THEORY OF GENETIC ALGORITHMS WITH α-SELECTION

André Neubauer

Information Processing Systems Lab
Münster University of Applied Sciences
Stegerwaldstraße 39, D-48565 Steinfurt, Germany
Email: `andre.neubauer@fh-muenster.de`

ABSTRACT

Genetic algorithms are random heuristic search (RHS) algorithms for adaptive systems with a wide range of applications in search, optimisation, pattern recognition and machine learning as well as signal processing. Despite their widespread use a general theory is still lacking. A promising approach is offered by the dynamical system model which describes the stochastic trajectory of a population under the dynamics of a genetic algorithm with the help of an underlying deterministic heuristic function and its fixed points. However, even for the simple genetic algorithm (SGA) with fitness-proportional selection, crossover and mutation the determination of the population trajectory and the fixed points of the heuristic function is unfeasible for practical problem sizes. In order to simplify the mathematical analysis α-selection is introduced in this paper. Based on this strong selection scheme it is possible to derive the dynamical system model and the respective fixed points in closed form. In addition to the theoretical analysis experimental results are presented.

Index Terms— Genetic algorithm, α-selection, random heuristic search, dynamical system model.

1. INTRODUCTION

Genetic algorithms mimic the basic mechanisms of biological evolution and molecular genetics in simplified form. These random heuristic search (RHS) algorithms are based on populations of individual solutions which evolve according to selection and genetic operators like crossover and mutation. Since they were proposed by HOLLAND [1] genetic algorithms have been used in a wide range of applications, e.g. for search as well as numerical and combinatorial optimisation problems [2]. Despite their successful applications genetic algorithms still lack a general and stringent theory which guides the practitioner in tailoring the algorithm to a specific application domain [3]. The first attempt to theoretically describe the simple genetic algorithm (SGA) was the so-called *schema theorem* [1] which gives a lower bound for the expected number of individuals representing particular hyperplanes or *schemata* in the search space within the next gen-

eration. Recently, the schema theory and its implications like the so-called *building block-hypothesis* have been criticised for lack of explanatory power [3]. As an alternative, a dynamical system model was introduced for the SGA by VOSE and his co-workers [4, 5, 7, 8]. In this model the SGA's stochastic dynamics is modeled according to an underlying deterministic heuristic function which describes the expected population trajectory of the SGA. According to theoretical and empirical evidence this population trajectory is attracted by the fixed points of the heuristic function. The determination of these fixed points, however, becomes unfeasible for practical problem sizes. In order to simplify the mathematical analysis α-selection is introduced in this paper.

The paper is organised as follows. After a description of the SGA over search spaces of binary ℓ-tuples in section 2, the dynamical system model is summarised in section 3. The underlying heuristic function is formulated for the SGA with fitness-proportional selection, 1-point or uniform crossover and bitwise mutation. The genetic algorithm with α-selection is introduced in section 4 based on the notion of the best individual randomly mating with all other individuals in the current population. The dynamical system model of a genetic algorithm with α-selection is derived in section 5. The fixed points of the corresponding heuristic function are calculated analytically and compared to simulation results in section 6 showing a close agreement between theory and experiment. A brief summary and conclusion is given in section 7.

2. SIMPLE GENETIC ALGORITHM

In this section, the SGA with fitness-proportional selection, 1-point or uniform crossover and bitwise mutation [7] is described which is used for the maximisation of a fitness function f. This fitness function f is defined over the search space $\Omega = \mathbb{Z}_2^\ell = \{0, 1\}^\ell$ consisting of binary ℓ-tuples $a = (a_0, a_1, \ldots, a_{\ell-1})$, i.e. $f : \Omega \to \mathbb{R}$.

2.1. Algorithm

The SGA works over populations defined as multisets $P(t) = (a_0(t), a_1(t), \ldots, a_{r-1}(t))$ consisting of r individual binary

ℓ-tuples $a_k(t) = (a_{k,0}(t), a_{k,1}(t), \ldots, a_{k,\ell-1}(t)) \in \Omega$ with fitness values $f(a_k(t))$. For the creation of offspring individuals in each generation t random genetic operators like crossover χ_Ω and mutation μ_Ω are applied to parental individuals which are selected according to their fitness values as shown in Fig. 1. The population $P(0)$ is initialised appropriately, e.g. by randomly choosing individuals in Ω.

```
t := 0;
while end of adaptation ≠ true do
    for k = 0 to r − 1 do
        select parental ℓ-tuples b(t) and c(t);
        apply crossover χ_Ω and mutation μ_Ω
            a_k(t + 1) := μ_Ω (χ_Ω (b(t), c(t)));
        evaluate fitness f (a_k(t + 1));
    end
    t := t + 1;
end
```

Fig. 1. Simple genetic algorithm (SGA).

2.2. Selection

For fitness-proportional selection each individual $a_k(t)$ in the population $P(t)$ is selected with a probability $p_k(t)$ which is directly proportional to its fitness $f(a_k(t)) \geq 0$. The individual selection probability is then given by

$$p_k(t) = \frac{f(a_k(t))}{\sum_{j=0}^{r-1} f(a_j(t))} \quad .$$

2.3. Crossover

The crossover operator $\chi_\Omega : \Omega \times \Omega \rightarrow \Omega$ takes two selected individual ℓ-tuples $a = (a_0, a_1, \ldots, a_{\ell-1})$ and $b = (b_0, b_1, \ldots, b_{\ell-1})$ and randomly generates an offspring ℓ-tuple $c = (c_0, c_1, \ldots, c_{\ell-1})$ according to $c = \chi_\Omega(a, b)$.

For *1-point crossover* a position $1 \leq \lambda \leq \ell - 1$ is chosen with uniform probability and the two ℓ-tuples $(a_0, \ldots, a_{\lambda-1}, b_\lambda, \ldots, b_{\ell-1})$ and $(b_0, \ldots, b_{\lambda-1}, a_\lambda, \ldots, a_{\ell-1})$ are generated one of which is randomly chosen as offspring c. The 1-point crossover operator is applied with crossover probability χ. In case of *uniform crossover* each bit of the offspring c is chosen with probability χ from the parental ℓ-tuple a and with probability $1 - \chi$ from the parental ℓ-tuple b or vice versa.

2.4. Mutation

In the SGA with binary ℓ-tuples the bitwise mutation operator $\mu_\Omega : \Omega \rightarrow \Omega$ is defined by randomly flipping each bit of the individual ℓ-tuple $a = (a_0, a_1, \ldots, a_{\ell-1})$ with small mutation probability μ. A typical value is $\mu \sim \frac{1}{\ell}$.

3. DYNAMICAL SYSTEM MODEL FOR THE SIMPLE GENETIC ALGORITHM

In view of the search space Ω of binary ℓ-tuples we identify each binary ℓ-tuple $(a_0, a_1, \ldots, a_{\ell-1})$ with the integer number $a = a_0 \cdot 2^{\ell-1} + a_1 \cdot 2^{\ell-2} + \ldots + a_{\ell-2} \cdot 2^1 + a_{\ell-1} \cdot 2^0$. The respective fitness is given by $f_a = f(a)$. This leads to the search space $\Omega = \{0, 1, \ldots, n-1\}$ with cardinality $|\Omega| = n = 2^\ell$. Based on this binary number representation we define the bitwise modulo-2 addition $a \oplus b$, bitwise modulo-2 multiplication $a \otimes b$ and bitwise binary complement $\overline{a}$, e.g. $3 \oplus 5 = 011 \oplus 101 = 110 = 6$, $3 \otimes 5 = 011 \otimes 101 = 001 = 1$ and $\overline{3} = \overline{011} = 100 = 4$ for ℓ-tuples of length $\ell = 3$. Vice versa, the integer $a \in \Omega$ is also viewed as a column vector $(a_0, a_1, \ldots, a_{\ell-1})^\mathrm{T}$ so that we can define the scalar product $a^\mathrm{T} b = a_0 \cdot b_0 + a_1 \cdot b_1 + \ldots + a_{\ell-2} \cdot b_{\ell-2} + a_{\ell-1} \cdot b_{\ell-1}$. In the following we will also make use of the all-one ℓ-tuple $\mathbf{1}$ which corresponds to the integer $n-1 = 2^\ell - 1$. Furthermore, the indicator function $[i = j] = \delta_{i,j}$ is defined by $[i = j] = 1$ if $i = j$ and 0 if $i \neq j$.

In the dynamical system model [7] the SGA's dynamics is compactly formulated by defining the population vector $\boldsymbol{p} = (p_0, p_1, \ldots, p_{n-1})^\mathrm{T}$. Each component $p_i = \frac{1}{r} \sum_{j \in P}[j = i]$ gives the proportion of the element $i \in \Omega$ in the current population P. The population vector $\boldsymbol{p}$ is an element of the simplex $\Lambda = \{\boldsymbol{p} \in \mathbb{R}^n : p_i \geq 0 \wedge \sum_{i \in \Omega} p_i = 1\}$. For a population of size r the number of possible population vectors is given by $\binom{n+r-1}{r}$. In the limit of infinite populations with $r \rightarrow \infty$ the population vectors are dense in the simplex Λ. For simplicity we will take the simplex Λ as the defining region of the population vector $\boldsymbol{p}$. Therefore, the results obtained are strictly valid only for large populations with $r \gg 1$.

The SGA is now described as an instance of RHS $\tau : \Lambda \rightarrow \Lambda$ according to $\boldsymbol{p}(t + 1) = \tau(\boldsymbol{p}(t))$ with τ depending on the random selection and genetic operators. As outlined in [7] τ can be equivalently represented by a suitable heuristic function $\mathcal{G} : \Lambda \rightarrow \Lambda$ which for a given population vector $\boldsymbol{p}$ yields the probability distribution $\mathcal{G}(\boldsymbol{p})$. This probability distribution $\mathcal{G}(\boldsymbol{p})_i = \Pr\{\text{individual } i \text{ is sampled from } \Omega\}$ is sampled to generate the next population as illustrated in Fig. 2. For the SGA the heuristic function $\mathcal{G} = \mathcal{M} \circ \mathcal{F}$ is given by the composition of the selection operation $\mathcal{F}$ and the mixing operation $\mathcal{M}$ comprising crossover χ_Ω and mutation μ_Ω. The transition probabilities of the RHS τ are given by the formula

$$\Pr\{\tau(\boldsymbol{p}) = \boldsymbol{q}\} = r! \prod_{i \in \Omega} \frac{\mathcal{G}(\boldsymbol{p})_i^{rq_i}}{(rq_i)!} \quad .$$

The trajectory $\boldsymbol{p}$, $\tau(\boldsymbol{p})$, $\tau^2(\boldsymbol{p})$, $\ldots$ approximately follows the trajectory $\boldsymbol{p}$, $\mathcal{G}(\boldsymbol{p})$, $\mathcal{G}^2(\boldsymbol{p})$, $\ldots$ of the deterministic dynamical system defined by the heuristic function $\mathcal{G}$ with $\mathrm{E}\{\tau(\boldsymbol{p})\} = \mathcal{G}(\boldsymbol{p})$. Because of the mean quadratic deviation $\mathrm{E}\{\|\tau(\boldsymbol{p}) - \mathcal{G}(\boldsymbol{p})\|^2\} = \frac{1}{r}(1 - \|\mathcal{G}(\boldsymbol{p})\|^2)$ the RHS τ behaves like the deterministic dynamical system model in the limit of
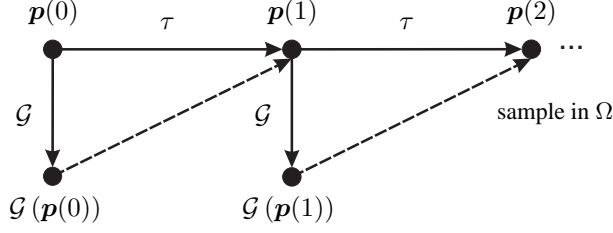
Fig. 2. Genetic algorithm as RHS τ with heuristic function $\mathcal{G}$.

infinite populations with $r \to \infty$. According to experimental evidence the heuristic function $\mathcal{G}$ shows punctuated equilibria, i.e. phases of relative stability nearby the fixed points $\boldsymbol{\omega} = \mathcal{G}(\boldsymbol{\omega})$ of the heuristic function $\mathcal{G}$ disrupted by sudden transitions to another dynamical equilibrium near another fixed point. We call this the *fixed point hypothesis* of genetic algorithms.

3.1. Selection

For fitness-proportional selection the corresponding heuristic function $\mathcal{F}(\boldsymbol{p})_i = \Pr\{\text{individual } i \text{ selected}\}$ is given by [7]

$$\mathcal{F}(\boldsymbol{p})_i = \frac{f_i \cdot p_i}{\sum_{j \in \Omega} f_j \cdot p_j} = \frac{f_i \cdot p_i}{\boldsymbol{f}^{\mathrm{T}} \boldsymbol{p}}$$

with fitness vector $\boldsymbol{f} = (f_0, f_1, \ldots, f_{n-1})^{\mathrm{T}}$ and individual fitness values $f_i \geq 0$.

3.2. Crossover

For the crossover operator χ_Ω applied to the individual ℓ-tuples a and b with crossover probability χ we make use of a random crossover mask $m \in \Omega$ according to $\chi_\Omega(a, b) = a \otimes m \oplus \overline{m} \otimes b$ or $\chi_\Omega(a, b) = a \otimes \overline{m} \oplus m \otimes b$ each with probability $\frac{1}{2}$. For 1-point crossover the crossover mask m is randomly chosen from Ω according to the probability distribution [7]

$$\chi_m = \left\{ \begin{array}{ccl} 1 - \chi & , & m = 0 \\ \chi/(\ell - 1) & , & m = 2^\lambda - 1 \text{ with } 1 \leq \lambda \leq \ell - 1 \\ 0 & , & \text{otherwise} \end{array} \right. .$$

In case of uniform crossover the probability distribution is given by [7]

$$\chi_m = \left\{ \begin{array}{ccl} 1 - \chi + \chi \cdot 2^{-\ell} & , & m = 0 \\ \chi \cdot 2^{-\ell} & , & m > 0 \end{array} \right. .$$

3.3. Mutation

The mutation operator μ_Ω can be described with the help of the random mutation mask $m \in \Omega$ according to $\mu_\Omega(a) = a \oplus m$. The mutation mask m is randomly chosen from Ω according to the probability distribution [7]

$$\mu_m = \mu^{\mathbf{1}^{\mathrm{T}} m} \cdot (1 - \mu)^{\ell - \mathbf{1}^{\mathrm{T}} m} .$$

3.4. Heuristic

The mixing operation $\mathcal{M}$ comprises crossover χ_Ω and mutation μ_Ω. With the help of the probability distributions for crossover and mutation masks we obtain

$$\Pr\{c = \mu_\Omega(\chi_\Omega(a, b))\} = \sum_{j \in \Omega} \mu_j \cdot \Pr\{\chi_\Omega(a, b) = c \oplus j\}$$

$$= \sum_{j \in \Omega} \mu_j \sum_{i \in \Omega} \frac{\chi_i + \chi_{\bar{i}}}{2} \cdot [a \otimes i \oplus \bar{i} \otimes b = c \oplus j] .$$

This yields $\Pr\{c = \mu_\Omega(\chi_\Omega(a, b))\} = M_{a \oplus c, b \oplus c}$ with the help of the $n \times n$ mixing matrix

$$M_{i,j} = \sum_{u,v \in \Omega} \mu_v \cdot \frac{\chi_u + \chi_{\bar{u}}}{2} \cdot [i \otimes u \oplus \bar{u} \otimes j = v] .$$

The heuristic function $\mathcal{G}$ follows according to

$$\begin{aligned} \mathcal{G}(\boldsymbol{p})_i &= \Pr\{\text{individual } i \text{ is sampled from } \Omega\} \\ &= \sum_{u,v \in \Omega} \mathcal{F}(\boldsymbol{p})_u \cdot \mathcal{F}(\boldsymbol{p})_v \cdot \Pr\{\mu_\Omega(\chi_\Omega(u, v)) = i\} \\ &= \sum_{u,v \in \Omega} \mathcal{F}(\boldsymbol{p})_u \cdot \mathcal{F}(\boldsymbol{p})_v \cdot M_{u \oplus i, v \oplus i} \\ &= \sum_{u,v \in \Omega} \mathcal{F}(\boldsymbol{p})_{u \oplus i} \cdot \mathcal{F}(\boldsymbol{p})_{v \oplus i} \cdot M_{u,v} . \end{aligned}$$

By defining the mixing operation $\mathcal{M} : \Lambda \to \Lambda$ according to

$$\mathcal{M}(\boldsymbol{p})_i = \sum_{u,v \in \Omega} p_{u \oplus i} \cdot p_{v \oplus i} \cdot M_{u,v}$$

the heuristic function is given by the composition $\mathcal{G} = \mathcal{M} \circ \mathcal{F}$.

In order to determine the heuristic function $\mathcal{G}$ the mixing matrix M has to be calculated. This can be done efficiently with the help of the WALSH transform [6]. For a matrix A the WALSH transform is $\widehat{A} = W \cdot A \cdot W$ with the orthogonal $n \times n$ WALSH matrix $W_{i,j} = n^{-1/2} \cdot (-1)^{i^{\mathrm{T}} j}$. Correspondingly, the WALSH transform of a vector $\boldsymbol{x}$ yields $\widehat{\boldsymbol{x}} = W \cdot \boldsymbol{x}$. In Fig. 3 the WALSH matrix is illustrated for $n = 2^6 = 64$.

With the WALSH transform of the mutation mask distribution $\widehat{\mu}_i = n^{-1/2} \cdot (1 - 2\mu)^{\mathbf{1}^{\mathrm{T}} i}$ the WALSH transform of the mixing matrix M is given by

$$\widehat{M}_{i,j} = \frac{\sqrt{n}}{2} \cdot [i \otimes j = 0] \cdot \widehat{\mu}_{i \oplus j} \cdot \sum_{k \in \Omega_{\bar{i} \otimes \bar{j}}} (\chi_{k \oplus i} + \chi_{k \oplus j})$$

with $\Omega_k = \{i \in \Omega : i \otimes \bar{k} = 0\}$. Although the mixing matrix M can be calculated efficiently (for 1-point crossover a closed-form expression for the WALSH transform $\widehat{M}$ is known [7]) the fixed point analysis of the heuristic function $\mathcal{G}$ becomes infeasible even for moderate ℓ due to its dependence on the population vector $\boldsymbol{p}$ via $\mathcal{F}(\boldsymbol{p})$ and the cardinality $n = 2^\ell$ of the search space Ω. Therefore, in the next section we introduce the genetic algorithm with α-selection which yields a simpler heuristic function and even allows to derive a closed-form solution for the fixed points.
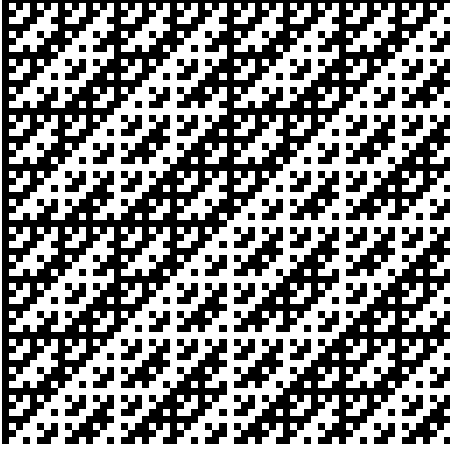
Fig. 3. Illustration of the WALSH matrix W for $n = 2^6 = 64$.

4. GENETIC ALGORITHM WITH α-SELECTION

Let the current population be described by the population vector $\boldsymbol{p}$ and let

$$b = \text{argmax}\,\{f_i : i \in \Omega \wedge p_i > 0\}$$

be the best individual in the current population. Clearly, this best individual or α-*individual* b depends on the population vector $\boldsymbol{p}$ (for ease of notation we omit this dependence in the following). In the genetic algorithm with α-selection the α-individual b is mated with individuals randomly chosen from the current population with uniform probability. The genetic algorithm with α-selection is shown in Fig. 4.

> $t := 0$;
> **while** end of adaptation $\neq$ true **do**
>   select α-individual $b(t)$ as first parent;
>   **for** $k = 0$ to $r - 1$ **do**
>   select second parent $c(t)$ randomly;
>   apply crossover χ_Ω and mutation μ_Ω
> $a_k(t+1) := \mu_\Omega\left(\chi_\Omega\left(b(t), c(t)\right)\right)$;
>   evaluate fitness $f\left(a_k(t+1)\right)$;
>   **end**
> $t := t + 1$;
> **end**

Fig. 4. Genetic algorithm with α-selection.

5. DYNAMICAL SYSTEM MODEL FOR THE GENETIC ALGORITHM WITH α-SELECTION

In the following we determine the dynamical system model for the genetic algorithm with α-selection by deriving the heuristic function $\mathcal{G}\left(\boldsymbol{p}\right)$ for a given population vector $\boldsymbol{p}$.

5.1. Heuristic

The α-individual b is selected as the first parent for creation of a new offspring, whereas the second parent is chosen uniformly at random from the current population according to the probability distribution $\text{Pr}\{\text{individual } j \text{ is selected}\} = p_j$. The heuristic function $\mathcal{G}$ is then given by

$$\begin{aligned}
\mathcal{G}\left(\boldsymbol{p}\right)_i &= \sum_{j \in \Omega} p_j \cdot \text{Pr}\{\mu_\Omega\left(\chi_\Omega(b, j)\right) = i\} \\
&= \sum_{j \in \Omega} p_j \cdot M_{b \oplus i, j \oplus i} \;.
\end{aligned}$$

With the $n \times n$ system matrix

$$A_{i,j} = M_{b \oplus i, j \oplus i}$$

we obtain the linear system of equations for the new population vector

$$\boldsymbol{q} = \mathcal{G}\left(\boldsymbol{p}\right) = A \cdot \boldsymbol{p} \;.$$

With the permutation matrix $(\sigma_b)_{i,j} = [i \oplus j = b]$ and the twist $(M^*)_{i,j} = M_{i \oplus j, i}$ of the symmetric mixing matrix $M = M^{\mathrm{T}}$ the system matrix A can also be expressed as

$$A = \sigma_b \cdot M^* \cdot \sigma_b \;.$$

Due to the α-selection scheme we gain a simpler heuristic function $\mathcal{G}$ which is completely described by the α-individual b and the mixing matrix M or system matrix A, respectively.

5.2. Fixed Points

Because of the linear relationship $\mathcal{G}\left(\boldsymbol{p}\right) = A \cdot \boldsymbol{p}$ in case of a given α-individual b the fixed points $\boldsymbol{\omega} = \mathcal{G}\left(\boldsymbol{\omega}\right)$ of the heuristic function $\mathcal{G}$ are obtained from the eigenvectors of the system matrix A to eigenvalue $\lambda = 1$, i.e. $A \cdot \boldsymbol{\omega} = \boldsymbol{\omega}$. To this end, we consider the WALSH transform of both sides of the equation $\boldsymbol{q} = A \cdot \boldsymbol{p}$ according to $\widehat{\boldsymbol{q}} = W \cdot \boldsymbol{q} = W \cdot A \cdot \boldsymbol{p} = W \cdot A \cdot W \cdot W \cdot \boldsymbol{p} = \widehat{A} \cdot \widehat{\boldsymbol{p}}$ where we have made use of the orthogonality of the WALSH matrix W. For an eigenvector $\boldsymbol{v}$ with eigenvalue λ we obtain $A \cdot \boldsymbol{v} = \lambda \cdot \boldsymbol{v} \Rightarrow \widehat{A} \cdot \widehat{\boldsymbol{v}} = \lambda \cdot \widehat{\boldsymbol{v}}$. The system matrix A and its WALSH transform $\widehat{A}$ have the same eigenvalues with eigenvectors which are also related by the WALSH transform. It is an easy but lengthy task to calculate the WALSH transform of the system matrix $A_{i,j} = M_{b \oplus i, j \oplus i}$ leading to

$$\widehat{A}_{i,j} = (-1)^{b^{\mathrm{T}}(i \oplus j)} \cdot \widehat{M}_{i \oplus j, j} \;.$$

The system matrix A as well as its WALSH transform $\widehat{A}$ depend on the α-individual b.

For crossover χ_Ω and mutation μ_Ω defined by masks it can be shown that the WALSH transform $\widehat{A}$ is a lower triangular matrix with eigenvalues λ_i given by the diagonal elements $\lambda_i = \widehat{A}_{i,i} = \widehat{M}_{0,i}$. This yields the eigenvalues

$$\lambda_i = \frac{(1 - 2\mu)^{\mathbf{1}^{\mathrm{T}} i}}{2} \cdot \sum_{k \in \Omega_{\bar{i}}} \left(\chi_k + \chi_{k \oplus i}\right)$$

of the system matrix A. It is a remarkable fact that the eigenvalues do not depend on the α-individual b. Taking into account $\sum_{i \in \Omega} \chi_i = 1$ and $\mu < \frac{1}{2}$ it is evident from this expression that the eigenvalues are bounded by $0 \leq \lambda_i \leq 1$. The system matrix A and its WALSH transform $\widehat{A}$ are therefore positive semidefinite. Because of $\lambda_0 = 1$ and $0 \leq \lambda_i \leq 1-2\mu$ for $1 \leq i \leq n-1$ there exists a single eigenvector $\boldsymbol{\omega}$ which is a fixed point of the heuristic function $\boldsymbol{\omega} = \mathcal{G}(\boldsymbol{\omega}) = A \cdot \boldsymbol{\omega}$. Both the system matrix A and the fixed point $\boldsymbol{\omega}$ depend on the α-individual b. According to the *fixed point hypothesis* the population will stay near this fixed point $\boldsymbol{\omega}$. If a new best or α-individual $b \in \Omega$ is found the genetic algorithm will converge to the new fixed point thus showing punctuated equilibria, i.e. phases of relative stability nearby a fixed point disrupted by sudden transitions to another fixed point.

The fixed point $\boldsymbol{\omega}$ can be determined explicitly with the help of the WALSH transform. First, we note that due to $\widehat{\boldsymbol{\omega}} = \widehat{A} \cdot \widehat{\boldsymbol{\omega}}$ and the lower triangular matrix $\widehat{A}_{i,j} = (-1)^{b^{\mathrm{T}}(i \oplus j)} \cdot \widehat{M}_{i \oplus j, j}$ we can recursively calculate the WALSH transform of the fixed point according to

$$\widehat{\omega}_i = \frac{1}{1 - \widehat{A}_{i,i}} \cdot \sum_{j=0}^{i-1} \widehat{A}_{i,j} \cdot \widehat{\omega}_j$$

for $1 \leq i \leq n-1$ starting with $\widehat{\omega}_0 = n^{-1/2}$ which ensures that $\sum_{i \in \Omega} \omega_i = 1$. The fixed point is then obtained via the inverse WALSH transform $\boldsymbol{\omega} = W \cdot \widehat{\boldsymbol{\omega}}$.

6. EXPERIMENTAL RESULTS

For the shifted and inverted ONEMAX problem with fitness function $f_i = \mathbf{1}^{\mathrm{T}}(\overline{i \oplus i_0})$ with $i_0 = 128$ (corresponding to the number of 0's in the binary representation of $i \oplus i_0 \in \Omega$) Fig. 5 shows the theoretical and estimated population vectors $\boldsymbol{p}$ in generation $t = 25$. Here, we have assumed a genetic algorithm with α-selection for $\ell = 8$ and $n = 2^8 = 256$ using 1-point crossover with crossover probability $\chi = 1$, bitwise mutation with mutation probability $\mu = \frac{1}{\ell}$ and population size $r = 100$. The EUCLIDean distance of the simulated population vector $\boldsymbol{p}$ and the fixed point $\boldsymbol{\omega}$ is 0.0932. There is a close match between the theoretical prediction and the experimental result.

7. CONCLUSION

The dynamical system model represents the current state of the art of genetic algorithm theory [3, 7]. For practical problem sizes, however, the determination of the population trajectory and the fixed points of the underlying heuristic function $\mathcal{G}$ becomes unfeasible. In this paper, α-selection has been introduced in order to simplify the mathematical analysis and to enable the derivation of the dynamical system model and the fixed points of the heuristic function in closed form. Experimental results show close agreement to the theoretical
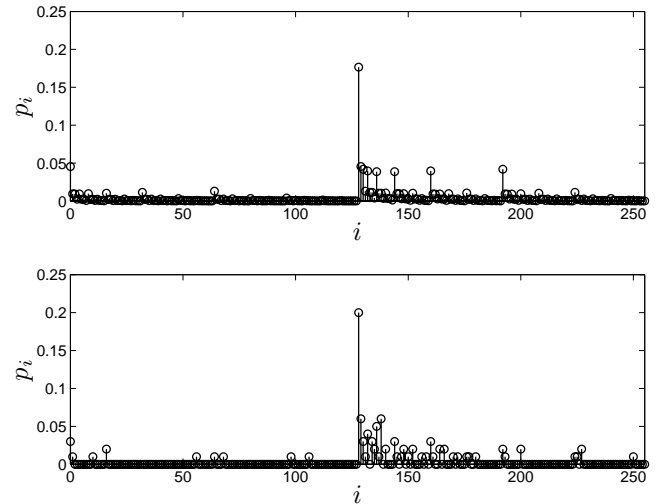

Fig. 5. Population vector according to the dynamical system model with heuristic function $\mathcal{G}$ (top) and its estimate from a single run of the genetic algorithm with α-selection (bottom).

predictions. The dynamical system model for the genetic algorithm with α-selection depends on the α-individual b. If in the next generation this individual is lost or a better individual is sampled from the search space Ω the heuristic function $\mathcal{G}$ changes due to the dependence of the system matrix A on the α-individual b, thus showing punctuated equilibria. Since α-selection is a strong selection scheme future research will be oriented towards genetic algorithms with weaker selection.

8. REFERENCES

[1] HOLLAND, J.H.: *Adaptation in Natural and Artificial Systems – An Introductory Analysis with Applications to Biology, Control, and Artificial Intelligence.* Cambridge: First MIT Press Edition, 1992

[2] MITCHELL, M.: *An Introduction to Genetic Algorithms.* Cambridge: MIT Press, 1996

[3] REEVES, C.R.; ROWE, J.E.: *Genetic Algorithms – Principles and Perspectives, A Guide to GA Theory.* Boston: Kluwer Academic Publishers, 2003

[4] VOSE, M.D.; WRIGHT, A.H.: *Simple Genetic Algorithms with Linear Fitness.* In: *Evolutionary Computation.* Vol. 2, No. 4, pp. 347-368, 1994

[5] VOSE, M.D.: *Modeling Simple Genetic Algorithms.* In: *Evolutionary Computation.* Vol. 3, No. 4, pp. 453-472, 1995

[6] VOSE, M.D.; WRIGHT, A.H.: *The Simple Genetic Algorithm and the Walsh Transform – Part I, Theory.* In: *Evolutionary Computation.* Vol. 6, No. 3, pp. 253-273, 1998

[7] VOSE, M.D.: *The Simple Genetic Algorithm – Foundations and Theory.* Cambridge: MIT Press, 1999

[8] VOSE, M.D.: *Random Heuristic Search.* In: *Theoretical Computer Science.* Vol. 229, No. 1-2, pp. 103-142, 1999