

A Partial Order Approach to Noisy Fitness Functions

Günter Rudolph

Department of Computer Science

University of Dortmund

44221 Dortmund / Germany

rudolph@LS11.cs.uni-dortmund.de

Abstract-

If the fitness values are perturbed by noise then they do not have a definitive total order. As a consequence, traditional selection procedures in evolutionary algorithms may lead to obscure solutions. A potential remedy is as follows: Construct a partial order on the set of noisy fitness values and apply those evolutionary algorithms that have been designed for finding the minimal elements of partially ordered sets. These minimal elements are the only reasonable candidates for the unperturbed true solution. A method for reducing the number of candidate solutions is suggested. From a theoretical point of view it is worth mentioning that all convergence results for evolutionary algorithms with partially ordered fitness sets remain valid for the approach considered here.

1 Introduction

The Gaussian distribution is the predominant choice for modeling noise frequently observable in measurements of various kinds. Here, we hold the view that a noise distribution with unbounded support (like the Gaussian, Cauchy, Laplace, Logistic, and others) may be quite unrealistic. Actually it is at least equally plausible to assume that the noise cannot exceed certain limits due to technical characteristics of the involved measurement unit. Even if a distributional shape close to a Gaussian appears reasonable we can resort to a symmetrical Beta distribution which can converge weakly to a Gaussian distribution under continuously increasing but bounded support (see e.g. Evans et al. 1993, p. 36). This assumption will have significant theoretical and practical impacts on the evolutionary algorithms (EAs) considered here.

Traditional measures for coping with noisy fitness functions in evolutionary algorithms include the resampling of the random fitness value with averaging, the appropriate adjustment (i.e., enlargement) of the population size, and in case of continuous search spaces also the rescaling of inherited mutations; see Beyer (2000) for a summary of work on EAs for noisy fitness functions.

Here, we add yet another avenue for dealing with noisy fitness functions: Instead of using a selection procedure that is based on the totally ordered set of noisy fitness values we endow the probabilistic fitness set with an appropriate partial order and deploy EAs with those selection methods being explicitly designed for coping with arbitrary partially ordered fitness sets (Rudolph 1998, 2001; Rudolph and Agapie 2000).

Section 2 offers a brief introduction to partially ordered

sets in general and in particular to interval orders (Fishburn 1985) which constitute the first step towards the partial order to be used later on. Since the noise is supposed to have bounded support we can easily equip these intervals with a probability measure (representing the noise distribution). Thus, the interval order turns to a partial order on random variables. This subject is detailed in Section 3 which also contains the presentation of the EA along with a discussion of the inherited theoretical properties from the general case (Rudolph 2001). Preliminary experimental results can be found in Section 4.

2 Partially Ordered Sets

Let \mathcal{F} be a set. A reflexive, antisymmetric, and transitive relation “ \preceq ” on \mathcal{F} is termed a *partial order relation* whereas a *strict partial order relation* “ \prec ” must be antireflexive, asymmetric, and transitive. The latter relation may be obtained by the former relation by setting $x \prec y := (x \preceq y) \wedge (x \neq y)$. After these preparations one is in the position to turn to the actual objects of interest.

Definition 1 *Let \mathcal{F} be some set. If the partial order relation “ \preceq ” is valid on \mathcal{F} then the pair (\mathcal{F}, \preceq) is called a partially ordered set (or short: poset). If $x \prec y$ for some $x, y \in \mathcal{F}$ then x is said to dominate y . Distinct points $x, y \in \mathcal{F}$ are said to be comparable when either $x \prec y$ or $y \prec x$. Otherwise, x and y are incomparable which is denoted by $x \parallel y$. If each pair of distinct points of a poset (\mathcal{F}, \preceq) is comparable then (\mathcal{F}, \preceq) is called a totally ordered set or a chain. Dually, if each pair of distinct points of a poset (\mathcal{F}, \preceq) are incomparable then (\mathcal{F}, \preceq) is termed an antichain. \square*

For example, let $\Pi = \{[x_1, x_2] \subset \mathbb{R} : x_1 \leq x_2\}$ be the set of closed intervals of \mathbb{R} and define

$$\begin{aligned} [x_1, x_2] \prec [y_1, y_2] & \text{ iff } x_2 < y_1, \\ [x_1, x_2] = [y_1, y_2] & \text{ iff } x_1 = y_1 \wedge x_2 = y_2, \\ [x_1, x_2] \preceq [y_1, y_2] & \text{ iff } x \prec y \vee x = y. \end{aligned}$$

It is easily seen that (Π, \preceq) is a partially ordered set in which distinct intervals with a nonvoid intersection are incomparable. Similarly, the infinitely large but countable set $(\Pi_\varepsilon, \preceq)$ with $\Pi_\varepsilon = \{[x - \varepsilon, x + \varepsilon] \subset \mathbb{R} : x \in \mathbb{N}_0\}$ with $\varepsilon > 1/2$ is a poset with incomparable elements whereas $(\Pi_\varepsilon, \preceq)$ with $\varepsilon < 1/2$ is totally ordered and therefore a chain. An example for an antichain is the set of “minimal elements” introduced next.

Definition 2 An element $x^* \in \mathcal{F}$ is called a minimal element of the poset (\mathcal{F}, \preceq) if there is no $x \in \mathcal{F}$ such that $x \prec x^*$. The set of all minimal elements, denoted $\mathcal{M}(\mathcal{F}, \preceq)$, is said to be complete if for each $x \in \mathcal{F}$ there is at least one $x^* \in \mathcal{M}(\mathcal{F}, \preceq)$ such that $x^* \preceq x$. \square

In contrast to infinitely large posets the completeness of $\mathcal{M}(\mathcal{F}, \preceq)$ is guaranteed for finitely large posets. Of course, completeness of infinitely large posets is not precluded. For example, the set $(\Pi_\varepsilon, \preceq)$ with $\varepsilon = 2/3$ is infinitely large and the set of minimal elements $\mathcal{M}(\Pi_\varepsilon, \preceq) = \{[-\frac{2}{3}, \frac{2}{3}], [\frac{1}{3}, \frac{5}{3}]\}$ is complete.

3 Coping with Noisy Fitness Functions

3.1 Assumptions

Let S be the finite search set and assume that the deterministic fitness function $f : S \rightarrow \mathbb{R}$ is perturbed by additive noise Z , i.e., $\tilde{f}(x) = f(x) + Z$ for $x \in S$. As mentioned earlier, here we insist that random variable Z has bounded and known support in form of a closed interval of \mathbb{R} . For example, Z may have a uniform or symmetric beta distribution on its support $[-a, a]$ with $a > 0$.

At first it is assumed that every point/individual $x \in S$ is evaluated only once. Later on this assumption is dropped.

3.2 Partial Order Approach

When an individual $x \in S$ is evaluated via $\tilde{f}(x) = f(x) + Z$ then the noisy fitness value is an element of the interval $[f(x) - a, f(x) + a]$. Since the EA only has knowledge of the support bound $a > 0$ and in no case of the true fitness value $f(x)$, the noisy evaluation of $x \in S$ only leads to the information that the true fitness value $f(x)$ must be in the interval $[\tilde{f}(x) - a, \tilde{f}(x) + a]$. Thus, each point or individual is associated with a realization of a random interval.

Next we declare a strict partial order on these intervals and thereby also a strict partial order on the individuals. Let $x, y \in S$ and w.l.o.g. $\tilde{f}(x) < \tilde{f}(y)$. If

$$\tilde{f}(x) + a < \tilde{f}(y) - a \quad (1)$$

then we define $\tilde{f}(x) \prec \tilde{f}(y)$ and thereby $x \prec y$. This choice is reasonable because we can immediately infer from $x \prec y$ that $f(x) < f(y)$ with probability 1. One should mention that this partial order is a special case of a partial order introduced in Guddat et al. (1985), p. 29. Moreover, notice that the connection to interval orders gets evident by the equivalence between equation (1) and

$$[\tilde{f}(x) - a, \tilde{f}(x) + a] \cap [\tilde{f}(y) - a, \tilde{f}(y) + a] = \emptyset. \quad (2)$$

Thus, whenever two intervals as those above have a nonvoid intersection then the noisy fitness values and therefore also the individuals are incomparable, in symbols: $\tilde{f}(x) \parallel \tilde{f}(y)$ resp. $x \parallel y$.

It remains to examine whether the set of minimal elements of such posets represents a reasonable and useful set of candidate solutions. For this purpose define

$$f^* = \min\{f \in \mathcal{F}\} \quad \text{with} \quad \mathcal{F} = \{f(x) : x \in S\} \quad \text{and} \\ \tilde{f}^* = \min\{\tilde{f} \in \tilde{\mathcal{F}}\} \quad \text{with} \quad \tilde{\mathcal{F}} = \{\tilde{f}(x) : x \in S\}.$$

In general, \tilde{f}^* and $\tilde{\mathcal{F}}$ are random objects. But since it is assumed that each element $x \in S$ is evaluated only once, one can hold the view that *each* element of S has been evaluated already *before* the EA is run such that the set $\tilde{\mathcal{F}}$ and the quantity \tilde{f}^* are deterministic during the run of the EA. In this manner one obtains a unique partial order on $\tilde{\mathcal{F}}$ and on S for each run. The set of minimal elements is then given by

$$\begin{aligned} \tilde{\mathcal{F}}^* &= \{\tilde{f} \in \tilde{\mathcal{F}} \mid \nexists \tilde{f}' \in \tilde{\mathcal{F}} : \tilde{f}' \prec \tilde{f}\} \\ &= \{\tilde{f} \in \tilde{\mathcal{F}} \mid \tilde{f} \leq \tilde{f}^* + 2a\}. \end{aligned} \quad (3)$$

Needless to say, it is reasonable to postulate that the noisy image $\tilde{f}(x^*)$ of an unperturbed optimal point $x^* \in S$ is contained in the set of minimal elements. As shown below, this requirement is fulfilled.

Theorem 1

For all $x^* \in S$ with $f(x^*) = f^*$ holds $\tilde{f}(x^*) \in \tilde{\mathcal{F}}^*$ regardless of the value of $a > 0$.

Proof: First notice the equivalence

$$\tilde{f}(x^*) \in \tilde{\mathcal{F}}^* \Leftrightarrow \tilde{f}(x^*) \leq \tilde{f}^* + 2a$$

which is easily deduced from equation (3). Since the support of random variable Z is $[-a, a]$ the relation

$$\tilde{f}(x^*) = f(x^*) + Z = f^* + Z \leq f^* + a \quad (4)$$

holds with probability 1. For the same reason one obtains $\tilde{f}^* > f^* - a \Leftrightarrow f^* < \tilde{f}^* + a$. Insertion in equation (4) leads to $\tilde{f}(x^*) \leq f^* + a < (\tilde{f}^* + a) + a = \tilde{f}^* + 2a$. \square

The next result offers an assessment of the 'solutions' contained in the set $\tilde{\mathcal{F}}^*$.

Theorem 2

$\max\{\tilde{f} \in \tilde{\mathcal{F}}^*\} \leq f^* + 3a$.
Proof: Owing to equation (3) each element of $\tilde{\mathcal{F}}^*$ is upper bounded by $\tilde{f}^* + 2a$. Since the support of Z is $[-a, a]$ one obtains $\tilde{f}^* \leq f^* + a$. Putting all together yields the desired inequality $\max\{\tilde{f} \in \tilde{\mathcal{F}}^*\} \leq \tilde{f}^* + 2a \leq f^* + 3a$. \square

Notice that under this partial order the set of minimal elements of a given finite population P is determinable in linear time: Find the individual with the smallest perturbed fitness value \tilde{f}^* in the population. This takes $\Theta(|P|)$ time. Each individual with fitness $\tilde{f} < \tilde{f}^* + 2a$ moves to the set of minimal elements. Since this takes $\Theta(|P|)$ time the entire run time is $\Theta(|P|)$.

3.3 The Base Algorithm

The pseudo code given in Figure 1 is taken from Rudolph (2001). Notice that an individual $p \in P_t$ of a population at generation $t \geq 0$ gathers all quantities of interest, i.e., $p \in S \times \tilde{\mathcal{F}} \times \dots$. The partial ordering of the individuals is based on their noisy fitness values $\tilde{f} \in \tilde{\mathcal{F}}$ of course. The expression $Q = \text{offspring}(P_t)$ in phase 1 encapsulates the operation of generating λ new individuals from the current population P_t of size μ with $\mu \leq \lambda$ at generation $t \geq 0$.

```

initialize  $P_0$ ; set  $t = 0$ 
repeat
  (* PHASE 1 *)
   $Q = \text{offspring}(P_t)$ 
   $Q^* = \mathcal{M}(Q, \preceq)$ 
   $Q = Q \setminus Q^*$ 
   $P' = Q' = \emptyset$ 
  (* PHASE 2 *)
  for each  $q \in Q^*$ :
     $D(q) = \{p \in P_t : q \prec p\}$ 
    if  $D(q) \neq \emptyset$  then
       $P_t = P_t \setminus D(q)$ 
       $P' = P' \cup \{q\}$ 
       $Q^* = Q^* \setminus \{q\}$ 
    endif
    if  $D(q) = \emptyset \wedge q \parallel p$  for all  $p \in P_t$  then
       $Q' = Q' \cup \{q\}$ 
       $Q^* = Q^* \setminus \{q\}$ 
    endif
  endfor
  (* PHASE 3 *)
   $P_{t+1} = P_t \cup P'$ 
  if  $|P_{t+1}| < \mu$  then
    fill  $P_{t+1}$  with elements from:
    1.  $Q'$ 
    2.  $Q^*$ 
    3.  $Q$ 
    until  $|P_{t+1}| = \mu$ 
  endif
   $t = t + 1$ 
until stopping criterion fulfilled

```

Figure 1: Pseudo code of the evolutionary algorithm with partially ordered fitness.

3.4 The Theoretical Property Inherited

Since the base algorithms' properties are valid for arbitrary partially ordered fitness sets, any instantiation inherits the theoretical properties from the general case (Rudolph 2001).

Theorem 3 *Let the search space of the base algorithm in Figure 1 be finite and the partial order of the fitness set be as described in Section 3.2. If every collection of offspring*

can be generated from an arbitrary collection of parents with some positive minimum probability, then the entire population will enter the set $\tilde{\mathcal{F}}^$ after a finite number of generations with probability 1 and stays there forever.* \square

As we know from Theorems 1 and 2 the set of minimal elements $\tilde{\mathcal{F}}^*$ contains the noisy version of the global minimum and each member of $\tilde{\mathcal{F}}^*$ is at most $3a$ away from the true minimum f^* . In this sense, we may call $\tilde{\mathcal{F}}^*$ also the set of ε -optimal solutions with $\varepsilon = 3a$.

3.5 The Virtue of Resampling Revisited

A $(3a)$ -optimal solution may be sufficient or may not. For the latter case one should look for a method of decreasing this bound. Here we use the technique of resampling that is common practice in EAs with noisy fitness values. Usually each point/individual is sampled several times and the thereby obtained noisy fitness values are averaged. This makes the estimator of the true fitness value more and more reliable by reducing its variance by a factor of $1/\sqrt{n}$ in case of n samples. Nevertheless, the estimator's value is afflicted with some uncertainty and every statement concerning its deviation from the true value may be false with some probability. For example, in case of Gaussian noise with known variance σ^2 we only can say that the true value f is in the interval

$$\left[\bar{f} - \frac{\sigma}{\sqrt{n}} \Phi^{-1} \left(\frac{1+\gamma}{2} \right), \bar{f} + \frac{\sigma}{\sqrt{n}} \Phi^{-1} \left(\frac{1+\gamma}{2} \right) \right]$$

with probability $\gamma > 0$. Here, \bar{f} denotes the average of n fitness samples and $\Phi^{-1}(\cdot)$ is the inverse of the cumulative distribution function (c.d.f.) of the standard normal distribution.

A moment reflection reveals that the partial order introduced in Section 3.2 is actually a partial ordering of confidence intervals. Since the support of the noise is bounded the confidence intervals for the true value have a confidence level of $\gamma = 1$ in contrast to $\gamma < 1$ for Gaussian noise. But here we don't use resampling for averaging; rather, the information gained from the additional samples is used to narrow the uncertainty interval containing the true fitness value.

Let $\tilde{f}_n = f + Z_n$ denote the n th sample of the noisy fitness function at a certain point in the search space. The first sample $\tilde{f}_1 = f + Z_1$ leads to the initial confidence interval $[\tilde{f}_1 - a, \tilde{f}_1 + a]$ for the true value f . Since each sample leads to a different confidence interval in general and f must be contained in each of these intervals we immediately obtain

$$\begin{aligned}
f &\in \bigcap_{k=1}^n [\tilde{f}_k - a, \tilde{f}_k + a] \\
&= \left[\max_{k \leq n} \{\tilde{f}_k\} - a, \min_{k \leq n} \{\tilde{f}_k\} + a \right] \\
&= \left[f + \max_{k \leq n} \{Z_k\} - a, f + \min_{k \leq n} \{Z_k\} + a \right] \\
&= [f, f] + [Z_{n:n} - a, Z_{1:n} + a] \tag{5}
\end{aligned}$$

where $Z_{k:n}$ denotes the k th smallest outcome of n samples in total. Thus, after n samples one knows for sure that the true value f is somewhere in the interval given in equation (5). The uncertainty interval $[Z_{n:n} - a, Z_{1:n} + a]$ shrinks to $[0, 0]$ for $n \rightarrow \infty$. The speed of narrowing can be determined as follows: Let $L_n = Z_{n:n} - a$ and $R_n = Z_{1:n} + a$. Then $|[L_n, R_n]|/(2a)$ is the relative size of the uncertainty or incomparability interval $[L_n, R_n]$ after n samples and the probability that it is then still larger than 100ε percent of its initial size is given by

$$\begin{aligned} \mathbf{P} \left\{ \frac{|[L_n, R_n]|}{2a} > \varepsilon \right\} &= \mathbf{P} \left\{ \frac{R_n - L_n}{2a} > \varepsilon \right\} \\ &= \mathbf{P} \left\{ 1 - \frac{Z_{n:n} - Z_{1:n}}{2a} > \varepsilon \right\} \\ &= \mathbf{P} \{ 1 - W_n > \varepsilon \} \\ &= \mathbf{P} \{ W_n < 1 - \varepsilon \} \end{aligned}$$

where $W_n = B_{n:n} - B_{1:n}$ is known as the *sample range* of n samples of random variable B . Here it is assumed that B has support $[0, 1]$ such that the noise $Z = a(2B - 1)$ has support $[-a, a]$. The distribution of the sample range can be determined via

$$\mathbf{P} \{ W_n \leq w \} = n \int_{-\infty}^{\infty} [F_B(x+w) - F_B(x)]^{n-1} f(x) dx$$

(see e.g. Arnold et al. 1992, p. 31) such that

$$\begin{aligned} \mathbf{P} \{ W_n < 1 - \varepsilon \} &= n(1 - \varepsilon)^{n-1} - (n-1)(1 - \varepsilon)^n \\ &\leq n(1 - \varepsilon)^{n-1} \end{aligned}$$

if B is the uniform distribution on $[0, 1]$. In this case the relative size of the uncertainty interval after n samples is given by $1 - W_n$ with $\mathbf{E}[1 - W_n] = 2/(n+1)$ and $\mathbf{V}[1 - W_n] \leq 2/n^2$. If B is a symmetric Beta random variable with probability density function

$$f_B(x) = \frac{\Gamma(2b)}{\Gamma(b)^2} x^{b-1} (1-x)^{b-1} \cdot 1_{[0,1]}(x) \quad (6)$$

then $\mathbf{E}[1 - W_n] \approx 2n^{-1/b}$. Thus, the closer B should resemble Gaussian noise ($b \rightarrow \infty$) the slower is the narrowing of the uncertainty or incomparability interval.

4 First Numerical Experiments

4.1 Instantiation of the Base Algorithm

The search space is a finite subset of the d -dimensional set of integers with box constraints. An individual is represented by a d -tuple of integers (the chromosomes) and the bounds of the confidence interval that is needed for the comparison of the individuals according to equation (2). Mutations of the chromosomes obey a bilateral geometrical distribution on the integers (Rudolph 1994) that is truncated at the box constraints. If the self-adaptation of the mutation distribution is

switched on then an additional parameter (chromosome) must be represented in the individual. The recombination of two chromosomes is realized by uniform crossover, the potential real-valued chromosomes for self-adaptation are averaged.

4.2 Realization of the Noise Generator

The noise used here is represented by a Beta random $Z = a(2B - 1)$. A Beta random variable B on $[0, 1]$ with parameter $b \in \mathbb{N}$ and probability density function as given in equation (6) can be generated via $B = G_1/(G_1 + G_2)$ where

$$G_1 = -\log \left(\prod_{i=1}^b U_i \right) \quad \text{and} \quad G_2 = -\log \left(\prod_{j=1}^b U_j \right)$$

with uniformly distributed random numbers U_i, U_j .

4.3 Preliminary Results

The search space is $S = \mathbb{Z}^{30} \cap [-1000, 1000]^{30}$ for the test function

$$f(x) = \sum_{i=1}^{30} \left(\frac{x_i}{10} \right)^2$$

with optimal solution $x^* = (0, \dots, 0)$ and $f(x^*) = 0$.

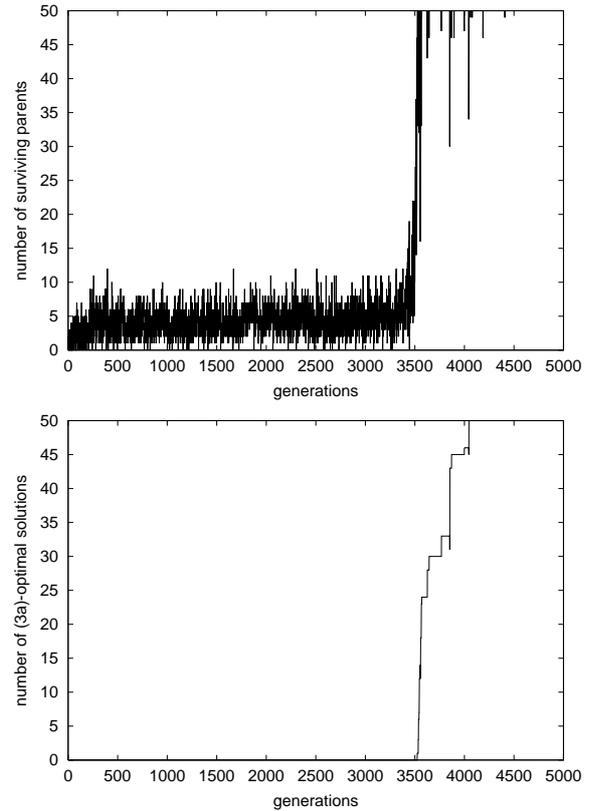


Figure 2: Number of surviving parents and $(3a)$ -optimal solutions for noise bound $a = 1$.

Figure 2 shows a typical run for a population size of $\mu = \lambda = 50$ and a support bound $a = 1$ for the uniformly distributed noise ($b = 1$), whereas the support bound was increased to $a = 5$ in Figure 3 and $a = 10$ in Figure 4.

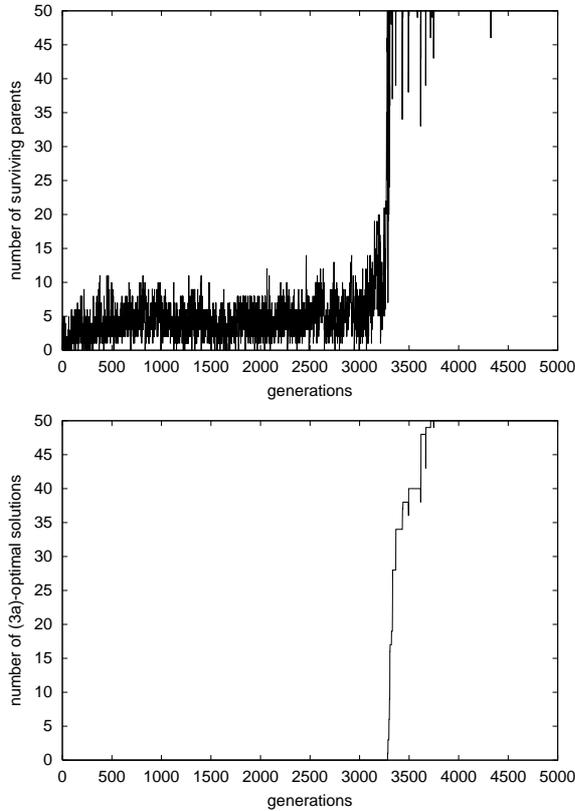


Figure 3: Number of surviving parents and $(3a)$ -optimal solutions for noise bound $a = 5$.

As can be seen in all three cases the number of surviving parents rapidly increases at about the maximum size as soon as the population enters the set of $(3a)$ -optimal solutions. It appears plausible that this happens the earlier the larger is the noise bound a , but it must be noted that there is no statistical support for this conjecture at the moment.

Nevertheless, we dare to use the number of surviving parents as an indicator for the event of entering the $(3a)$ -optimal set. When this happens then the individuals cannot be compared due to the size of the noise interval. Therefore, the re-sampling of the fitness values should begin right now. Many rules for the indicator mechanism are possible.

Here, we introduce a width parameter ω that is initially set to $\omega = 2a$. If two individuals are incomparable and at least one of them has a confidence interval larger than ω , then the individual with the largest interval is re-evaluated and its interval bounds are updated. This is repeated as long as the individuals are incomparable or both confidence intervals are smaller than ω . The decrease of ω by 10 % is triggered as soon as 90 % of the parents have survived, provided that $\omega \geq 1/100$. The EA is stopped as soon as $\omega < 1/100$ and no

parent has been replaced in the last selection process.

Of course, these parameters are chosen arbitrarily for the next experiment—other values may yield far better results. The identification of “good” parameter settings is not the goal here; rather, we like to gain first insights concerning the behavior of this EA.

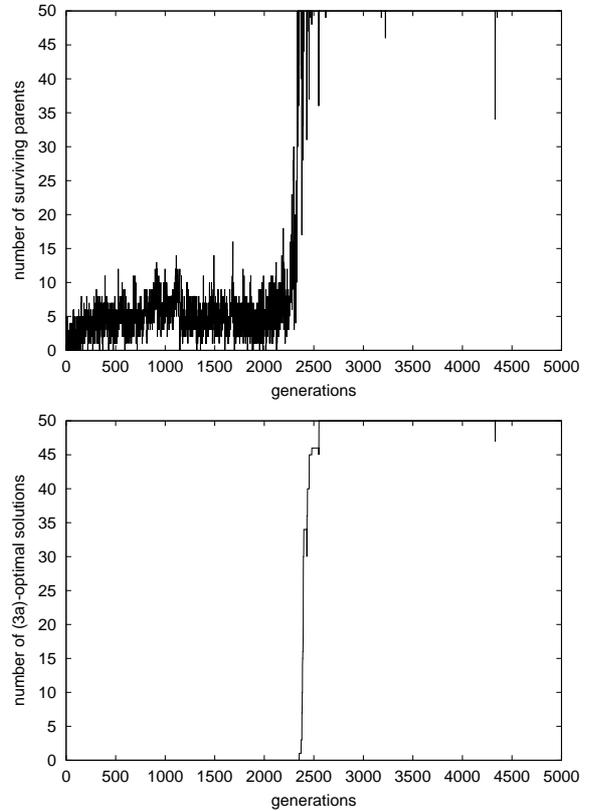


Figure 4: Number of surviving parents and $(3a)$ -optimal solutions for noise bound $a = 10$.

For this purpose the noise bound is set to $a = 10$ and the EA is run. Figure 5 shows some characteristic quantities recorded during a typical run of the EA.

At about generation 2000 the population begins to enter the set of 30 -optimal solutions. No re-evaluation of the individuals was triggered until now. The number of surviving parents oscillates considerably while the width parameter is continuously decreased. This leads to a rapid increase of the number of re-evaluations along with a significant improvement of the solutions as can be seen from the number of parents below certain (normally unavailable) true fitness values.

The number of re-evaluations, however, is much too large for practical use. This might be caused by an unlucky choice of the EA parameters. In the next experiment the width threshold was chosen a magnitude larger, namely, $\omega = 1/10$. Table 1 summarizes the results obtained from 50 independent runs. The number of re-evaluations is considerably smaller now, but the optimal true solution was never contained in the final population, in contrast to some other tests with $\omega = 1/100$.

	mean	std. dev.	skew.
generations	3231	740.6	0.46
re-evaluations	2.2×10^6	6.5×10^5	0.85
best f of last pop.	0.13	0.031	0.17
worst f of last pop.	0.25	0.039	0.46

Table 1: Summary of results for 50 runs with $\omega = 1/10$.

Needless to say, some parameter studies are necessary to find a useful parametrization of this EA.

Acknowledgments

This work was supported by the Deutsche Forschungsgemeinschaft (DFG) as part of the *Collaborative Research Center "Computational Intelligence" (SFB 531)*.

Bibliography

- B. C. Arnold, N. Balakrishnan, and H. N. Nagaraja (1992). *A First Course in Order Statistics*. New York: Wiley.
- H.-G. Beyer (2000). Evolutionary algorithms in noisy environments: Theoretical issues and guidelines for practice. *Computer Methods in Applied Mechanics and Engineering* 186(2-4), 239–267.
- M. Evans, N. Hastings, and B. Peacock (1993). *Statistical Distributions* (2nd ed.). New York: Wiley.
- P. C. Fishburn (1985). *Interval Orders and Interval Graphs: A Study of Partially Ordered Sets*. New York: Wiley.
- J. Guddat, F. Guerra Vasquez, K. Tammer, and K. Wendler (1985). *Multiobjective and Stochastic Optimization Based on Parametric Optimization*. Berlin: Akademie-Verlag.
- G. Rudolph (1994). An evolutionary algorithm for integer programming. In Y. Davidor, H.-P. Schwefel, and R. Männer (Eds.), *Parallel Problem Solving From Nature*, 3, pp. 139–148. Berlin and Heidelberg: Springer.
- G. Rudolph (1998). Evolutionary search for minimal elements in partially ordered finite sets. In V. W. Porto, N. Saravanan, D. Waagen, and A. E. Eiben (Eds.), *Evolutionary Programming VII, Proceedings of the 7th Annual Conference on Evolutionary Programming*, pp. 345–353. Berlin: Springer.
- G. Rudolph (2001). Evolutionary search under partially ordered fitness sets. In *Proceedings of the International Symposium on Information Science Innovations in Engineering of Natural and Artificial Intelligent Systems (ENAI 2001)*. ICSC Academic Press.
- G. Rudolph and A. Agapie (2000). Convergence properties of some multi-objective evolutionary algorithms. In A. Zalzal et al. (Eds.), *Proceedings of the 2000 Congress on Evolutionary Computation (CEC 2000), Vol. 2*, pp. 1010–1016. Piscataway (NJ): IEEE Press.

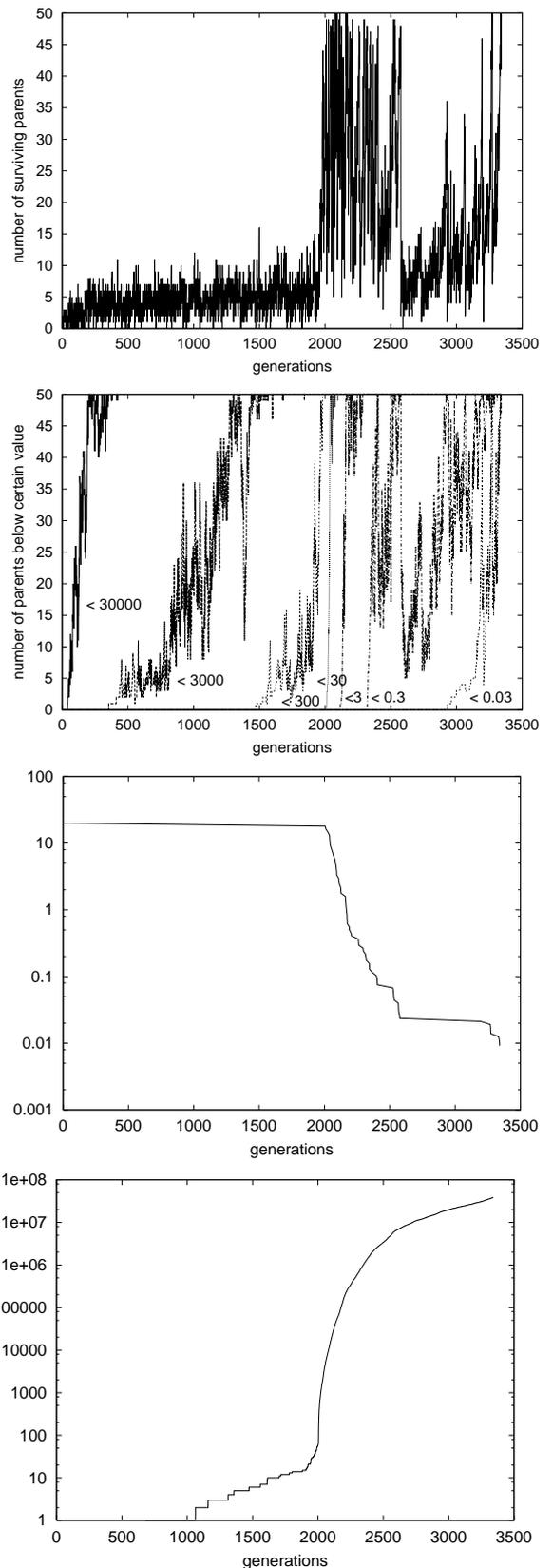


Figure 5: From top to bottom: Number of surviving parents, number of parents below certain fitness thresholds, decay of width threshold and cumulated number of re-evaluations.