On the Mean Convergence Time of Multi-parent Genetic Algorithms Without Selection

Chuan-Kang Ting

International Graduate School of Dynamic Intelligent Systems,
University of Paderborn, 33100 Paderborn, Germany
`ckting@upb.de`

Abstract. This paper investigates genetic drift in multi-parent genetic algorithms (MPGAs). An exact model based on Markov chains is proposed to formulate the variation of gene frequency. This model identifies the correlation between the adopted number of parents and the mean convergence time. Moreover, it reveals the pairwise equivalence phenomenon in the number of parents and indicates the acceleration of genetic drift in MPGAs. The good fit between theoretical and experimental results further verifies the capability of this model.

1 Introduction

Multi-parent genetic algorithms (MPGAs) are genetic algorithms using multi-parent crossovers. Traditionally, genetic algorithms (GAs) adopt two parents in crossover to reproduce offspring. This idea is reasonable because, to the best of our knowledge, the form of sexual reproduction on the Earth is absolutely of two parents. Multi-parent crossovers break through this natural limitation by allowing more than two parents in the process of crossover. In a sense, multi-parent genetic algorithms can be viewed as multi-parent generalization of genetic algorithms. In light of MPGAs, several multi-parent crossovers have been proposed and shown their power in a variety of optimization problems [5,7,15]. However, most of these crossovers are validated empirically; only a few theoretical analyses of multi-parent crossovers are conducted.

Genetic drift is a phenomenon that a finite population, even if no selection pressure is applied, will ultimately converge to a uniform population. The rate of genetic drift serves as an important index of how fast population diversity is lost. Schippers [14] studied the genetic drift of two multi-parent crossovers: *uniform scanning crossover* (U-Scan) and *occurrence based scanning crossover* (OB-Scan). His work revealed that U-Scan has no influence on genetic drift whilst OB-Scan induces severe genetic drift, as the number of parents is increased. Nevertheless, the strength of genetic drift in his work is determined by comparing the probabilities of drift in and drift out. The rate of genetic drift in MPGAs is still an open question.

This paper investigates in theory the rate of genetic drift in MPGA using OB-Scan. Specifically, we propose an exact model for the mean convergence

time — the principal measure of genetic drift rate [2]. First, we analyze the gene frequency altered by multi-parent crossovers. Based on gene frequency, we model the behavior of MPGAs through Markov chains. This Markov model affords the correlation between parent numbers and mean convergence time. In addition, the theoretical results reveal the pairwise equivalence of parent numbers in OB-Scan. These theoretical results are further verified by a series of experiments.

The rest of this paper is organized as follows. Section 2 describes OB-Scan and Section 3 analyzes its effect on gene frequency. Next, we model MPGAs with Markov chains. Theoretical results and experimental validation are presented in Section 5. Finally, conclusions are drawn in Section 6.

2   Occurrence Based Scanning Crossover (OB-Scan)

Occurrence based scanning crossover is one of the *scanning crossovers* proposed by Eiben et al. [6]. Scanning crossovers are multi-parent generalization of *uniform crossover* — a widely used crossover in GAs. In uniform crossover, the donor for each locus is randomly picked from two selected parents. Extended to more than two parents, scanning crossovers choose the donor at random or using heuristics. According to different strategies, Eiben et al. proposed three variations of scanning crossovers: *uniform scanning crossover* (U-Scan), *occurrence based scanning crossover* (OB-Scan), and *fitness based scanning crossover* (FB-Scan). In this paper, we only discuss OB-Scan.

Rather than at random, OB-Scan determines offspring genes depending on the occurrence of parental genes at that locus. Specifically, it picks the *majority* of parental gene values as the offspring gene for each locus. Note that in this paper OB-Scan is defined to break ties by randomly[1] choosing a binary. Examples of 2-parent OB-Scan (corresponding to uniform crossover) and 4-parent OB-Scan are given in Fig. 1. The formal definitions of the components of GAs and OB-Scan are drawn below.
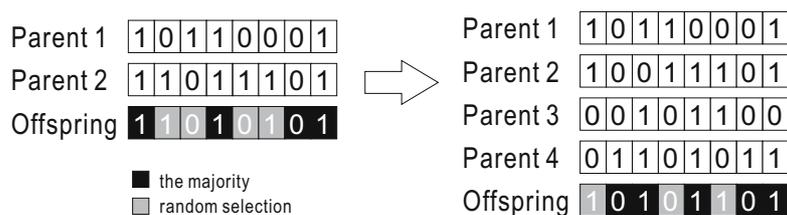

Fig. 1. Examples of 2-parent OB-Scan (left) and 4-parent OB-Scan (right)

---

[1] In the original version of OB-Scan [6], OB-Scan breaks ties by directly inheriting the genes of the first selected parent. However, random tie break conforms to generalization of uniform crossover.

Definition 1 (Chromosome and Population).

1. *A* chromosome $\boldsymbol{c}$ *is encoded as a bit string, i.e.* $\boldsymbol{c} \stackrel{\text{def}}{=} (c_1, \ldots, c_l) \in \{0,1\}^l$, *where* c_i *denotes a* gene *and l is the* chromosome length.
2. *A* population C *is a set of chromosomes:* $C \stackrel{\text{def}}{=} \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m\}$, *where* $\boldsymbol{c}_i \in \{0,1\}^l$ *and m is the* population size.

Definition 2 (OB-Scan). *Given n parents* $\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n \in C$ *selected from population* C, *OB-Scan reproduces the offspring* $\boldsymbol{c}' = \mathcal{X}_{\text{ob}}(\boldsymbol{c}_1, \ldots, \boldsymbol{c}_n) = (c'_1, \ldots, c'_l)$ *by*

$$c'_i = \begin{cases} 0 & \text{if } \sum_{j=1}^{n} (\boldsymbol{c}_j)_i < \frac{n}{2} \\ 1 & \text{if } \sum_{j=1}^{n} (\boldsymbol{c}_j)_i > \frac{n}{2} \quad \text{for } i = 1, \ldots, l, \\ \text{rand}(0,1) & \text{otherwise} \end{cases}$$

where $(\boldsymbol{c}_j)_i$ *denotes the* i^{th} *gene of the chromosome* $\boldsymbol{c}_j$, *and* $\text{rand}(0,1)$ *is a binary random function.*

3   Variation of Gene Frequency

Gene frequency is widely used as a quantitative measure of genetic variation in population genetics [9]. It also suffices to clue us in on the course of evolution in GAs. In this section we analyze the variation of gene frequency caused by OB-Scan.

Definition 3 (Gene Frequency). *The gene frequency* $p_k(\alpha, t)$ *is defined as the proportion of allele* α *at locus k in the population at time t. Let* $C = \{\boldsymbol{c}_1, \ldots, \boldsymbol{c}_m\}$ *be a population at time t and let* $C_k(\alpha) = \{\boldsymbol{c} \in C \mid c_k = \alpha\}$ *be the subset of population where chromosomes possess allele* α *at locus k. The gene frequency*

$$p_k(\alpha, t) \stackrel{\text{def}}{=} \frac{|C_k(\alpha)|}{|C|},$$

where $|C|$ *and* $|C_k(\alpha)|$ *represent the cardinality of set C and* $C_k(\alpha)$.

As above defined, chromosomes are represented as binary strings. Thus there exists exactly two gene frequencies $p_k(1, t)$ and $p_k(0, t)$ with $p_k(1, t) = 1 - p_k(0, t)$ for every locus k and time t. For simplicity, we refer to the gene frequency $p_k(1, t)$ as $p_k(t)$ and refer to $p_k(0, t)$ as $(1 - p_k(t))$.

Definition 4 (Variation of Gene Frequency in GAs). *Let* $p_k^s(t), p_k^x(t), p_k^m(t)$ *be the gene frequencies after performing selection, crossover, and mutation at generation t. The variation of gene frequency in GAs can be expressed as*

$$p_k(t) \xrightarrow{\text{selection}} p_k^s(t) \xrightarrow{\text{crossover}} p_k^x(t) \xrightarrow{\text{mutation}} p_k^m(t) \xrightarrow{\text{survivor}} p_k(t+1), \quad (1)$$

To investigate genetic drift, *random selection* and *no mutation* is assumed. In this paper we focus on generational GAs. As a result, the gene frequency in

the process of GA is only affected by the sampling of random selection and the process of crossover. These influences will be analyzed in Lemma 1. Incidentally, the symbol $p_k(t)$ is referred to as p_k while the indication of time t is not in effect.

Before conducting the analysis of OB-Scan, we introduce the incomplete beta function for simplifying the expression of equations .

Definition 5 (Incomplete Beta Function). *The* incomplete beta function *is defined as*

$$I_x(a,b) \stackrel{\text{def}}{=} \frac{1}{\text{Beta}(a,b)} \int_0^x t^{a-1}(1-t)^{b-1}dt,$$

where $a, b > 0$ and $\text{Beta}(a,b)$ *is the* beta function.

The incomplete beta function holds the following properties:

1. (26.5.24 [1]) For binomial distribution $B(n,p)$,

$$\sum_{i=a}^{n} B(i;n,p) = I_p(a, n-a+1). \tag{2}$$

2. (26.5.16 [1])

$$I_x(a,b) = \frac{1}{a \cdot \text{Beta}(a,b)} x^a (1-x)^b + I_x(a+1,b). \tag{3}$$

Using the above definition and properties, we embark on the analysis of OB-Scan's influence on gene frequency.

Lemma 1. *Suppose we have the gene frequency p_k of the population. The gene frequency, denoted by p_k^{ob}, of the offspring reproduced by n-parent OB-Scan $\mathcal{X}_{\text{ob}}$ with $n \in \mathbb{N}_{>1}$ is*

$$p_k^{\text{ob}} = I_{p_k}(a,a),$$

where I_p denotes the incomplete beta function and $a = \lceil \frac{n}{2} \rceil$.

Proof. Let X be the number of parents possessing the allele 1 at locus k among n selected parents. Since the process of random selection is independent, it is a Bernoulli process. Performing this selection n times, the number X holds a binomial distribution with probability mass function

$$\Pr(X = x) = B(x;n,p_k) = \binom{n}{x} (p_k)^x (1-p_k)^{n-x}.$$

Let $\mathfrak{D}_1$ denote the event that OB-Scan assigns the allele 1 to the offspring locus k. According to Definition 2, OB-Scan yields

$$\Pr(\mathfrak{D}_1 \mid X = x) = \begin{cases} 1 & \text{if } x > n/2, \\ 0 & \text{if } x < n/2, \\ \frac{1}{2} & \text{if } x = n/2. \end{cases}$$

For OB-Scan with an odd number of parents ($n = 2a - 1$ with $a \in \mathbb{N}_{>1}$), from (2) we have

$$p_k^{\mathrm{ob}} = \Pr(\mathfrak{D}_1) = \sum_{x=0}^{n} \Pr(\mathfrak{D}_1 \mid X = x) \cdot \Pr(X = x)$$
$$= \sum_{x=a}^{2a-1} B(x; 2a - 1, p_k) = I_{p_k}(a, a).$$

Similarly, for OB-Scan with an even number of parents ($n = 2a$ with $a \in \mathbb{N}$),

$$p_k^{\mathrm{ob}} = \sum_{x=a+1}^{2a} B(x; 2a, p_k) + \frac{1}{2} B(a; 2a, p_k)$$
$$= I_{p_k}(a, a) - \frac{1}{a \mathrm{Beta}(a, a)} (p_k)^a (1 - p_k)^a + \frac{1}{2} \binom{2a}{a} (p_k)^a (1 - p_k)^a$$
$$= I_{p_k}(a, a) + \left[-\frac{\Gamma(2a)}{a \Gamma(a) \Gamma(a)} + \frac{1}{2} \frac{(2a)!}{a! a!} \right] (p_k)^a (1 - p_k)^a$$
$$= I_{p_k}(a, a).$$

$\square$

Corollary 1 (Pairwise Equivalence). *Let $p_k^{\mathrm{ob}(n)}$ be the gene frequency p_k^{ob} corresponding to n-parent OB-Scan. For $n \in 2\mathbb{N}$ and $n \geq 4$,*

$$p_k^{\mathrm{ob}(n)} = p_k^{\mathrm{ob}(n-1)}.$$

Proof. Trivial (since $\left\lceil \frac{n}{2} \right\rceil = \left\lceil \frac{n-1}{2} \right\rceil$ in Lemma 1). $\square$

4   Modeling with Markov Chains

Markov chains have been used to model the exact behavior of GAs [2,8,10] and to analyze the convergence of GAs [4,11,13]. In this paper, we use Markov chains to model the evolution of gene frequency. From this Markov model, the mean convergence time can be derived.

In light of gene frequency, a GA can be viewed as a stochastic process manipulating the number of allele 1 (or 0) in the population: Let random variables $G_k(t) \in \{0, 1, \ldots, m\}$ be the number of allele 1 at locus k at generation t. The process of GAs on gene frequency can be represented as $\{G_k(t) : t \in \mathbb{Z}_*\}$. Since for every $i_0, i_1, \cdots, i_{t+1} \in \{0, 1, \ldots, m\}$ the process $\{G_k(t)\}$ satisfies

$$\Pr\{G_k(t+1) = i_{t+1} \mid G_k(t) = i_t, G_k(t-1) = i_{t-1}, \ldots, G_k(0) = i_0\}$$
$$= \Pr\{G_k(t+1) = i_{t+1} \mid G_k(t) = i_t\},$$

the process $\{G_k(t)\}$ is a Markov chain. The formal definition of the Markov chain for gene frequency is given as follows.

Definition 6 (Markov Chain for Gene Frequency). *In the Markov chain* $\{G_k(t)\}$ *for gene frequency at locus* $k \in \{1, \ldots, l\}$ *in GAs,*

1. *the* state *is defined as the number of allele 1 at locus* k *in the population and the state space is thus* $\{0, 1, \ldots, m\}$. *A state* i *in* $\{G_k(t)\}$ *implies the gene frequency at locus* k *is*

$$p_k = \frac{i}{m}.$$

2. *The* transition matrix *of* $\{G_k(t)\}$ *is defined as* $\mathbf{P} \overset{\text{def}}{=} (\rho_{ij})$, *where* ρ_{ij} *is the transition probability of state* i *to state* j:

$$\rho_{ij} \overset{\text{def}}{=} \Pr\{G_k(t+1) = j \mid G_k(t) = i\}.$$

The previous section has shown how OB-Scan changes the gene frequency. From those formulae, we derive the transition probabilities of the Markov chain $\{G_k(t)\}$ for the evolution of gene frequency in MPGAs.

Theorem 1. *For a GA using random selection, n-parent OB-Scan, and no mutation, the transition probability* ρ_{ij} *of the Markov chain* $\{G_k(t)\}$ *corresponding to this GA is*

$$\rho_{ij} = B(j; m, p'_k)$$

with

$$p'_k = I_{\frac{i}{m}}\left(\left\lceil\frac{n}{2}\right\rceil, \left\lceil\frac{n}{2}\right\rceil\right).$$

Proof. The state i of transition probability ρ_{ij} suggests the gene frequency $p_k = \frac{i}{m}$. Given this frequency p_k, from Lemma 1 we can obtain the gene frequency p'_k of the offspring reproduced by a GA using random selection, n-parent OB-Scan, and no mutation:

$$p'_k = p_k^{\text{ob}} = I_{\frac{i}{m}}\left(\left\lceil\frac{n}{2}\right\rceil, \left\lceil\frac{n}{2}\right\rceil\right).$$

In generational GAs, population is completely replaced with the subpopulation, consisting of m offspring reproduced by m times of *selection-crossover-mutation* process. Since this process is independent, the number of allele 1 holds a binomial distribution $B(m, p'_k)$. Therefore, the transition probability

$$\begin{aligned}
\rho_{ij} &= \Pr\{G_k(t+1) = j \mid G_k(t) = i\} \\
&= B(j; m, p'_k).
\end{aligned}$$

$\square$

For the Markov chain $\{G_k(t)\}$, of particular interest to us is, if at all, the *convergence* of $\{G_k(t)\}$ — at that time the population turns out to be either all-zeros or all-ones at each locus. For this, first we introduce a special kind of Markov chains, called *absorbing* Markov chains [3], which have this convergence property, i.e. absorption. Next, we will prove the Markov chain corresponding to the aforementioned GA belongs to such kind of Markov chains; that is to say, the corresponding GA will converge. From the properties of absorbing Markov chains we can further derive the mean convergence time.

Definition 7 (Absorbing States).

1. The closed set S^c is a set of states whose transition probabilities $\rho_{ij} = 0$ for all $i \in S^c, j \notin S^c$.
2. A state i is said to be absorbing if and only if $\exists S^c : S^c = \{i\} \iff \rho_{ii} = 1$.
3. A Markov chain with absorbing states is called an absorbing Markov chain.

Proposition 1. *For the GA given in Theorem 1, its corresponding Markov chain $\{G_k(t)\}$ is an absorbing Markov chain with exactly two absorbing states: 0 and m.*

Proof. According to the definition of absorption and Theorem 1, we know

$$\{G_k(t)\} \text{ is absorbing} \iff \exists i : \rho_{ii} = 1 \iff \exists i : B(i; m, p'_k) = 1. \qquad (4)$$

The solutions of $B(i; m, p'_k) = 1$ subject to $p'_k = I_{\frac{i}{m}}\left(\left\lceil \frac{n}{2} \right\rceil, \left\lceil \frac{n}{2} \right\rceil\right)$ and $n \in \mathbb{N}_{>1}$ are (i) $i = 0$ with $p'_k = 0$ and (ii) $i = m$ with $p'_k = 1$. This leads to, for the Markov chain $\{G_k(t)\}$

$$\rho_{00} = \rho_{mm} = 1 \implies \{G_k(t)\} \text{ is absorbing} .$$

Thus we complete the proof that the Markov chain $\{G_k(t)\}$ is absorbing with exactly two absorbing states 0 and m. □

The above proposition indicates the chain $\{G_k(t)\}$ will get absorbed into state 0 or m. It means that the process of the predefined GA will 'drift' into all-zeros or all-ones for each locus, that is, reaching convergence. In addition to the existence of convergence, we are interested in the mean time to reach it. To compute the mean time for a chain to get absorbed, we introduce the fundamental matrix [3] and its related property below.

Definition 8 (Fundamental Matrix). *For a Markov chain with b absorbing states, the transition matrix can be rewritten as*

$$\mathbf{P} = \begin{pmatrix} \mathbf{I}_b & \mathbf{0} \\ \mathbf{R} & \mathbf{Q} \end{pmatrix}, \qquad (5)$$

where $\mathbf{I}_b$ is a $b \times b$ identity matrix. The fundamental matrix *for this absorbing Markov chain is defined as*

$$\mathbf{F} \stackrel{\text{def}}{=} (\mathbf{I} - \mathbf{Q})^{-1}.$$

Theorem 2 ([12]). *Let $\mathbf{F}$ be the fundamental matrix of an absorbing Markov chain. The fundamental matrix $\mathbf{F}$ stands for the mean time τ_{ij} that the process spends at transient state j starting from transient state i.*

Theorem 3 (The Mean Convergence Time of MPGA). *Suppose we have the GA given in Theorem 1. Let $\mathbf{F} = (\tau_{ij})$ be the fundamental matrix of the Markov chain $\{G_k(t)\}$ corresponding to this GA. For some locus $k \in \{1, \ldots, l\}$, given the initial state distribution $\boldsymbol{\pi}(0) = (\pi_0(0), \ldots, \pi_m(0))$, the mean convergence time*

$$\tau = \sum_{i=1}^{m-1} \sum_{j=1}^{m-1} \pi_i(0)\tau_{ij}.$$

Proof. Let A be the set of absorbing states in $\{G_k(t)\}$. Proposition 1 gives $A \equiv \{0, m\}$. According to Theorem 2, the fundamental matrix of $\{G_k(t)\}$ represents the mean time τ_{ij} for $i, j \in \overline{A}$. Hence, the mean time τ_i that the process spends among transient states starting from transient state i can be derived by

$$\tau_i = \sum_{j \in \overline{A}} \tau_{ij}.$$

Given the initial state distribution $\boldsymbol{\pi}(0)$, we have the mean convergence time

$$\tau = \sum_{i=1}^{m-1} \tau_i \Pr(i \mid t = 0) = \sum_{i=1}^{m-1} \tau_i \pi_i(0) = \sum_{i=1}^{m-1} \sum_{j=1}^{m-1} \pi_i(0) \tau_{ij}. \qquad \square$$

5   Theoretical Results and Experimental Validation

This section demonstrates theoretical results obtained from the above theorems. Moreover, we conduct experiments on single locus ($l = 1$) to verify these theoretical results. The setting of MPGA used in our experiments is generational GA, bit-string representation, random selection, and no mutation. Each experiment setting includes 1000 independent runs.

Figure 2 compares the mean convergence time obtained from Theorem 3 and from experiments. First, this figure shows that the theoretical and the experimental results fit very well. In addition, it shows that for $n \in 2\mathbb{N}$ a MPGA using n-parent OB-Scan performs correspondingly to that using $(n-1)$-parent OB-Scan, which confirms the *pairwise equivalence* claimed in Corollary 1. Second, this figure indicates the fact that, compared with two parents, using more than two parents in OB-Scan causes a drastic decrease in mean convergence
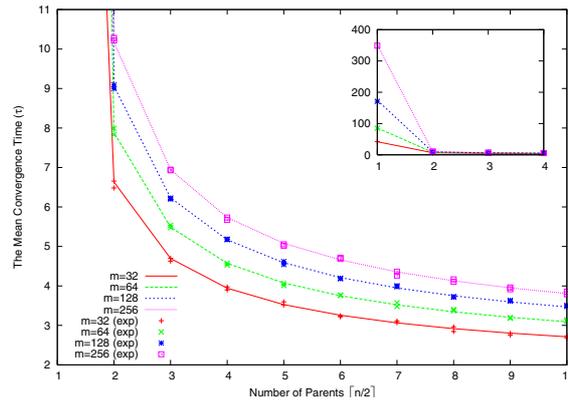
Fig. 2. Comparison of the theoretical (lines) and the experimental (symbols) mean convergence time τ for n-parent OB-Scan and population size m
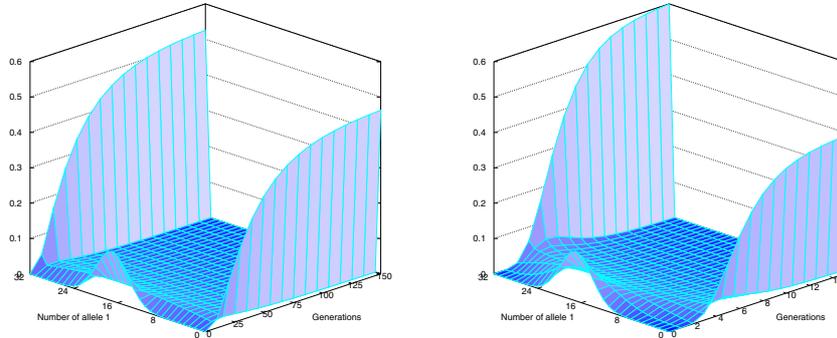
Fig. 3. The progress of genetic drift of 2-parent OB-Scan (left) and 3-parent OB-Scan (right) for population size $m = 32$ with initialization bias $\beta = \frac{1}{32}$

time. Nonetheless, further raising parents yields a relatively small difference. The mean convergence time ($m = 256$) for $n = 2$, for example, needs 351.55 generations while it takes only 10.17 generations for $n = 3$ (or 4) and 6.94 generations for $n = 5$ (or 6). This speedup in convergence reflects that OB-Scan with more than two parents accelerates genetic drift.

Next, we examine the progress of genetic drift in case of a initialization bias. We denote by β the bias of initial gene frequency to the allele 1. Figure 3 compares the progress of genetic drift of uniform crossover (i.e. 2-parent OB-Scan) and 3-parent OB-Scan for population size $m = 32$ under initialization bias $\beta = \frac{1}{32}$. As aforementioned, the genetic drift of 3-parent OB-Scan is much faster than that of uniform crossover. Interestingly, the distribution of convergence probability of uniform crossover differs from that of 3-parent OB-Scan either. Asoh and Mühlenbein [2] have shown the convergence probability of uniform crossover equals the initialization probability, which is reflected in Fig. 3. Yet, adopting more parents does not follow this rule. The 3-parent OB-Scan gives a probability (≈ 0.6) higher than the initialization probability ($\frac{17}{32} \approx 0.53$). This outcome suggests that using more parents in OB-Scan will intensify the preference of the initialization.

6   Conclusions

This paper presents an exact model for exploration of genetic drift in MPGAs. First we analyze the gene frequency altered by OB-Scan. Based on gene frequency, we model the behavior of MPGAs through Markov chains. The mean convergence time is further derived from this model.

The theoretical results demonstrate that raising parents in OB-Scan shortens the mean convergence time; that is, it accelerates genetic drift. This outcome not only reconfirms Schippers' claims about the genetic drift of scanning crossovers, but also gives the expected time of convergence. In addition, our analysis reveals the pairwise equivalence in OB-Scan: n-parent OB-Scan performs analogously

with $(n-1)$-parent OB-Scan for $n \in 2\mathbb{N}$. Moreover, the progress of genetic drift under initialization bias suggests raising parents in OB-Scan will intensify the preference of initialization for allele 0 or 1. The good fit between theoretical and experimental results validates our theoretical arguments and the capability of the proposed model.

References

1. M. Abramowitz and I.A. Stegun, editors. *Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables.* Dover Publications, 1972. ninth Dover printing.
2. H. Asoh and H. Mühlenbein. On the mean convergence time of evolutionary algorithms without selection and mutation. In *Parallel Problem Solving from Nature – PPSN III*, volume 866 of *LNCS*, pages 88–97, Berlin, 1994. Springer.
3. P. Brémaud. *Markov Chains: Gibbs Fields, Monte Carlo Simulation, and Queues.* Springer Verlag, 1999.
4. T.E. Davis and J.C. Principe. A markov chain framework for the simple genetic algorithm. *Evolutionary Computation*, 1(3):269–288, 1993.
5. A.E. Eiben. Multiparent recombination in evolutionary computing. In *Advances in Evolutionary Computing*, pages 175–192. Springer, 2002.
6. A.E. Eiben, P.-E. Rau'e, and Zs. Ruttkay. Genetic algorithms with multi-parent recombination. In *Parallel Problem Solving from Nature – PPSN III*, volume 866 of *LNCS*, pages 78–87. Springer, 1994.
7. A.E. Eiben and C.H.M. van Kemenade. Diagonal crossover in genetic algorithms for numerical optimization. *Journal of Control and Cybernetics*, 26(3):447–465, 1997.
8. D.E. Goldberg and P. Segrest. Finite markov chain analysis of genetic algorithms. In *Proceedings of the 2nd International Conference on Genetic Algorithms and their Applications*, pages 1–8. Lawrence Erlbaum Associates, 1987.
9. D.L. Hartl and A.G. Clark. *Principles of Poppulation Genetics.* Sinauer Associates, 1 edition, 1989.
10. J. Horn. Finite markov chain analysis of genetic algorithms with niching. In S. Forrest, editor, *Proceedings of the Fifth International Conference on Genetic Algorithms*, pages 110–117. Morgan Kaufmann Publishers, 1993.
11. A.E. Nix and M.D. Vose. Modeling genetic algorithms with markov chains. *Annals of Mathematics and Artificial Intelligence*, 5:78–88, 1992.
12. A. Papoulis and S.U. Pillai. *Probability, Random Variables, and Stochastic Processes.* McGraw-Hill, New York, NY, USA, 4 edition, 2002. ISBN 0-07-048477-5.
13. G. Rudolph. Convergence analysis of canonical genetic algorithms. *IEEE Transactions on Neural Networks*, 5(1):96–101, 1994.
14. C.A. Schippers. Multi-parent scanning crossover and genetic drift. In *Theoretical Aspects of Evolutionary Computing*, pages 307–330. Springer, 2001.
15. S. Tsutsui and A. Ghosh. A study on the effect of multi-parent recombination in real coded genetic algorithms. In *Proceedings of International Conference on Evolutionary Computation*, pages 828–833, 1998.