and sequential niching methods," in *Proc. Int. Conf. Genet. Algorithms*, 1995, pp. 136–143.
- [3] R. Morrison and K. de Jong, "Measurement of population diversity," in *Proc. 5th Eur. Conf. Artif. Evol.*, 2001, pp. 31–41.
- [4] S. Gustafson, "An analysis of diversity in genetic programming," Ph.D. dissertation, School Comput. Sci., Univ. Nottingham, Nottingham, U.K., 2004.
- [5] J. Lozano, P. Larrañaga, I. Inza, and E. Bengoetxea, Eds., *Toward a New Evolutionary Computation. Advances in Estimation of Distribution Algorithms*. Berlin, Germany: Springer, 2006.
- [6] J. Lozano and P. Larrañaga, Eds., *Estimation of Distribution Algorithms. A New Tool for Evolutionary Computation*. Dordrecht, The Netherlands: Kluwer, 2002.
- [7] N. Hansen, "The CMA evolution strategy: A comparing review," in *Toward a New Evolutionary Computation. Advances on Estimation of Distribution Algorithms*, J. Lozano, P. Larrañaga, I. Inza, and E. Bengoetxea, Eds. Berlin, Germany: Springer, 2006, pp. 75–102.
- [8] S. Wright, *Evolution and the Genetics of Populations: Genetics and Biometric Foundations*. Chicago, IL: Univ. Chicago Press, 1984.
- [9] D. E. Goldberg and K. Deb, "A comparative analysis of selection schemes used in genetic algorithms," in *Foundations of Genetic Algorithms*. San Francisco, CA: Morgan Kaufmann, 1991, pp. 69–93.
- [10] T. Bäck, *Evolutionary Algorithms in Theory and Practice*. Oxford, U.K.: Oxford Univ. Press, 1996.
- [11] H. Mühlenbein and D. Schlierkamp-Vosen, "The science of breeding and its application to the breeder genetic algorithm," *Evol. Comput.*, vol. 1, no. 4, pp. 335–360, 1993.
- [12] T. Bickle, "Theory of evolutionary algorithms and application to system synthesis," Ph.D. dissertation, Dept. Inform. Technol. Elect. Eng., Swiss Federal Inst. Technol., Zurich, Switzerland, 1996.
- [13] T. Bickle and L. Thiele, "A mathematical analysis of tournament selection," in *Proc. Int. Conf. Genet. Algorithms*, 1995, pp. 9–16.
- [14] H. Xie, M. Zhang, and P. Andreae, "Another investigation on tournament selection: Modeling and visualization," in *Proc. Genet. Evol. Comput. Conf.*, 2007, pp. 1468–1475.
- [15] C. Stevens and H. Walbroeck, "Schemata evolution and building blocks," *Evol. Comput.*, vol. 7, no. 2, pp. 109–124, 1999.
- [16] D. Goldberg, *Genetic Algorithms in Search, Optimization and Machine Learning*. Reading, MA: Addison-Wesley, 1989.
- [17] N. Radcliffe, "Forma analysis and random respectful recombination," in *Proc. Int. Conf. Genet. Algorithms*, 1991, pp. 222–229.
- [18] M. Vose, *The Simple Genetic Algorithm: Foundations and Theory*. Cambridge, MA: MIT Press, 1999.
- [19] E. Baake and H.-O. Georgii, "Mutation, selection, and ancestry in branching models: A variational approach," *J. Math. Biol.*, vol. 54, no. 2, pp. 257–303, Feb. 2007.
- [20] E. Baake and R. Bialowons, "Ancestral processes with selection: Branching and Moran models," *Banach Center Publications*, vol. 80, pp. 33–52, 2008.
- [21] H. Mühlenbein and T. Mahnig, "Evolutionary algorithms: From recombination to search distributions," in *Theoretical Aspects of Evolutionary Computing*. Berlin, Germany: Springer, 2000, pp. 137–176.
- [22] X. Qi and F. Palmieri, "Theoretical analysis of evolutionary algorithms with an infinite population size in continuous space. Part I: Basic properties of selection and mutation," *IEEE Trans. Neural Netw.*, vol. 5, no. 1, pp. 102–119, Jan. 1994.
- [23] T. Nomura, "An analysis on crossovers for real number chromosomes in an infinite population size," in *Proc. Int. Joint Conf. Artif. Intell.*, 1997, pp. 936–941.
- [24] A. Prügel-Bennett, "Modeling finite populations," in *Foundations of Genetic Algorithms*. San Francisco, CA: Morgan Kaufmann, 2002, pp. 99–114.
- [25] I. Karcz-Dulęba, "Dynamics of infinite populations evolving in a landscape of uni and bimodal fitness functions," *IEEE Trans. Evol. Comput.*, vol. 5, no. 4, pp. 398–409, Aug. 2001.
- [26] H.-G. Beyer, H.-P. Schwefel, and I. Wegener, "How to analyze evolutionary algorithms," *Theor. Comput. Sci.*, vol. 287, no. 1, pp. 101–130, Sep. 2002.
- [27] N. Eldredge and S. Gould, "Punctuated equilibria: An alternative to phyletic gradualism," in *Models in Paleobiology*, T. Shopf, Ed. San Francisco, CA: Freeman, 1972, pp. 82–115.
- [28] D. Pollard, *Convergence of Stochastic Processes*. Berlin, Germany: Springer, 1984.
- [29] T. Bäck, D. Fogel, and Z. Michalewicz, Eds., *Handbook of Evolutionary Computation*. Oxford, U.K.: Oxford Univ. Press, 1996.
- [30] N. Johnson and S. Kotz, *Continuous Univariate Distributions*, vol. 1. New York: Wiley, 1970.
- [31] F. Lobo, C. Lima, and Z. Michalewicz, Eds., "Parameter setting in evolutionary algorithms," in *Studies in Computational Intelligence*, vol. 54. Berlin, Germany: Springer, 2007.



Jarosław Arabas was born in Poland in 1970. He received the M.Sc., Ph.D., and D.Sc. degrees from the Warsaw University of Technology (WUT), Warsaw, Poland, in 1993, 1996, and 2005, respectively.

Since 1993, he has been with the Faculty of Electronics and Computer Engineering, WUT, where he is currently a Professor. His main areas of research interests include evolutionary computation, optimization methods, and learning systems. Since 1995, he has been consulting on applications of these methods for power plant control, decision making, and risk management in electricity markets.