Soft Computing
https://doi.org/10.1007/s00500-025-10652-6

FOUNDATION, ALGEBRAIC, AND ANALYTICAL METHODS IN SOFT COMPUTING

Analysis of selection noise in genetic algorithms

Nataliya M. Gulayeva[1,2] · **Joaquín Borrego-Díaz**[2] · **Fernando Sancho-Caparrini**[2]

Accepted: 18 January 2025

Abstract

Selection is often considered as a fundamental force in the evolutionary process. Genetic drift, or selection noise, is an important characteristic of selection methods. It has a direct effect on the performance of genetic algorithms. In this paper, a brief review of methods to analyze genetic drift is given, and known estimations of selection noise of various selection schemes used in genetic algorithms are presented. After that, genetic drift of widely used proportional, ranking, and tournament selection schemes is thoroughly studied. To this end, two new measures for selection noise analysis are proposed, namely the noise takeover time and pure reproduction rate. Using these measures, the effect of population size, chromosome length, and selection scheme parameters on genetic drift is analyzed. Also, selection schemes known as being selection pressure equivalent are tested for selection noise equivalence. Both theoretical and experimental approaches are used for the analysis. The results obtained are presented in tabular form. Wherever possible, it is indicated whether the obtained results are identical or different from the results of previous studies. Since no comprehensive study of selection noise has been conducted previously, this indication concerns only some of the results. Although our results differ at some points from those presented earlier, they are consistent on both measures.

Keywords Genetic algorithm · Genetic drift · Selection noise · Proportional selection · Tournament selection · Ranking selection

1 Introduction

Genetic drift, also called *genetic sampling error*, is a well-known phenomenon in population genetics, referring to changes in the gene pool that take place strictly by chance; drift happens faster in small populations than in larger ones (Masel 2011). Genetic drift phenomenon is observed in genetic algorithms due to the stochastic nature of selection operators. In a finite size population, a random selection among individuals of equal fitness leads to a disproportion between the expected and actual number of copies of an individual in the mating pool; in other words, random selection adds noise to the expected number of copies of an individual. This explains why, in genetic algorithms, genetic drift is also referred to as *selection noise*.

At low selection pressure, genetic drift can lead to a loss of population diversity and, consequently, to premature convergence of the algorithm, since operation of the algorithm is virtually independent of fitness of individuals. Reducing selection noise is one of the population diversification techniques aiming at preventing premature convergence of genetic algorithms (Hill and O'Riordan 2016). On the other hand, evolutionary algorithms may benefit from the stochasticity in selection process (Branke and Schmidt 2003; Carvalho and Nolfi 2023; Hancock 1997). Therefore, an important task is not to eliminate selection noise but to control its level. To provide mechanisms of such control, we need to understand how various parameters of genetic algorithms affect selection noise. Since selection noise is primarily a property of selection scheme, a tool to estimate selection noise regarding selection scheme parameters as well as genetic algorithm parameters is needed, and estimation of selection noise for various parameter combinations should be done.

✉ Nataliya M. Gulayeva
gulayeva@ukma.edu.ua

[1] Department of Informatics, National University of Kyiv-Mohyla Academy, Kyiv, Ukraine

[2] Departamento de Ciencias de la Computación e Inteligencia Artificial, Universidad de Sevilla, E.T.S. Ingeniería Informática, Avda. Reina Mercedes s/n, 41012 Sevilla, Spain

In this paper, we introduce two measures to estimate selection noise property of selection schemes. Using these measures, we study the effect of population size, chromosome length, and selection scheme parameters on genetic drift of various proportional, ranking, and tournament selection schemes. We also compare selection noise of selection schemes known as being selection pressure equivalent. We summarize obtained results in tabular form. Wherever possible, we indicate whether our results coincide with the results of previous works or differ from them. This indication applies only to some of our results, since no comprehensive study of selection noise has been conducted previously.

The rest of this paper is organized as follows. Section 2 provides a brief review of methods used to study selection noise and results obtained earlier for various selection schemes. Section 3 describes the most widely used selection schemes, presents known theoretical results on their selection pressure equivalence, introduces the noise takeover time and pure reproduction rate measures to be used for selection noise analysis, and provides theoretical estimates of the pure reproduction rate values for described selection schemes. Section 4 describes the experimental conditions and parameters whose effect on selection noise is studied. Section 5 provides and discusses results obtained for various selection schemes, with the selection pressure equivalent schemes considered separately. Finally, Section 6 concludes the paper.

2 Related work

Formal criteria to evaluate disproportion between the expected and actual number of copies of an individual in the mating pool were proposed in (Baker 1987; Brindle 1980). To introduce them, let us note that any selection scheme can be considered as a two-step procedure (Brindle 1980; Schell and Wegenkittl 2001). In the first step, the selection probabilities $P = (p_1, \ldots, p_N)$, $p_i \in [0, 1]$, $\sum_{i=1}^{N} p_i = 1$, are calculated for all individuals i, $1 \leq i \leq N$, of the current population. In the second step, a sampling algorithm, often called a sampling strategy, draws N individuals from the current population to the mating pool. Let e_i be the expected number of copies of the i-th individual in the mating pool, and let o_i be the actual number of its copies. Then $e_i = N * p_i$, expectation $E[o_i] = e_i$, and $\sum_{i=1}^{N} e_i = \sum_{i=1}^{N} o_i = N$. Thus, o_i is an estimator for e_i, and selection noise can be measured in terms of the vector $(o_1, \ldots, o_N)$ distribution characteristics, namely:

- *bias* is the average absolute difference between the actual and expected number of copies of an individual in the mating pool (Brindle 1980);
- *variance* of number of copies of an individual in the mating pool (Brindle 1980);
- *spread* is the range of possible number of copies of an individual in the mating pool (Baker 1987).

Following Brindle (1980), bias and variance are measures to estimate accuracy; following Baker (1987), bias indicates accuracy and spread indicates precision.

Criteria above were used in (Baker 1987; Brindle 1980) to evaluate the genetic drift of proportional selection schemes. In particular, it was shown that the most popular proportional selection schemes, roulette-wheel selection (RWS) and stochastic universal selection (SUS), both have zero bias, SUS has the minimum spread (from $\lfloor e_i \rfloor$ to $\lceil e_i \rceil$), and RWS has the maximum spread (from 0 to N). In addition, Baker (1987) used an empirical approach to evaluate bias of proportional selection schemes. He conducted runs of a genetic algorithm using the selection scheme under analyzes and no other genetic operators; the bias factor (the ratio of the actual number of copies of an individual in the mating pool to its expected value) and fertility factor (the percentage of individuals in a population that were selected into the mating pool) were calculated at each iteration of the algorithm.

Schell and Wegenkittl (2001) proposed the following noise estimation method based on calculation of the chi-square Goodness-of-Fit statistic. Let all individuals are grouped into d mutually disjoint classes $C_1, C_2, \ldots, C_d$. Let $E_j = \sum_{i \in C_j} e_i$ denotes the overall expectation in each class, and $O_j = \sum_{i \in C_j} o_i$ denotes the overall observation in each class. Then the chi-square statistic $\chi = \sum_{j=1}^{d} \frac{(O_j - E_j)^2}{E_j}$ can be used as a measure for accuracy, with the higher value of χ meaning lower accuracy. It is shown theoretically and empirically that $\chi \approx 0$ for SUS and $\chi \approx E[\chi^2_{d-1}] = d - 1$ for RWS (Schell and Wegenkittl 2001). To evaluate other selection schemes, a sequence of independent replications of vector $(o_1, \ldots, o_N)$ is generated. After that, sample expectation and variance are calculated as approximations for the bias and variance criteria. The empirical distribution function is also calculated and compared with theoretical chi-square distribution of a benchmark selection scheme that uses the same probability distribution. SUS and RWS are commonly used as benchmark selection schemes since they are often referred to as extreme sampling schemes with high and low sampling accuracy, respectively.

Based on the technique above, it is shown in (Schell and Wegenkittl 2001) that binary tournament selection with replacement is almost perfectly resembled by RWS using the same probability distribution and has lower accuracy in comparison with SUS using the same probability distribution; binary tournament selection with partial replacement provides significantly smaller bias and spread in comparison with RWS using the same probability distribution but has lower accuracy in comparison with SUS using the same probability distribution; the accuracy of binary probability tournament selection with partial replacement, $p > 0.5$, does not depend on p and cannot be distinguished from RWS using the same probability distribution. The chi-square statistic was also used to evaluate genetic drift of mixed size tournament selection scheme (msTS) (Huber and Schell 2002) and of several recently proposed selection schemes such as stairwise selection (SWS) (Haq et al 2019), split rank selection (SRS) (Hussain and Muhammad 2020), split-based selection (SBS) (Hussain and Cheema 2020), fitness-based selection (FBS) (Naqvi and Shad 2022), and round-robin based tournament selection (RRTS) (Hussain et al 2022).

Rogers and Prügel-Bennett (1999) proposed an analytical estimation of genetic drift in terms of reduce in population fitness variance after one generation. They introduced the *rate of genetic drift* measure as the ratio of the expectation of fitness variance in a given generation to the fitness variance in the previous generation. The rate is computed assuming the action of selection operator only: to avoid the effect of selection pressure, fitness value is not taken into account in the selection process. It is shown that the rate of genetic drift depends on the population size and selection scheme only, and equations for tournament selection schemes used in generational and generation gap genetic algorithms are derived.

Hill and O'Riordan (2016) analyzed genetic drift empirically by comparing the average and the best fitness of a population (*on-line* and *off-line*, respectively). Analysis in (Hill and O'Riordan 2016) was performed for SUS, stochastic remainder selection (SRS), and their scaled versions (linear scaling, window scaling, sigma truncation, and Boltzmann scaling) using a four-bit changing deceptive function; no statistically significant differences (Wilcoxon signed rank test) among these methods were revealed.

Another approach for empirical analysis of selection noise of selection schemes is to use a flat surface for evolutionary process being under the action of selection operator only. Mahfoud (1996) used this approach to demonstrate differences in selection noise between RWS and SUS in terms of time needed for the population to become homogeneous. In (Chisari and Amadio 2018), to compare genetic drift of RWS, SUS, and binary tournament selection schemes, genetic algorithm runs on a flat landscape are depicted on graphs with "number of generation" and "standard deviation" axes; these graphs visualize the population loss of diversity on the phenotype level. In (Hancock 1994), a function with greater range of fitness values is used, and it is shown graphically that for SUS compared to RWS the number of best individuals in the population grows faster while the fitness of the worst individual remains lower during algorithm execution. Unfortunately, no statistical analysis is provided in the mentioned studies.

Genetic drift analysis can also be done by calculating Markov chain transition matrices and finding the time required for the system to reach a homogeneous population (Holland 1992). However, results obtained in this manner are approximations and are difficult to generalize, as Petrowski and Ben Hamida (2016) state.

In (Wiles et al 2001) it is argued that tracking the progress of genetic algorithm towards population homogeneity could provide valuable information about genetic drift; the average fitness of the population when an allele becomes homogeneous is proposed as a quantitative measure of diversity. With this approach, several proportional and tournament selection schemes were analyzed empirically using a needle in a haystack problem. Based on the calculations averaged over 10 trials, authors conclude that deterministic and probabilistic binary tournament selection schemes have the highest genetic drift, RWS with sigma truncation and elitist deterministic tournament selection schemes have the lowest genetic drift, and RWS and roulette with ranking (stochastic binary tournament) are characterized by slightly higher genetic drift compared to the lowest one. The conclusion seems surprising since RWS is usually referred to as a scheme with the highest genetic drift. A possible explanation is that the results in (Wiles et al 2001) are obtained on a single parameter set, calculations are averaged over 10 trials only, and no statistical analysis is performed. Also, a crossover operator is used whose effect on selection noise remains an open question.

From the above we see that there exist different approaches to study genetic drift of selection schemes, both analytical and empirical ones. However, genetic drift of only some selection schemes has been evaluated analytically. Empirical analysis of genetic drift has been conducted for several popular selection schemes, but statistical analysis of obtained results has not always been performed. We still know little about the effects of various parameters on genetic drift and how selection schemes differ in selection noise while this information could shed light on parameter tuning in genetic algorithms (Gibbs et al 2008; Huber and Schell 2002).

3 Materials and methods

3.1 Selection schemes to be analyzed

In this paper, we analyze selection noise of the most popular selection schemes widely used in generational genetic algorithms. These are SUS, RWS, and these schemes using fitness scaling techniques such as linear scaling, power low scaling, and sigma truncation; linear and exponential ranking selection schemes based on SUS and RWS; deterministic and probabilistic tournament selection schemes with and without replacement. Recall that in a generational genetic algorithm the entire population of size N is replaced during each generation: N individuals are selected to the mating pool by use of a selection scheme, and these individuals produce N children that replace entirely the population of parents.

To avoid ambiguity, we briefly describe the above mentioned selection schemes below. We also present theoretical results on selection pressure equivalence for these schemes.

3.1.1 Description of selection schemes

In the RWS scheme, the probability p_i of an individual X_i to be selected into the mating pool is proportional to its fitness $f(X_i)$: $p_i = \frac{f(X_i)}{\sum_{j=1}^{N} f(X_j)}$. To implement, a segment of length p_i on a segment of unit length is assigned to each individual X_i. Then, a random number $0 < \xi \leq 1$ is generated, and the individual corresponding to the segment containing generated value is selected into the mating pool. This procedure is repeated N times.

The SUS scheme uses the same probability distribution as the RWS, but a segment of length N is divided into N segments of length $N * p_i$, a random number $0 < \xi \leq 1$ is generated only once, and N individuals that correspond to the segments containing points of the form $\xi + i, 0 \leq i \leq N-1$, are selected into the mating pool.

When a fitness scaling technique is used, RWS and SUS deal with fitness values scaled according to the applied scaling method. In particular, in linear scaling, $f_{scaled}(X) = a * f(X) + b$, where coefficients a and b may be chosen in a number of ways. In sigma truncation, $f_{scaled}(X) = f(X) - (\overline{f} - c * \sigma)$, where $\overline{f}$ is the mean fitness of the population, σ is the standard fitness deviation in the population, and c is a small integer constant. The power low scaling changes fitness values as $f_{scaled}(X) = f^k(X)$, where k is a small constant. When $f_{scaled}(X) \leq 0$, we put $f_{scaled}(X) = d$, where d is a small constant.

Suppose that individuals are sorted by their fitness values, and ranks are assigned to the individuals such that rank zero is assigned to the least fit individual and rank $N-1$ is assigned to the fittest one. In a ranking selection scheme, RWS or SUS is used as a sampling strategy. However, the probability p_i of an individual X_i to be drawn into the mating pool is no longer proportional to its fitness, but depends on its rank, i.e., the relative position of the individual in the sorted list. In linear ranking scheme, $p_i = \frac{(2-\beta)}{N} + \frac{2rank(X_i)*(\beta-1)}{N*(N-1)}$, where β is a constant, $1 < \beta \leq 2$. In exponential ranking scheme, $p_i = \frac{c-1}{c^{N-1}} * c^{N-rank(X_i)-1}$, where c is a constant, $0 < c < 1$. Hereinafter, the linear ranking scheme with RWS sampling strategy is denoted by *LinRank_RWS*, and the linear ranking scheme with SUS sampling strategy is denoted by *LinRank_SUS*. The exponential ranking schemes with RWS and SUS sampling strategies are denoted by *ExpRank_RWS* and *ExpRank_SUS*, respectively. Note that in ranking selection schemes, different probabilities are assigned to individuals with the same fitness. Therefore, the population should be shuffled at each iteration of the algorithm to avoid randomly favoring any genotypes.

In tournament selection schemes, the probabilities of individuals to be selected into the mating pool are not explicitly assigned. Instead, a group of t individuals is randomly selected from the population, and the fittest individual is inserted into the mating pool. The procedure is repeated N times. The fittest individual can be selected deterministically or with a probability p, $0.5 < p < 1.0$; in the latter case binary tournaments are usually used. There are also tournaments with replacement, when all tournament participants are returned back to the population, and without replacement, when tournament participants are withdrawn from the population. In the latter case, the population is copied t times and each copy is used to select individuals until it is exhausted. Another way of implementation is to join t copies of the population into the one of size $t * N$. This approach is used in (Schell and Wegenkittl 2001) and called the tournament with partial replacement selection scheme. In the sequel, we refer to the deterministic tournament selection scheme as *TournDet_With* if selection is made with replacement, *TournDet_Without* if selection is made without replacement, and *TournDet_Part* if selection is made with partial replacement. Similarly, the notations *TournProb_With*, *TournProb_Without*, and *TournProb_Part* are used for the probabilistic tournament selection schemes.

3.1.2 Selection pressure equivalence

In this subsection, we describe selection schemes known as being selection pressure equivalent, i.e., schemes that provide the same probability distributions. First of all, these are ranking selection schemes with different sampling strategies (SUS or RWS) and tournament selection schemes with, without, and with partial replacement. Note that tournament selection schemes with different replacement strategies actually provide different probability distributions, but this difference is usually considered to be negligible (Schell and Wegenkittl 2001).

It is shown in (Bäck 1996; Julstrom 1999) that deterministic tournament selection with replacement and exponential ranking selection schemes associate approximately the same probabilities with chromosomes in a population when

$$c \approx \left(1 - \frac{1}{N}\right)^t \tag{1}$$

It is stated in (Blickle and Thiele 1996; Goldberg and Deb 1991; Julstrom 1999) that individual's selection probabilities are the same for deterministic tournament selection without replacement, $t = 2$, and linear ranking selection, $\beta = 2$. Also, it is known from (Julstrom 1999) that the distribution of probabilistic tournament selection without replacement is equivalent to that of linear ranking selection when

$$\beta = 2p \tag{2}$$

Combining the statements above we also get the selection pressure equivalence for linear ranking and exponential ranking schemes when $\beta = 2$ and

$$c \approx \left(1 - \frac{1}{N}\right)^2 \tag{3}$$

3.2 Proposed selection noise estimation measures

Now let us introduce measures to be used for selection noise estimation. These measures are inspired by the biological genetic drift counterpart.

Similar to the takeover time criterion τ used to evaluate selection pressure (Grefenstette 1997), we introduce the noise takeover time criterion to evaluate selection noise. The *noise takeover time*, denoted by η, is the number of generations needed for a population to become homogeneous on a flat fitness function when the evolutionary process is driven by selection operator only. Obviously, η differs for different starting points. Therefore, the same initial populations should be used to analyze the effect of scheme parameters on selection noise or compare selection schemes with respect to genetic drift. In the following, we also use the notation η_i to denote the noise takeover time for populations of size N_i.

The lower the selection noise, the larger η, and vice versa. For selection schemes with very low selection noise, η can be very large. Moreover, for some selection schemes the population will never become homogeneous on a flat fitness landscape, so $\eta \to \infty$. This means that an appropriate budget for maximum number of iterations should be predefined to estimate η experimentally. This budget dependence is a limitation for the use of the criterion.

In addition to the noise takeover time, we introduce the *pure reproduction rate PRR* measure defined as the fraction of individuals selected into the mating pool when evolutionary process is driven by selection operator only and effect of selection pressure is excluded. Note that similar measures were used earlier for selection pressure analysis. For example, Baker (1989) introduced the *reproduction rate RR* measure as the percentage of individuals in a population that were selected into the mating pool; in his earlier papers, the percent involvement (Baker 1985) or fertility factor (Baker 1987) terms were used instead. In (Blickle and Thiele 1996), the *loss of diversity LD* measure was defined as the proportion of individuals that were not selected to reproduce. From the definitions it follows that $RR = 100 * (1 - LD)$. Baker (1989) used *RR* as a dynamic convergence measure. Blickle and Thiele (1996) viewed *LD* as a function of selection intensity. We propose to compute *PRR* on a flat fitness landscape to quantify the selection noise of selection schemes. Note that $PRR \leq 1$, and the less noisy the selection scheme, the larger *PRR*.

3.3 Methods of analysis

In this subsection, we describe the empirical approach used in the paper to estimate the noise takeover time and pure reproduction rate values. Then we give theoretical estimates of the proposed measures for the schemes described in Subsection 3.1.1.

3.3.1 Empirical approach

We use the following empirical approach to study genetic drift. We conduct multiple runs of genetic algorithm with a selection scheme under analysis; no other genetic operator is used, thus, the population of children matches the mating pool. Runs are conducted on a flat fitness function, hence, the effect of selection pressure is eliminated. Algorithm stops when population becomes homogeneous or when allocated time resources are spent. For each run, the number of iterations performed by the algorithm and *PRR* average over generations are saved. Afterwards, the minimum, maximum, average, and standard deviation of number of iterations and of *PRR* average are calculated for the runs when the algorithm stops under the population homogeneity condition. In the sequel, we denote by *AvgNI* the average number of iterations and by *AvgPRR* the average *PRR* average over runs. Also, the notation $AvgNI(N)$ is used for *AvgNI* computed for populations of size N.

3.3.2 Theoretical estimations

By construction, fitness scaling techniques accentuate small differences in objective values by mapping fitness values to an interval. On a flat objective function, these techniques do not change the probability distribution of individuals to be selected into the mating pool. Thus, RWS and SUS with fitness scaling have the same selection noise, i.e., the same noise takeover time and pure reproduction rate, as RWS and SUS without scaling, respectively. Also, on a flat fitness landscape, determining the fittest individual among a group of individuals is stochastic for any tournament selection scheme. Thus, probabilistic tournament selection schemes with different values of p should have the same noise takeover time and pure reproduction rate as the appropriate binary deterministic tournament selection schemes.

With a flat fitness function, the expected number of copies of the i-th individual in the mating pool for proportional selection schemes is given by $e_i = N * p_i = N * \frac{1}{N} = 1$. Recall that SUS is known as a scheme with the highest sampling accuracy: it has zero bias and minimum spread that ranges from $\lfloor e_i \rfloor$ to $\lceil e_i \rceil$, i.e., from 1 to 1. Therefore, we estimate $PRR = 1$ for SUS and its scaled versions. Also, on a flat fitness landscape no convergence of genetic algorithms using SUS or its scaled versions is expected, i.e., $\eta \to \infty$.

We have no theoretical estimates of η for other selection schemes. Theoretical estimates of *PRR* are given below.

For RWS, selection of an individual on a flat fitness landscape can be considered as a Bernoulli trial with probability of success $\frac{1}{N}$. Therefore, the probability of an individual not to be selected into the mating pool is determined by binomial distribution and is equal to $\binom{N}{0} * \left(\frac{1}{N}\right)^0 * \left(1 - \frac{1}{N}\right)^N = \left(\frac{N-1}{N}\right)^N$, where $\binom{N}{0}$ is the binomial coefficient. Thus, for RWS and its scaled versions we get

$$PRR = 1 - \left(\frac{N-1}{N}\right)^N \tag{4}$$

Eq. 4 can also be obtained by calculating the probability of an individual to be selected into the mating pool at least once. In this case, we get sum of terms of the geometric sequence:

$$\begin{aligned} PRR &= \frac{1}{N} + \frac{1}{N} * \left(1 - \frac{1}{N}\right)^1 + \frac{1}{N} * \left(1 - \frac{1}{N}\right)^2 + \\ &+ \cdots + \frac{1}{N} * \left(1 - \frac{1}{N}\right)^N = \\ &= \frac{\frac{1}{N} * \left(\left(1 - \frac{1}{N}\right)^N - 1\right)}{\left(1 - \frac{1}{N}\right) - 1} = 1 - \left(\frac{N-1}{N}\right)^N \end{aligned}$$

Eq. 4 is valid for tournament selection schemes with replacement since on a flat fitness function the probability of success equals $\frac{t}{N} * \frac{1}{t} = \frac{1}{N}$ for all individuals. This result differs from that of Blickle and Thiele (1996), where non-flat fitness landscapes are considered.

Recall that number of successful draws without replacement from a finite population is modeled by Hypergeometric(M, K, m) distribution, where M is the population size, K is the number of objects with a specified feature, and m is the number of draws from the population.

Let $N = k * t + i$, where $k = \lfloor \frac{N}{t} \rfloor$. For tournament selection schemes without replacement, k different individuals are drawn without replacement from each of t copies of the population. Hence, on a flat fitness landscape the probability of an individual to be selected into the mating pool is described by the Hypergeometric(N, 1, k) distribution. In our implementation, the remaining i individuals needed to fill the mating pool are drawn without replacement from another copy of the population, and the probability of an individual to be selected into the mating pool is described by the Hypergeometric(N, 1, i) distribution. From the above it follows that the probability of an individual not to be selected into the mating pool is equal to the probability of being selected zero times from all population copies and is equal to $\left(\frac{\binom{1}{0} * \binom{N-1}{k}}{\binom{N}{k}}\right)^t * \left(\frac{\binom{1}{0} * \binom{N-1}{i}}{\binom{N}{i}}\right) = \left(\frac{N-k}{N}\right)^t * \left(\frac{N-i}{N}\right)$. Thus, for TournDet_Without we get

$$PRR = 1 - \left(\frac{N-k}{N}\right)^t * \left(\frac{N-i}{N}\right) \tag{5}$$

where $k = \lfloor \frac{N}{t} \rfloor$, $i = N - k * t$. Note that when $i = 0$, we get $N = k * t$, and $\left(\frac{N-k}{N}\right)^t * \left(\frac{N-i}{N}\right) = \left(\frac{k*t-k}{k*t}\right)^t * \left(\frac{N-0}{N}\right) = \left(\frac{t-1}{t}\right)^t$. Therefore, we get

$$PRR = 1 - \left(\frac{t-1}{t}\right)^t \tag{6}$$

for TournDet_Without when $\frac{N}{t}$ is an integer.

For tournament selection schemes with partial replacement, N individuals are drawn without replacement from t joined copies of the population of size N. Therefore, the probability of an individual to be selected into the mating pool is described by the Hypergeometric($N * t$, t, N) distribution, and for TournDet_Part we obtain

$$\begin{aligned} PRR &= 1 - \left(\frac{\binom{t}{0} * \binom{t*N-t}{N}}{\binom{t*N}{N}}\right) = \\ &= 1 - \prod_{i=0}^{t-1} \frac{(t-1) * N - i}{t * N - i} \end{aligned} \tag{7}$$

Binary tournaments are the most widely used in theoretical and practical research, see (Chisari and Amadio 2018; Schell and Wegenkittl 2001; Wiles et al 2001). Therefore, values of *PRR* at $t = 2$ are of particular interest to us.

From Eq. 4 it follows that *PRR* tends to 0.63 as N increases. The value does not depend on t and is valid for binary TournDet_With and TournProb_With.

For binary tournaments without and with partial replacement, the *PRR* tends to 0.75 as N increases. Indeed, from Eq. 5 we get for binary TournDet_Without and TournProb_Without

$$PRR = 0.75 \tag{8}$$

when N is even, and

$$PRR = 0.75 - \frac{N^2 - N - 1}{4N^3} \tag{9}$$

when N is odd. From Eq. 7 we get for binary TournDet_Part and TournProb_Part

$$PRR = 1 - \frac{N-1}{4N-2} \tag{10}$$

For ranking selection schemes, the probabilities of individuals to be selected into the mating pool depend on the rank only, thus, estimates obtained earlier for *RR* or *LD* should also be valid under a flat fitness function. Using results from (Blickle and Thiele 1996), we get for LinRank_SUS

$$PRR = 1 - \frac{\beta - 1}{4} \tag{11}$$

and for ExpRank_SUS

$$PRR = 1 - \left(\frac{1 - \ln \frac{\alpha - 1}{\alpha \ln \alpha}}{\ln \alpha} - \frac{\alpha}{\alpha - 1} \right) \tag{12}$$

where $\alpha = c^N$. Values of *PRR* for LinRank_RWS and ExpRank_RWS should be smaller and bounded from above by Eq. 4.

4 Experiments

In this paper, experiments are conducted under the following conditions.

As a flat fitness landscape, the function $FCONST(X) = 100$ defined on binary strings of fixed length is used. Initial population of size N is generated randomly; each gene is set to 1 with the probability 0.5 or to 0 with the same probability. The algorithm stops when the population becomes homogeneous (that is, all strings representing individuals in the population become identical) or after 100000 iterations.

For each parameter set, a series of 100 runs of genetic algorithm using the selection scheme under analysis is performed. To avoid randomly assigning a better starting point to any algorithm, i-th runs of the compared algorithms are performed using the same initial population, $i = 1, \ldots, 100$. Criteria mentioned in Subsection 3.3.1 are computed over runs to estimate the noise takeover time and pure reproduction rate values.

We have conducted two stages of experiments. First, we study the effect of various parameters on genetic drift of the selection schemes under analysis. Secondly, we test selection schemes that provide the same probability distributions (the same selection pressure, see Subsection 3.1.2) to have the same selection noise.

At the first stage, the following parameters are considered:

- population size N;
- chromosome length l;
- parameters specific for selection schemes.

The reasons to choose the above parameters are as follows. It is shown analytically in (Rogers and Prügel-Bennett 1999) that the rate of genetic drift depends on the population size: the larger N, the smaller the genetic drift, hence, the larger the value of η. It is also well known that in nature genetic drift happens faster in small populations than in larger ones (Masel 2011). The diversity of chromosomes of the initial population affects the noise takeover time η: the greater the chromosomal diversity of the initial population, the larger the value of η. Chromosomal diversity depends on the initialization method and length of chromosomes. Indeed, when $2^l < N$, the population will contain identical chromosomes regardless of the initialization method used. To study the effect of initial population diversity we use different values of l. We use chromosome length values such that the possible number of different chromosomes is equal to 2, is about 30% of the population size, and is greater than the population size. Finally, parameters specific for selection schemes such as t and p in tournament selection schemes or β and c in ranking selection schemes are usually referred to as parameters affecting selection pressure (Bäck 1996; Baker 1985; Blickle and Thiele 1995, 1996; Goldberg and Deb 1991; Huber and Schell 2002; Julstrom 1999; Julstrom and Robinson 2000), whereas their effect on genetic drift is an open question.

All experiments at the first stage are conducted using the following parameter values:

- $N = 100$, $N = 200$, $N = 400$;
- $l = 1, l = \log(0.32N), l = 10, l = 100$;
- values of parameters specific for selection schemes are given in Table 1.

In the sequel, we refer to this set of parameter values as the *ParamSuite 1*. If necessary, *ParamSuite 1* is extended to refine the analyzed dependencies; the extensions are listed additionally for each particular case.

It is noted in Subsection 3.3.2 that RWS and SUS with fitness scaling have the same selection noise as RWS and SUS without scaling, respectively. Nevertheless, we conduct additional runs for RWS and SUS with fitness scaling to get more data for analysis and ensure the consistency of results obtained for proportional selection schemes. Similarly, we conduct additional runs for probabilistic tournament selection schemes with different values of p, although on a flat fitness landscape the probabilistic tournament selection schemes have the same selection noise as the appropriate deterministic tournament selection schemes, see Subsection 3.3.2.

For the second stage of experiments, the *ParamSuite 2* set of parameter values is used:

- $N = 100$, $N = 200$, $N = 300$, $N = 400$;
- $l = 1, l = 10, l = 100$;
- values of parameters specific for selection schemes are given in Table 2.

Table 1 Values of parameters specific for selection schemes (*ParamSuite 1*)

Scheme	Parameter values
SUS/RWS with linear scaling	$a = 1, b = 1; a = 1, b = -1; a = 2, b = 1; a = 2, b = -1$
SUS/RWS with sigma truncation	$c = 1; c = 2; c = 3; c = 4; c = 5$
SUS/RWS with power low scaling	$k = 1.005; k = 1.01; k = 1.05; k = 1.1; k = 1.5; k = 2$
LinRank_SUS/_RWS	$\beta = 2; \beta = 1.8; \beta = 1.6; \beta = 1.2$
ExpRank_SUS/_RWS	$c = 0.809; c = 0.9; c = 0.945; c = 0.955; c = 0.979; c = 0.996$
TournDet_With/_Without/_Part	$t = 2; t = 3; t = 4; t = 5; t = 6; t = 7; t = 8; t = 9; t = 10; t = 20$
TournProb_With/_Without/_Part	$p = 0.9; p = 0.8; p = 0.75; p = 0.7; p = 0.6$

Table 2 Values of parameters specific for selection schemes with the same selection pressure (*ParamSuite 2*)

Scheme 1	Scheme 2	Parameter values
LinRank_SUS	LinRank_RWS	$\beta = 2; \beta = 1.8; \beta = 1.6; \beta = 1.5; \beta = 1.4; \beta = 1.2$
ExpRank_SUS	ExpRank_RWS	$c = 0.9950; c = 0.9933; c = 0.9925; c = 0.9900; c = 0.9876; c = 0.9867; c = 0.9851; c = 0.9834; c = 0.9826; c = 0.9802; c = 0.9801; c = 0.9777; c = 0.9769; c = 0.9753; c = 0.9752; c = 0.9736; c = 0.9704; c = 0.9703; c = 0.9672; c = 0.9655; c = 0.9607; c = 0.9606; c = 0.9559; c = 0.9512; c = 0.9511; c = 0.9510; c = 0.9415; c = 0.9354; c = 0.9321; c = 0.9227; c = 0.9135; c = 0.9046; c = 0.9044; c = 0.8179$
TournDet_Without	TournDet_Part	t from Table 1
TournProb_Without	TournProb_Part	p from Table 1
TournDet_Without/_Part	TournDet_With	t from Table 1
TournProb_Without/_Part	TournProb_With	p from Table 1
TournDet_With/_Without/_Part	ExpRank_SUS/_RWS	t from Table 1 c is calculated for each t, N by Eq. 1
TournDet_With/_Without/_Part	LinRank_SUS/_RWS	$t = 2$ $\beta = 2$
LinRank_SUS	ExpRank_SUS	$\beta = 2$ c is calculated for each N by Eq. 3
LinRank_RWS	ExpRank_RWS	$\beta = 2$ c is calculated for each N by Eq. 3
TournProb_With/_Without/_Part	LinRank_SUS/_RWS	p from Table 2 β is calculated by Eq. 2

Parameter values for the second stage of experiments are chosen with regard to the results obtained in the first stage, and the data obtained in the first stage are reused where possible.

To evaluate the effect of selection scheme specific parameters, ANOVA is used. Cochran's C test is used to verify the equality of variances condition required for ANOVA to be applied. In case of strong differences in variances, the Kruskal-Wallis H test (by ranks) is additionally used. However, both the H test and ANOVA gave similar results in all the cases. To evaluate the effect of other parameters, two-sample t-tests are performed for two series of runs, each using the same parameter values except for the one under analysis.

Table 3 Effect of population size N: test results for RWS, $l = 100$

N_1	*AvgNI*, $N = N_1$	N_2	*AvgNI*, $N = N_2$	p-value
100	205.35	200	393.72	$8.38 * 10^{-13}$
100	205.35	400	817.53	$1.26 * 10^{-23}$
200	393.72	400	817.53	$2.55 * 10^{-12}$

Also, t-tests are used to compare selection noise of selection schemes with the same selection pressure and in some cases to clarify the effect of scheme specific parameters. We use $\alpha = 0.01$ significance level in our analysis.

5 Results and discussion

5.1 Effect of parameters on selection noise

In this subsection, we analyse the effect of population size, chromosome length, and selection scheme parameters on genetic drift of various proportional (Subsection 5.1.1), ranking (Subsection 5.1.2), and tournament (Subsection 5.1.3) selection schemes. Subsection 5.1.4 summarizes obtained results and compares them with those obtained in earlier studies.

5.1.1 Proportional selection schemes

All runs of SUS and its scaled versions were stopped after execution of 100000 iterations; $AvgPRR = 1$. This experimental result is consistent with our theoretical estimates, see Subsection 3.3.2.

To analyze the effect of population size N on η for RWS and RWS with scaling, two-sample t-tests are performed for series of runs, each series in a pair differs only in N; pairs with $N = 100$ and $N = 200$, $N = 200$ and $N = 400$, and $N = 100$ and $N = 400$ are compared. In all our experiments, values of *AvgNI* differ for different N. Table 3 illustrates the results obtained for RWS, $l = 100$.

The experiments have also shown that *AvgNI* grows proportionally to the population size growth. To confirm, we used two-sample t-tests for $AvgNI(N_j) * \frac{N_i}{N_j}$ and $AvgNI(N_i)$ values. Thus, we conclude that

$$\frac{N_i}{N_j} \approx \frac{\eta_i}{\eta_j} \tag{13}$$

To analyze the effect of chromosome length l on η for RWS and RWS with scaling, two-sample t-tests are performed for series of runs using the same parameter sets but different chromosome lengths. Our experiments revealed statistically significant difference in *AvgNI* only for pairs with $l = 1$ and $l \geq \log(0.32N)$, i.e., when numbers of different chromosomes in initial populations differ significantly. Table 4 shows the results obtained for RWS, $N = 100$.

Table 4 Effect of chromosome length l: test results for RWS, $N = 100$

l_1	*AvgNI*, $l = l_1$	l_2	*AvgNI*, $l = l_2$	p-value
1	129.56	5	188.95	$2.81 * 10^{-05}$
1	129.56	10	174.67	$1.14 * 10^{-03}$
1	129.56	100	205.35	$1.01 * 10^{-06}$
5	188.95	10	174.67	0.275
5	188.95	100	205.35	0.258
10	174.67	100	205.35	0.033

Experimentally obtained values of *AvgPRR* correspond to the values from Eq. 4 and are about 0.63 for all N.

5.1.2 Ranking selection schemes

To examine the effect of chromosome length l on η, two-sample t-tests are performed for series of runs using the same parameter sets but different chromosome lengths. For all ranking selection schemes, the effect of l on η is confirmed for pairs with $l = 1$ and $l \geq \log(0.32N)$ and is not confirmed for pairs where both $l \geq \log(0.32N)$.

For linear ranking selection schemes LinRank_SUS and LinRank_RWS, the effect of population size N on η with the growth of η in proportion to the growth of N, see Eq. 13, is confirmed by two-sample t-tests performed for series of runs using the same parameter sets but different values of N.

For exponential ranking selection schemes ExpRank_SUS and ExpRank_RWS, the effect of population size N on η was not clear from the analysis of runs based on *ParamSuite 1*. Therefore, additional series of runs were conducted using the extended set of parameter values, namely:

- $N = 40$, $N = 80$, $N = 100$, $N = 200$, $N = 300$, $N = 400$;
- $l = 10$, $l = 100$;
- 34 different values of parameter c given in Table 2.

As before, two-sample t-tests were performed for different series of runs. Parameter values for these series differed in the population size only, and pairs N_i and N_j such that $\frac{N_i}{N_j} \geq 2$ were used for the analysis.

Experiments have shown that for ExpRank_RWS the value of *AvgNI* increases with increasing N, and this increase is more significant for larger values of c. Namely, for $N_i > N_j$ it is true that $AvgNI(N_i) > AvgNI(N_j)$ for larger values of c, and either $AvgNI(N_i) > AvgNI(N_j)$ or the

Table 5 Effect of population size N: test results for ExpRank_RWS, $l = 100$

c	*AvgNI*, $N = 40$	*AvgNI*, $N = 300$	p-value
0.9950	74.15	507.22	$1.70 * 10^{-33}$
0.9933	76.42	472.30	$6.02 * 10^{-39}$
0.9925	76.14	444.33	$7.85 * 10^{-33}$
0.9801	72.56	180.38	$6.66 * 10^{-20}$
0.9777	80.73	173.08	$1.03 * 10^{-17}$
0.9769	74.65	170.17	$2.50 * 10^{-14}$
0.9046	35.73	41.13	$6.04 * 10^{-2}$
0.9044	40.23	44.70	$1.98 * 10^{-1}$
0.8179	18.13	19.80	$2.47 * 10^{-1}$

Table 6 Effect of population size N: test results for ExpRank_SUS, $l = 100$

c	*AvgNI*, $N = 40$	*AvgNI*, $N = 300$	p-value
0.9950	1595.42	1539.02	$6.61 * 10^{-1}$
0.9933	1155.50	1143.56	$8.95 * 10^{-1}$
0.9925	993.63	1048.91	$4.64 * 10^{-1}$
0.9801	378.21	291.90	$1.17 * 10^{-3}$
0.9777	353.11	250.63	$6.96 * 10^{-6}$
0.9769	297.70	233.62	$1.44 * 10^{-3}$
0.9046	60.55	36.09	$2.59 * 10^{-11}$
0.9044	55.68	42.32	$3.62 * 10^{-4}$
0.8179	23.67	20.50	$6.90 * 10^{-2}$

results are not statistically significant for smaller values of c. It depends on N what values of c are considered as larger or smaller ones. Table 5 illustrates the above.

Experiments conducted for ExpRank_SUS have revealed another form of dependency. For $N_i > N_j$ we get either $AvgNI(N_i) < AvgNI(N_j)$ or the results are not statistically significant. Thus, for ExpRank_SUS the estimated noise takeover time η decreases with increasing N. Table 6 shows an illustrative example of the effect of N.

The result is unexpected, since it is well known that drift occurs faster in small populations than in larger ones (Masel 2011). In addition, it was shown by Rogers and Prügel-Bennett (1999) that increasing the population size reduces the genetic drift for any selection scheme. At the same time, our empirical results can be explained using pure reproduction rate measure. It follows from Eq. 12 that *PRR* decreases significantly as N increases, see Table 7 as an illustration. This means that the larger N is, the fewer is the proportion of individuals providing their copies to the mating pool. From the above and the fact that assignment of probabilities is random on a flat fitness landscape, it follows that η decreases and genetic drift increases with increasing N.

Table 7 PRR values for ExpRank_SUS calculated by Eq. 12

N	$c = 0.809$	$c = 0.945$	$c = 0.996$
50	0.317	0.680	0.975
100	0.191	0.480	0.950
200	0.112	0.303	0.901
300	0.081	0.226	0.853
400	0.064	0.182	0.806

For ExpRank_SUS and ExpRank_RWS, the effect of scheme parameter c on genetic drift is approved experimentally, with an increase in c leading to an increase in η and *PRR*. Tables 6 and 7 illustrate the dependency.

For LinRank_SUS, the effect of scheme parameter β on genetic drift is confirmed, with a decrease in β leading to an increase in η. A similar effect of β on η is expected for LinRank_RWS, but due to the high stochasticity of the RWS sampling strategy this dependence is weakly expressed, that is, the results of pairwise comparisons of *AvgNI* obtained for different values of β are often not statistically significant. The same dependence on β is observed for *PRR*.

5.1.3 Tournament selection schemes

For all tournament selection schemes, the effect of population size N on η and proportional relationship given by Eq. 13 are confirmed by our experiments. Similar to other selection schemes, the effect of chromosome length l on η is approved empirically only for pairs with $l = 1$ and $l \geq \log(0.32N)$.

As presumed in Subsection 3.3.2, our experiments revealed no effect of parameter p on η for TournProb_With, TournProb_Without, and TournProb_Part selection schemes.

We found no effect of parameter t on η for TournDet_With selection scheme; we got p-value > 0.05 in all our experiments. *PRR* is also independent of t, see Eq. 4. Thus, we conclude that there is no effect of t on selection noise of TournDet_With selection scheme, and this conclusion is confirmed by both the η and *PRR* measures.

For TournDet_Without and TournDet_Part, we expect the selection noise to increase as t increases, since the larger t is, the larger fraction of the population is eliminated during each tournament. Our experiments showed a general trend for *AvgNI* to decrease as t increases, but to analyze the dependence of *AvgNI* on t we had to conduct more runs of the algorithm with various t, N values. Let $AvgNI_{t_i}$ denotes *AvgNI* obtained in experiments with tournaments of size t_i. For all $t_i > t_j$ we either get $AvgNI_{t_i} < AvgNI_{t_j}$ or the difference in *AvgNI* is not statistically significant. We obtained statistically significant relationships $AvgNI_2 > AvgNI_{t_i}$

Table 8 Values of selection noise measures for TournDet_Part, $N = 100, l = 10$

t	*AvgNI*	*AvgPRR*	*PRR*[a]
2	396.14 ± 237.53	0.751 ± 0.001	0.751
3	297.58 ± 173.97	0.705 ± 0.002	0.705
4	253.53 ± 140.52	0.685 ± 0.002	0.685
5	252.09 ± 136.66	0.674 ± 0.002	0.674
6	260.72 ± 131.16	0.667 ± 0.002	0.667
7	228.07 ± 139.19	0.662 ± 0.002	0.662
8	234.40 ± 146.81	0.658 ± 0.002	0.658
9	225.22 ± 111.79	0.655 ± 0.002	0.655
10	227.57 ± 120.54	0.653 ± 0.002	0.653
20	208.93 ± 113.24	0.643 ± 0.003	0.643
50	196.01 ± 105.80	0.638 ± 0.003	0.638
100	207.88 ± 102.96	0.636 ± 0.002	0.636

[a] Calculated by Eq. 7

for all $t_i > 2$ at 0.05 significance level and all $t_i > 3$ at 0.01 significance level. We noticed that the difference between $AvgNI_{t_i}$ and $AvgNI_{t_j}$ was never statistically significant when $t_i, t_j > 0.2N$. Thus, for *AvgNI* values we observe a sharp decrease when t changes from 2 to 3, followed by a smooth decrease when t increases from 4 to N. The same form of dependence on t is confirmed for *PRR* theoretically (see Eq. 5 and 7) and empirically. Table 8 illustrates the above for TournDet_Part selection scheme.

5.1.4 Summary remarks

Table 9 summarizes the results described above for the noise takeover time measure η. For all selection schemes, the effect of chromosome length l on η is revealed only for pairs with $l = 1$ and $l \geq \log(0.32N)$ and is not observed if the initial populations contain sufficiently large numbers of different chromosome types (more than 30% of the population). For this reason, the parameter l is not included in the table. Also, the effect of scheme parameters on pure reproduction rate *PRR* coincides with their effect on η and is not mentioned in the table.

For most selection schemes, we observe growing η with growing population size N. This result agrees with the well-known biological phenomenon that drift is faster in small populations than in larger ones (Masel 2011). Another form of dependence on N for ExpRank_SUS is explained by specificity of the method, since the probabilities assigned by this method lead to a significant decrease in *PRR* values as N increases. Note that *PRR* decreases with increasing N for many other methods as well. However, this decrease is inconsiderable and does not lead to an inverse dependence of η on N.

Our finding of no effect of fitness scaling techniques on selection noise is consistent with the results of Hill and O'Riordan (2016) whose analysis of genetic drift revealed no differences between SUS and its scaled versions. At the same time, it is claimed by Wiles et al (2001) that RWS with sigma truncation has lower genetic drift than RWS. We attribute the discrepancy in results to the small experimental base and effect of crossover. In any case, the methodology proposed by Wiles et al (2001) requires further investigation.

For linear ranking selection schemes, selection noise decreases as β decreases. For exponential ranking selection schemes, selection noise decreases as c increases. Selection noise increases with increasing t for TournDet_Without and TournDet_Part selection schemes and does not depend on t for TournDet_With selection scheme. The parameter p has no effect on selection noise for TournProb_With, TournProb_Without, and TournProb_Part selection schemes. Note that the independence of selection noise from p for TournProb_Part selection scheme was previously stated by Schell and Wegenkittl (2001).

5.2 Selection noise of schemes with equivalent selection pressure

In this subsection, we test selection pressure equivalent schemes for selection noise equivalence. Subsection 5.2.1 tests ranking selection schemes with SUS and RWS sampling strategies. Subsection 5.2.2 tests tournament selection schemes with, without, and with partial replacement. Subsection 5.2.3 tests ranking and tournament selection schemes whose selection pressure equivalence is established by Eq. 1, 2, and 3. Finally, Subsection 5.2.4 summarizes our findings and compares them with those obtained in earlier studies, see Tables 12 and 13.

5.2.1 Ranking selection schemes

LinRank_SUS and LinRank_RWS with the same values of β as well as ExpRank_SUS and ExpRank_RWS with the same values of c provide the same probability distributions, see Subsection 3.1.1. That is, these schemes have the same selection pressure. However, the schemes have different selection noise because they use different sampling strategies.

It is assumed that schemes with SUS sampling strategy have lower genetic drift than schemes with RWS sampling strategy since SUS and RWS are known as high and low sampling accuracy schemes, respectively. This is true for linear ranking selection schemes: LinRank_SUS has larger values of η and *PRR* in comparison with LinRank_RWS. The result is confirmed by two-sample t-tests performed for series of runs differing by the selection strategy only. Note that the difference in η increases as β decreases, see Table 10 for an illustration. The same is true for the difference in *PRR*. The

Table 9 Effect of various parameters on the noise takeover time η

Selection scheme	Effect of scheme parameters	Effect of population size N	$\frac{N_i}{N_j} \approx \frac{\eta_i}{\eta_j}$
RWS	-	Yes, η increases as N increases	Yes
RWS with scaling	No	Yes, η increases as N increases	Yes
LinRank_SUS	Yes η increases as β decreases	Yes, η increases as N increases	Yes
LinRank_RWS	Yes η increases as β decreases (weakly expressed)	Yes, η increases as N increases	Yes
ExpRank_SUS	Yes η increases as c increases	Yes, η decreases as N increases (weakly expressed)	No
ExpRank_RWS	Yes η increases as c increases	Yes, η increases as N increases (expressed for large c)	No
TournDet_With	No	Yes, η increases as N increases	Yes
TournDet_Without	Yes η increases as t decreases	Yes, η increases as N increases	Yes
TournDet_Part	Yes η increases as t decreases	Yes, η increases as N increases	Yes
TournProb_With	No	Yes, η increases as N increases	Yes
TournProb_Without	No	Yes, η increases as N increases	Yes
TournProb_Part	No	Yes, η increases as N increases	Yes

Table 10 Effect of sampling strategy: test results for LinRank_SUS and LinRank_RWS, $l = 100$, $N = 300$

β	$AvgNI_S$ [a]	$AvgNI_R$ [b]	p-value	$\frac{AvgNI_S}{AvgNI_R}$
2	1195.29	448.22	$2.08 * 10^{-22}$	2.67
1.8	1396.93	462.43	$3.61 * 10^{-23}$	3.02
1.6	2017.95	541.24	$7.13 * 10^{-27}$	3.73
1.5	2347.19	498.63	$1.11 * 10^{-35}$	4.71
1.4	2894.71	579.98	$6.74 * 10^{-36}$	4.99
1.2	5063.77	586.49	$5.42 * 10^{-40}$	8.63

[a] $AvgNI_S$ denotes *AvgNI* obtained for LinRank_SUS
[b] $AvgNI_R$ denotes *AvgNI* obtained for LinRank_RWS

explanation is that when $\beta \approx 1$, LinRank_SUS approaches SUS and LinRank_RWS approaches RWS, the high and low sampling accuracy schemes, respectively.

Let $AvgNI_S$ denotes *AvgNI* obtained in experiments for ExpRank_SUS, and $AvgNI_R$ denotes *AvgNI* obtained in experiments for ExpRank_RWS with the same parameter values. For exponential ranking selection schemes, we get $AvgNI_S > AvgNI_R$ for larger values of c, and either $AvgNI_S > AvgNI_R$ or the results are not statistically significant for smaller values of c; values of c that can be considered as delimiting ones increase with N. Note that as c increases, the difference in η and *PRR* increases, but this change is much slower compared to the dependence of η and *PRR* distance on β for linear ranking selection schemes. It is shown in Subsection 5.1.2 that η increases for ExpRank_RWS and decreases for ExpRank_SUS when N increases. Hence, the difference in η also decreases as N increases, see Table 11 for an illustration.

Table 11 Effect of sampling strategy: test results for ExpRank_SUS and ExpRank_RWS, $c = 0.9777$, $l = 100$

N	$AvgNI_S$ [a]	$AvgNI_R$ [b]	p-value	$\frac{AvgNI_S}{AvgNI_R}$
40	353.11	80.73	$5.79 * 10^{-34}$	4.37
80	361.97	128.55	$6.87 * 10^{-26}$	2.82
100	352.52	136.29	$3.43 * 10^{-20}$	2.59
200	273.26	167.98	$9.52 * 10^{-11}$	1.63
300	250.63	173.08	$2.02 * 10^{-6}$	1.45
400	215.84	184.39	$5.21 * 10^{-2}$	1.17
1000	192.78	168.82	$1.00 * 10^{-1}$	1.14

[a] $AvgNI_S$ denotes *AvgNI* obtained for ExpRank_SUS
[b] $AvgNI_R$ denotes *AvgNI* obtained for ExpRank_RWS

It is shown in Subsection 3.1.2 that linear ranking and exponential ranking selection schemes have the same selection pressure when $\beta = 2$ and c is calculated by Eq. 3 for each N. We obtained selection noise equivalence for LinRank_SUS and ExpRank_SUS as well as for LinRank_RWS and ExpRank_RWS for the specified parameter values.

5.2.2 Tournament selection schemes

Although tournament selection schemes with and without replacement produce slightly different probability distributions, the difference is often considered negligible (Schell and Wegenkittl 2001). Therefore, we tested these selection schemes for selection noise equivalence.

First, we compared different implementations of tournament without replacement selection schemes. We found no statistically significant difference in *AvgNI* values between

TournDet_Part and TournDet_Without or TournProb_Part and TournProb_Without. *PRR* values for these methods are also very close. Thus, we consider tournament selection schemes without replacement and with partial replacement as selection noise equivalent.

Tournament without replacement selection schemes are often referred to as the schemes with lower genetic drift in comparison with tournament with replacement selection schemes (Hancock 1994; Petrowski and Ben Hamida 2016). Note that TournDet_Without is indistinguishable from TournDet_With at $t = N$ by scheme definition. We assume that TournDet_Without has lower selection noise than TournDet_With at small t, and the difference in selection noise between these methods decreases as t increases. This should also be true for TournDet_Part and TournDet_With since TournDet_Without and TournDet_Part have approximately the same selection noise as shown above. Our two-sample t-tests confirmed larger *AvgNI* values for TournDet_Without and TournDet_Part in comparison with TournDet_With when $t = 2$ or $t = 3$. We found no statistically significant differences in *AvgNI* when $t > 0.2N$. The same trend is observed for *PRR* values, see Eq. 4, 5, and 7. Therefore, our assumption is correct and supported by both η and *PRR* measures.

It is claimed in Subsection 3.3.2 that TournProb_With, TournProb_Without, and TournProb_Part have the same selection noise as TournDet_With, TournDet_Without, and TournDet_Part, respectively, $t = 2$, while these schemes have different selection pressures. Selection noise equivalence of deterministic and probabilistic tournament selection schemes is also claimed in (Wiles et al 2001) although it is not clear which form of replacement the authors use. From the above it follows and is confirmed statistically that TournProb_Without and TournProb_Part have lower selection noise than TournProb_With.

5.2.3 Ranking and tournament selection schemes

Recall that approximately the same individual's selection probabilities are assigned by the following selection schemes: TournDet_With and exponential ranking when c is calculated by Eq. 1 (Bäck 1996; Julstrom 1999), TournDet_Without and linear ranking when $t = 2$ and $\beta = 2$ (Blickle and Thiele 1996; Goldberg and Deb 1991; Julstrom 1999), TournProb_Without and linear ranking when β is calculated by Eq. 2 (Julstrom 1999), and linear ranking and exponential ranking when $\beta = 2$ and c is calculated by Eq. 3. We tested the above-mentioned selection pressure equivalent schemes to be selection noise equivalent. Since difference in selection pressure of tournament selection schemes with and without replacement is often considered to be negligible (Schell and Wegenkittl 2001), we used both replacement methods in our experiments.

Our two-sample t-tests confirmed larger *AvgNI* values for TournDet_With in comparison with ExpRank_RWS. It is shown in Subsections 5.1.2 and 5.1.3 that for ExpRank_RWS η increases as c increases, while for TournDet_With η does not depend on t. Thus, the difference in η changes, that is, decreases with increasing c (decreasing t), and attains its minimum at $t = 2$. A similar trend is observed for *PRR* values. It is argued in (Schell and Wegenkittl 2001) that binary TournDet_With has approximately the same genetic drift as RWS using the same probability distribution, i.e., ExpRank_RWS with c calculated by Eq. 1. We cannot confirm the result of Schell and Wegenkittl (2001) because we obtained less genetic drift for TournDet_With using both measures.

Values of *AvgNI* are larger for TournDet_With in comparison with ExpRank_SUS when $t > 4$ and smaller when $t = 2$. The result for $t = 2$ is consistent with the result from (Schell and Wegenkittl 2001), where it is stated that binary TournDet_With has lower accuracy in comparison with SUS using the same probability distribution, i.e., ExpRank_SUS with c calculated by Eq. 1. Differences in *AvgNI* at $t = 3$ and $t = 4$ observed in our experiments are not always statistically significant.

For TournDet_Without and TournDet_Part, we experimentally confirmed larger values of *AvgNI* in comparison with ExpRank_RWS for all t, and larger values of *AvgNI* in comparison with ExpRank_SUS for $t > 2$. Similar to TournDet_With, difference in *AvgNI* decreases as c increases (t decreases), and the same trend is observed for *PRR* values. At $t = 2$, *PRR* values are slightly larger for TournDet_Without and TournDet_Part in comparison with ExpRank_SUS, while *AvgNI* values are statistically indistinguishable.

We experimentally confirmed lower genetic drift for TournDet_With, TournDet_Without, and TournDet_Part at $t = 2$ in comparison with LinRank_RWS at $\beta = 2$. Note that the result for TournDet_Part was obtained earlier by Schell and Wegenkittl (2001). We also confirmed higher genetic drift for TournDet_With at $t = 2$ in comparison with LinRank_SUS at $\beta = 2$.

We found no statistically significant differences in *AvgNI* values for TournDet_Without, TournDet_Part, and LinRank_SUS at $t = 2$, $\beta = 2$. *PRR* values for these schemes are also very close and around 0.75, see Eq. 8, 9, 10, and 11. Thus, we consider TournDet_Without, TournDet_Part, and LinRank_SUS at $t = 2$, $\beta = 2$ as selection noise equivalent schemes. It is stated in (Schell and Wegenkittl 2001) that binary TournDet_Part has lower accuracy than SUS using the same probability distribution, i.e., LinRank_SUS at $\beta = 2$. This result is not supported by either our experimental or theoretical data.

We detected larger *AvgNI* and *PRR* values for TournProb_Without and TournProb_Part in comparison with LinRank_RWS, and the differences were statistically significant in all our experiments. This result differs from the result of Schell and Wegenkittl (2001) paper, where the selection noise equivalence for TournProb_Part and LinRank_RWS is stated. At the same time, we consider TournProb_With and LinRank_RWS as the schemes with approximately the same selection noise.

Values of *AvgNI* are smaller for TournProb_With, TournProb_Without and TournProb_Part in comparison with LinRank_SUS, and the difference increases as β decreases since for LinRank_SUS η increases as β decreases and for probabilistic tournament selection schemes η does not depend on p, see Subsections 5.1.2 and 5.1.3. Note that the difference is not always statistically significant for TournProb_Without and TournProb_Part when $p \geq 0.9$. The same trend is observed for *PRR* values. Thus, we state that TournProb_With has higher selection noise than LinRank_SUS, and TournProb_Without and TournProb_Part have higher selection noise than LinRank_SUS when $p < 0.9$.

5.2.4 Summary remarks

To summarize the results of comparing selection schemes based on their selection noise, we provide two tables. Table 12 presents pairs of selection schemes that are equivalent in selection pressure but different in selection noise. Table 13 presents pairs of selection schemes that are equivalent in both selection pressure and selection noise.

We added the well-known result of SUS and RWS comparison to Table 12 for completeness. This result is also supported by our experimental and theoretical study. We confirm that ranking selection schemes with SUS sampling strategy have lower genetic drift than ranking selection schemes with RWS sampling strategy. However, for exponential ranking selection schemes the difference is confirmed statistically for large c and small N only. We have found that tournaments without replacement have higher accuracy than tournaments with replacement only for small t, and different implementations of tournament without replacement selection scheme lead to approximately the same genetic drift.

We see that LinRank_SUS, ExpRank_SUS, TournDet_Without, and TournDet_Part are equivalent not only in selection pressure, but also in selection noise when $t = 2$, $\beta = 2$, and c is calculated by Eq. 3. At the same time, selection noise equivalence is not confirmed for most of the selection pressure equivalent schemes. Notice that in some cases, scheme parameters determine which of the two schemes has higher selection noise (e.g., TournDet_With and ExpRank_SUS).

6 Conclusion

In this paper, genetic drift as a property of selection schemes used in generational genetic algorithms is studied. First, a review of known results in this field is given. Secondly, two new measures, namely the noise takeover time η and pure reproduction rate *PRR*, are proposed to quantify selection noise of selection schemes. Thirdly, the effect of population size, chromosome length, and selection scheme parameters on genetic drift of various proportional, ranking, and tournament selection schemes is analyzed. Finally, selection pressure equivalent schemes are compared to be selection noise equivalent.

There exist a number of studies where schemes assigning the same probabilities to individuals in a population are considered equivalent (Bäck 1996; Goldberg and Deb 1991; Julstrom 1999). When the equivalence of selection pressure of tournament and ranking selection schemes is established, some researchers prefer tournaments. The reason is that tournaments are characterized by lower time complexity since they do not require sorting the population followed by computing and applying explicit probabilities (Bäck 1996; Goldberg and Deb 1991; Julstrom and Robinson 2000; Schell and Wegenkittl 2001). At the same time, other researches argue that time complexity of selection schemes should hardly be prioritized because fitness evaluation time often dominates the execution time of other parts of algorithms. In addition, there may be differences in selection noise of schemes with the same selection pressure (Hancock 1997; Huber and Schell 2002). Indeed, our comparison of selection pressure equivalent schemes revealed differences in selection noise for most of them, see Table 12. Thus, replacing one scheme with another should be done not only by comparing their selection pressure, but also by comparing their selection noise. Note that we used both η and *PRR* measures in our study, and both measures showed the same trends. Therefore, we conclude that these measures provide relevant and reliable estimates of selection noise although these estimates differ in some cases from those obtained in earlier studies, see Tables 12 and 13. Finding relationships between various measures estimating selection noise and other scheme or population properties is a task for future.

Genetic drift is a factor that affects the convergence of genetic algorithms (Rogers and Prügel-Bennett 1999). Gibbs et al (2008) proposed to control it by the choice of the population size N. We confirmed empirically that there is a strong relationship between η and N. Moreover, for most selection schemes we found a dependence given by Eq. 13, see Table 9. This dependence can be used to predict η for various population sizes and control genetic drift. Finding an equation expressing the relationship between η, N, and scheme parameters is a subject for future research.

We also argue that genetic drift can be controlled by choosing the selection scheme (Rogers and Prügel-Bennett 1999; Chisari and Amadio 2018) or by tuning the parameters of a selection scheme. Tables 9, 12, and 13 can help in providing this control. For example, in parallel island models, each

Table 12 Selection pressure equivalent schemes having different selection noise

Method 1 (noise smaller)	Method 2 (noise larger)	Condition	Note
SUS	RWS		Agrees with Baker (1987), Brindle (1980) Schell and Wegenkittl (2001),
SUS with scaling	RWS with Scaling		
LinRank_SUS	LinRank_RWS		The difference increases as β decreases
ExpRank_SUS	ExpRank_RWS		The difference increases as c increases and decreases as N increases
TournDet_Without	TournDet_With	$t < 4$	
TournDet_Part	TournDet_With	$t < 4$	
TournProb_Without	TournProb_With		
TournProb_Part	TournProb_With		
TournDet_Without	ExpRank_SUS	$t > 2$	The difference increases as t increases
TournDet_Part	ExpRank_SUS	$t > 2$	The difference increases as t increases
TournDet_Without	ExpRank_RWS		The difference increases as t increases
TournDet_Part	ExpRank_RWS		The difference increases as t increases
TournDet_With	ExpRank_SUS	$t > 4$	The difference increases as t increases
ExpRank_SUS	TournDet_With	$t = 2$	Agrees with Schell and Wegenkittl (2001)
TournDet_With	ExpRank_RWS	$t > 2$	The difference increases as t increases
TournDet_With	ExpRank_RWS	$t = 2$	In (Schell and Wegenkittl 2001) equality is asserted
LinRank_SUS	TournDet_With	$t = 2$	
TournDet_With	LinRank_RWS	$t = 2$	
TournDet_Without	LinRank_RWS	$t = 2$	
TournDet_Part	LinRank_RWS	$t = 2$	Agrees with Schell and Wegenkittl (2001)
LinRank_SUS	TournProb_With		The difference increases as β decreases
LinRank_SUS	TournProb_Without	$p < 0.9$	The difference increases as β decreases
LinRank_SUS	TournProb_Part	$p < 0.9$	The difference increases as β decreases
TournProb_Without	LinRank_RWS		
TournProb_Part	LinRank_RWS		In (Schell and Wegenkittl 2001) equality is asserted

Table 13 Selection pressure equivalent schemes having the same selection noise

Method 1	Method 2	Condition	Note
LinRank_SUS	ExpRank_SUS	$\beta = 2$	
LinRank_RWS	ExpRank_RWS	$\beta = 2$	
TournDet_Without	TournDet_Part		
TournProb_Without	TournProb_Part		
TournDet_Without	TournDet_With	$t > 0.2N$	
TournDet_Part	TournDet_With	$t > 0.2N$	
TournDet_Without	ExpRank_SUS	$t = 2$	
TournDet_Part	ExpRank_SUS	$t = 2$	
TournDet_Without	LinRank_SUS	$t = 2$	
TournDet_Part	LinRank_SUS	$t = 2$	Schell and Wegenkittl (2001) state a lower accuracy for TournDet_Part
TournProb_With	LinRank_RWS	$p < 0.9$	

island algorithm could have different selection noise and/or selection pressure. Providing procedures to control selection noise and find the pressure-noise balance for various computational models and problem types is a subject for future research.

Acknowledgements The authors thank the anonymous reviewers for their comments and suggestions that helped us improve the paper.

Funding Funding for open access publishing: Universidad de Sevilla/CBUA This work is funded by University of Seville plan to aid Ukraine and by Grant PID2023-147198NB-I00 funded by MICIU/AEI/10.13039/501100011033 (Agencia Estatal de Investigación), Spain, and by FEDER, UE.

Data Availability The datasets generated and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have no relevant financial or non-financial interests to disclose.

Ethical approval Ethical approval is not required for this type of study.

References

Bäck T (1996) Evolutionary Algorithms in Theory and Practice: Evolution Strategies, evolutionary programming, genetic algorithms. Oxford University Press Inc, New York, USA, Genetic Algorithms

Baker JE (1985) Adaptive selection methods for genetic algorithms. In: Greffenstette JJ (ed) Proceedings of the 1st International Conference on Genetic Algorithms and Their Applications. Lawrence Erlbaum Associates Inc., Hillsdale, NJ, pp 101–111

Baker JE (1987) Reducing bias and inefficiency in the selection algorithm. In: Greffenstette JJ (ed) Genetic algorithms and their applications: Proceedings of the 2nd International Conference on Genetic Algorithms. Lawrence Erlbaum Associates Inc., Hillsdale, NJ, pp 14–21

Baker JE (1989) An analysis of the effects of selection in genetic algorithms. PhD thesis, Graduate School of Vanderbilt University, Nashville, TN

Blickle T, Thiele L (1995) A mathematical analysis of tournament selection. In: Eshelman LJ (ed) Proceedings of the 6th International Conference on Genetic Algorithms. Morgan Kaufmann, San Mateo, CA, pp 9–16

Blickle T, Thiele L (1996) A comparison of selection schemes used in evolutionary algorithms. Evolutionary Computation 4(4):361–394. https://doi.org/10.1162/evco.1996.4.4.361

Branke J, Schmidt C (2003) Selection in the presence of noise. In: Cantú-Paz E, Foster JA, Deb K et al (eds) Genetic and Evolutionary Computation - GECCO 2003, vol 2723. Lecture Notes in Computer Science. Springer, Berlin, Heidelberg, pp 766–777

Brindle AF (1980) Genetic algorithms for function optimization. PhD thesis, University of Alberta, Edmonton, Canada

Carvalho JT, Nolfi S (2024) The role of morphological variation in evolutionary robotics: Maximizing performance and robustness. Evol Comput 32(2):125–142. https://doi.org/10.1162/evco_a_00336

Chisari C, Amadio C (2018) Tosca: a tool for optimisation in structural and civil engineering analyses. International Journal of Advanced Structural Engineering 10:401–419. https://doi.org/10.1007/s40091-018-0205-1

Gibbs MS, Dandy GC, Maier HR (2008) A genetic algorithm calibration method based on convergence due to genetic drift. Information Sciences 178(14):2857–2869. https://doi.org/10.1016/j.ins.2008.03.012

Goldberg DE, Deb K (1991) A comparative analysis of selection schemes used in genetic algorithms. In: Rawlins GJ (ed) Foundations of Genetic Algorithms. Morgan Kaufmann, San Mateo, CA, p 69–93, https://doi.org/10.1016/B978-0-08-050684-5.50008-2

Grefenstette J (1997) Proportional selection and sampling algorithms. In: Bäck T, Fogel DB, Michalewicz Z (eds) Handbook of Evolutionary Computation. Institute of Physics Publishing and Oxford University Press, Bristol and New York, p C2.2:1–C2.2:7

Hancock PJB (1994) An empirical comparison of selection methods in evolutionary algorithms. In: Fogarty TC (ed) AISB Workshop on Evolutionary Computing, vol 865. Lecture Notes in Computer Science. Springer, Berlin, pp 80–94

Hancock PJB (1997) A comparison of selection mechanisms. In: Bäck T, Fogel DB, Michalewicz Z (eds) Handbook of Evolutionary Computation. Institute of Physics Publishing and Oxford University Press, Bristol and New York, p C2.8:1–C2.8:11

Haq Eu, Ahmad I, Hussain A et al (2019) A novel selection approach for genetic algorithms for global optimization of multimodal continuous functions. Computational Intelligence and Neuroscience 2019. https://doi.org/10.1155/2019/8640218

Hill S, O'Riordan C (2016) Diversifying techniques & neutrality in genetic algorithms. In: Merelo J, Melício F, Cadenas JM, et al (eds) Proceedings of the 8th International Joint Conference on Computational Intelligence (IJCCI 2016) - ECTA, INSTICC, vol 3. SciTePress, pp 140–147, https://doi.org/10.5220/0006036201400147

Holland JH (1992) Adaptation in Natural and Artificial Systems: An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence. University of Michigan Press, Ann Arbor, MI, https://doi.org/10.7551/mitpress/1090.001.0001

Huber R, Schell T (2002) Mixed size tournament selection. Soft Computing 6:449–455. https://doi.org/10.1007/S00500-001-0160-8

Hussain A, Cheema SA (2020) A new selection operator for genetic algorithms that balances between premature convergence and population diversity. Croatian Operational Research Review 11:107–119. https://doi.org/10.17535/crorr.2020.0009

Hussain A, Muhammad YS (2020) Trade-off between exploration and exploitation with genetic algorithm using a novel selection operator. Complex & Intelligent Systems 6:1–14. https://doi.org/10.1007/s40747-019-0102-7

Hussain A, Riaz S, Amjad M et al (2022) Genetic algorithm with a new round-robin based tournament selection: Statistical properties analysis. PLOS ONE 17(9):e0274456. https://doi.org/10.1371/journal.pone.0274456

Julstrom BA (1999) It's all the same to me: revisiting rank-based probabilities and tournaments. In: Proceedings of the 1999 Congress on

Evolutionary Computation - CEC99 (Cat. No. 99TH8406), vol 2. IEEE Press, Piscataway, NJ, pp 1501–1505, https://doi.org/10.1109/CEC.1999.782661

Julstrom BA, Robinson DH (2000) Simulating exponential normalization with weighted k-tournaments. In: Proceedings of the 2000 Congress on Evolutionary Computation. CEC00 (Cat. No.00TH8512), vol 1. IEEE Press, Piscataway, NJ, pp 227–231, https://doi.org/10.1109/CEC.2000.870299

Mahfoud SW (1996) Niching methods for genetic algorithms. PhD thesis, University of Illinois at Urbana-Champaign, Urbana, USA, uMI Order No. GAX95-43663

Masel J (2011) Genetic drift. Current Biology 21(20):R837–R838. https://doi.org/10.1016/j.cub.2011.08.007

Naqvi FB, Shad MY (2022) A new fitness-based selection operator for genetic algorithms to maintain the equilibrium of selection pressure and population diversity. Croat Oper Res Rev (CRORR) 13(1):113–130. https://doi.org/10.17535/crorr.2022.0008

Petrowski A, Ben Hamida S (2016) Evolutionary algorithms. In: Siarry P (ed) Metaheuristics. Springer International Publishing, Cham, p 115–178, https://doi.org/10.1007/978-3-319-45403-0_6

Rogers A, Prügel-Bennett A (1999) Genetic drift in genetic algorithm selection schemes. IEEE Transactions on Evolutionary Computation 3(4):298–303. https://doi.org/10.1109/4235.797972

Schell T, Wegenkittl S (2001) Looking beyond selection probabilities: Adaptation of the χ^2 measure for the performance analysis of selection methods in GAs. Evolutionary Computation 9(2):243–256. https://doi.org/10.1162/106365601750190424

Wiles J, Schulz R, Bolland S et al (2001) Selection procedures for module discovery: exploring evolutionary algorithms for cognitive science. In: Moore JD, Stenning K (eds) Proceedings of the 23rd Annual Conference of the Cognitive Science Society (CogSci 2001). Lawrence Erlbaum Associates, Mahwah, New Jersey, pp 1124–1129