

# Stochastic Urn Models of Innovation and Search Dynamics

Werner Ebeling<sup>a</sup> Lutz Molgedey<sup>a</sup> Axel Reimann<sup>a</sup>

<sup>a</sup>*Humboldt-University Berlin, Institute of Physics, D-10115 Berlin, Germany*

---

## Abstract

This work is devoted to applications of the Ehrenfest urn model to innovation and search processes. In the first part we discuss systems of two urns serving as models of innovation processes. The elementary act of innovation is considered as a transition from old (technologies, way of production, behavior, decisions) to new. The survival probability of the new under the influence of stochastic effects is discussed. In the second part we study systems of  $s \gg 1$  urns serving as models for optimal solution searching in optimization problems. The problem is to find the minimum on a large set of real numbers  $U_i$  using a total of  $N$  seekers ( $N \simeq 2 - 100$ ) simultaneously. The potential  $U_i$  is defined on the integer set  $i = 1 \dots s$ , where  $s$  is extremely large. In particular we consider frustrated periodic strings and the merit problem. The known equations for thermodynamic search processes and for simple models of biological evolution are unified by defining a two-parameter family of equations which embeds both cases. The search parameters are controlled by means of seeker ensemble dispersion.

*Key words:* urn model, innovation, search process, variability, parameter control

*PACS:* 02.50.Ey, 05.10.-a, 05.10.Gg

---

## 1 Introduction

Our world appears to be the result of a special search process: natural evolution. There is no external program which controls this search for well-adapted solutions. Evolution is fundamentally based on the self-organization of matter which is a highly complex nonlinear process [1,2] As in earlier works, it is understood here as unlimited sequences of self-organization steps [2,3]. One of evolution's most evident features is the tendency to form well-adapted complex structures and behaviors in course of time. The understanding and explaining of the tendency to more complexity and optimization forms the heart of any research program dealing with problems of evolution. Modern research has

shown that the strategies developed in evolutionary processes might also be of interest for the design and construction of technical systems. Pioneering work in this direction has been done by BREMERMAN, HOLLAND, KIRKPATRICK, RECHENBERG, SCHWEFEL and others [4–9]. The mathematical model used in this work is basically a system of Ehrenfest’s urns with stochastic transitions [2,3,10,11]. This work is mainly methodologically oriented. We will show that the combination of different strategy types (thermodynamically as well as biologically oriented) may lead to a more effective search. In order to retain the possibility of extensive comparisons we will only treat relatively simple optimization problems with about  $2^{30}$  configurations here.

## 2 Ehrenfest’s Urn model: Dynamics on the occupation number space

At the beginning of the last century the Ehrenfests investigated an urn schema which is ideally suited to model the class of problems we have in mind here. Let us assume we have  $s$  urns filled with  $N_1, N_2, \dots, N_s$  balls. We now define a stochastic game where at random times balls are taken out of one urn to be transferred to another urn.

We restrict our study to games with binary exchange where the transition rate  $W$  of one ball from urn  $j$  to urn  $i$  catalyzed by  $k$  is given by

$$\begin{aligned} W(N_1 \dots N_i + 1 \dots N_j - 1 \dots N_s | N_1 \dots N_i \dots N_j \dots N_s) \\ = A_{ij} N_j + B_{ij} N_i N_j + \sum_k C_{ijk} N_i N_j N_k \end{aligned} \quad (1)$$

We note that our games conserve the total particle number  $N = \sum_{i=1}^s N_i$ . The game defined by the simple transition rule (1) is indeed surprisingly rich and motivated as follows:

- In the deterministic limit the transition rule includes many of the known evolution models as e.g. the Fisher-Eigen-Schuster dynamics and the Lotka-Volterra dynamics in a generalized form.
- The transition rule (1) can be simulated by very simple algorithms – only including the elements *selection* of a member, a pair or a triple, *comparisons* and *transitions* to other urns.

We can show easily that the mean values of the occupation numbers in the thermodynamic limit approximately follow the differential equation:

$$\frac{d}{dt} x_i = \sum_j \left( A_{ij} x_j + B_{ij} N x_i x_j + \sum_k C_{ijk} N^2 x_i x_j x_k \right) \quad (2)$$

with

$$x_i = \frac{\langle N_i \rangle}{N} \quad (3)$$

as the mean relative occupation number. This class of differential equations includes the Lotka-Volterra schema, the Eigen-Schuster schema and many others [2,3,12]. This shows us that the urn schema with binary exchange may be well suited to model simple evolutionary processes. The stochastic dynamics described by these rules does not – in general – admit analytical solutions. In the special case of two participants  $s = 2$  denoted by  $m = \text{master}$  and  $n = \text{newcomer}$ , an analytical solution for the survival probability of the innovative trait infecting the population in initially  $N_n$  exemplars is known [13,14].

For the transition from  $N_n$  to  $N_n + 1$  we assume the rates

$$A_{nm} = 0; \quad B_{nm} = \frac{E_n}{N}; \quad C_{nmk} = \frac{B_n}{N} \delta_{nk}; \quad k = \{n, m\} \quad (4)$$

which leads to

$$W(N_n + 1, N_m - 1 | N_n N_m) = E_n \frac{N_n N_m}{N} + B_n \frac{N_n^2 N_m}{N} \quad (5)$$

The probability  $\sigma_{N_n, N}$  that a population of  $N_n$  newcomers with parameters  $E_n, B_n, N$  infecting a population of  $N_m = N - N_n$  masters with parameters  $E_m, B_m$  survives in the competition is given by [1]:

$$\sigma_{N_n, N} = \left[ 1 + \sum_{j=1}^{N_n-1} \prod_{i=1}^j \frac{E_m + B_m(N-i)}{E_n + B_n i} \right] \cdot \left[ 1 + \sum_{j=1}^{N-1} \prod_{i=1}^j \frac{E_m + B_m(N-i)}{E_n + B_n i} \right]^{-1} \quad (6)$$

Equation (6) was derived in a biological context in [14] at first, and was so far applied to problems of technological innovations [1] only. We see that the newcomers' survival probability is a quite complicated function of  $N_n, N, E_n, E_m, B_n$  and  $B_m$  which is smooth however, and in tendency increasing with  $N_n, (E_n/E_m)$  and  $(B_n/B_m)$ . Applications of this expression to problems in economical, technological and scientific evolution were given in [1]. Here, we propose a new application to group decision making, a problem which recently attracted some interest [15]. Let us consider a group (a committee) of  $N$  people searching for a decision between "yes" or "no".

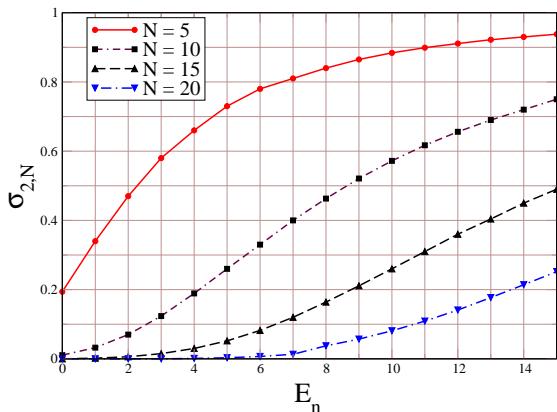


Fig. 1. Winning probability for 2 innovators in a committee consisting of  $N = 5, 10, 15, 20$  members. Parameters:  $E_m = 1, B_m \equiv B_n = 1$

$N_y(0) \gg N_n(0)$  is given. We see that the members voting for "yes" play the master's role in the game whereas the initially small group voting for "no" represents the innovators. After a long time period, the group as a whole votes for either "yes" or "no" – i.e. "yes" and "no" are absorbing states. However, the decision itself is stochastic with probability  $\sigma_{N_n(0),N}$  for "no" and  $1 - \sigma_{N_n(0),N} = \sigma_{N_y(0),N}$  for "yes". As a rule in our game the members voting for "yes" will most likely win since they start with the larger initial number. However, there is also a small probability that the voters for "no" make the game due to the fact that they have more / better arguments ( $E_n > E_y$ ). This probability is given by eq. (6) and represented in Fig. 1.

### 3 Stochastic modeling of mixed strategies

Now, let us study optimization problems. We assume that a set of real numbers  $U_i$  with  $i = 1, \dots, s$  is given ( $U_i$  - potentials). The problem is to find the minimal value  $U_{min}$ . Simulations and realistic search problems are mostly based on a finite number of seekers. Therefore, we transform the problem onto an urn model with  $s \gg 1$  urns and  $N \ll s$  balls. We introduce a set of integer numbers  $N_i(t)$  forming a lattice: the occupation number space. Any changes are restricted to one-step processes [2,3]

$$N_i \longrightarrow N_i \pm 1 \quad (7)$$

For the general case of mixed strategies we assume that the transition probabilities are given by eq. (1) including the following functions:

We model the process of finding the decision by a game with the two urns "y" and "n". The  $N$  participants consist at time  $t = 0$  of  $N_n(0)$  voters with the starting position "no", their position is described by positive real numbers  $E_n$  describing the strength of rational arguments to say "no" and  $B_n$ , describing the strength of a tendency for imitation of other group members. The corresponding figures for "yes" - voters are  $N_y(0), E_y, B_y$ . We start out with a situation where for some reason

$$A_{ij} = A_{ij}^0 G(U_j - U_i); \quad B_{ij} = F(U_j - U_i); \quad C_{ijk} \equiv 0 \quad (8)$$

The transition algorithm (1, 8) conserves the total number of seekers. The probability to find  $N_1$  seekers using the strategy 1 and  $N_i$  seekers using the strategy  $i$  at time  $t$  in the game:  $P(N_1 \dots N_i \dots N_s, t)$  satisfies a master equation [1-3,10,11]

$$\frac{\partial}{\partial t} P(N_1 \dots N_i \dots N_s, t) = \mathbf{W}P(N_1 \dots N_i \dots N_s, t) \quad (9)$$

where  $\mathbf{W}$  is the matrix operator of the transitions.

In the simplest case of a linear competition function we assume [2]:

$$F(\Delta U) = \gamma(\text{const} - \Delta U) \quad (10)$$

Another case of interest is hard competition – which is described by a step function like competition:

$$F(\Delta U) = \gamma(1 - \Theta(\Delta U)) = \gamma \Theta(-\Delta U) \quad (11)$$

In this case any improvement ( $\Delta U < 0$ ) is accepted, whereas all deteriorations ( $\Delta U \geq 0$ ) are rejected. Usually, the mutation is viewed as an undirected operation, whereas we bias the mutation operator in gradient direction. Therefore, the mutation operator is a product of a symmetric matrix  $A_{ij}^0 = A_{ji}^0$  and a non-increasing function of the potential difference  $G(U_j - U_i)$ . The special case  $G \equiv 1$  leads to a so called Darwin strategy. In the general case of mixed search strategies the mean value dynamics is given by eq. (2) with  $C_{ijk} \equiv 0$ . Here, we consider several special cases. In order to model Boltzmann type strategies, we introduce  $B_{ij} \equiv 0$  and unsymmetric temperature dependent transition matrices [16,17]. The simplest model of a Boltzmann process which tends to find minima of  $U_i$  – known as the Metropolis algorithm – is described by choosing

$$G(\Delta U) = \begin{cases} 1 & \text{if } \Delta U < 0 \\ \exp[-\beta \Delta U] & \text{if } \Delta U \geq 0 \end{cases} \quad (12)$$

In other words, a downhill transition is always carried out whereas an uphill transition occurs only with a small probability decreasing exponentially with the threshold's height. Now, we consider the dynamical equations for the mean values (2) which might be derived from the master equation (9). For the pure Boltzmann strategy the equations for the means read

$$\frac{d}{dt} x_i(t) = \sum_{j=1}^s (A_{ij}(\beta) x_j(t) - A_{ji}(\beta) x_i(t)) \quad (13)$$

In order to model mixed strategies we combine eqs. (10) and (12). Then, our model of a mixed strategy with the property to find minima of  $U_i$  leads to the following mean value eqs.:

$$\frac{d}{dt}x_i(t) = \gamma (\langle U \rangle - U_i) x_i(t) + \sum_{j=1}^s \left( A_{ij}(\beta) x_j(t) - A_{ji}(\beta) x_i(t) \right) \quad (14)$$

The new two-parameter family given by eq. (14) contains the Boltzmann strategy for  $\gamma = 0$  as a special case. The case of a pure Darwin strategy is obtained for

$$\gamma = 1, \beta \rightarrow 0. \quad (15)$$

The properties of eq. (14) in the limit  $\gamma = 0$  are well known. The equation has a Lyapunov functional (the Kullback entropy) which disappears for the Boltzmann distribution. In the other limit  $A_{ij}(\beta) \equiv 0$  the solution observes the relation

$$\frac{d}{dt} \langle U \rangle \leq 0 \quad (16)$$

while the target is the occupation of the minimal  $U_i$ . In the general case  $\gamma \neq 0$  and unsymmetric matrices  $A_{ij}(\beta) \neq A_{ji}(\beta)$  no general theorem is known yet, but periodic orbits are not likely to exist.

Finally, we may combine the step function eq. (11) with directed transitions which leads to the so called tournament strategies. Then, the mean values satisfy in average the following nonlinear equation:

$$\begin{aligned} \frac{d}{dt}x_i(t) = & \gamma \sum_{j=1}^s \left( \Theta(U_j - U_i) - \Theta(U_i - U_j) \right) x_i(t)x_j(t) \\ & + \sum_{j=1}^s \left( A_{ij}(\beta) x_j(t) - A_{ji}(\beta) x_i(t) \right) \end{aligned} \quad (17)$$

#### 4 Simulations – the dependence on search parameters

In this section we will test the effectiveness of the stochastic strategy versions introduced above by applying them to very simple model problems. In particular we will study the dependence of the search on the total number of seekers, temperature and mutation/selection rates.

The maximal computation for all simulations was fixed to a certain value  $t_{max}$  while the computation time  $t$  itself was simulated according to Gillespie [18,19].

In earlier works we considered the application of mixed Boltzmann-Darwin-strategies based on the discrete form of mixed strategies given by equations (14 - 17) to the traveling salesman problem (TSP) and to related street network problems [17,20]. It could be shown by simulations that in this respect mixed Boltzmann-Darwin-strategies have good search properties.

Let us consider the simple frustrated sequence game proposed by Engel: a sequence is constructed using 4 letters, e.g.  $\{A, B, C, D\}$ . This sequence gets credited a fitness point if any two letters appear in alphabetical order. (We also consider the wrapped order  $\{D, A\}$  to be alphabetical.) Additionally, the recurrence of a letter with period 5 is credited by one point. We simulated sequences of length 15. The maximum possible fitness in this configuration is 22 – acquired by an exhaustive search. The ratio of local to global optima is 24 : 4 or 6 : 1. As a second – more difficult – problem we investigated so called MERIT sequences (or Low Autocorrelation Binary Sequences; LABS) [21–23]. The optimization problem for a given sequence

$$S = S_1, S_2, \dots, S_L, \quad S_i = \pm 1 \quad (18)$$

is to find the realization with least autocorrelation or maximal Merit factor. The Merit factor  $F$ , introduced by GOLAY [22,24], serves as fitness measure:

$$F = \frac{L^2}{2E}; \quad \text{with: } E = \sum_{k=1}^{L-1} R_k^2 \quad \text{and} \quad R_k = \sum_{i=1}^{L-k} S_i S_{i+k} \quad (19)$$

In this work, we were using sequences of length  $L = 29$ . The best sequences found through exhaustive search have a fitness (Merit factor) of  $F = 6.78226$ . For our computer simulations the number of representatives (seekers)  $N$  in the ensemble has been varied between 1 and 100. Let us underline that  $N = 1$  always corresponds to a pure Boltzmann strategy and  $N > 1$  to a mixed Boltzmann-Darwin strategy. In this case we are simulating a population of  $N$  seekers searching simultaneously and being coupled by competition for best results. We always calculated the mean values and the best results obtained for a fixed maximum number of fitness function evaluations. The results of the *Boltzmann strategy* applied to Engel sequences are displayed in Fig. 2.

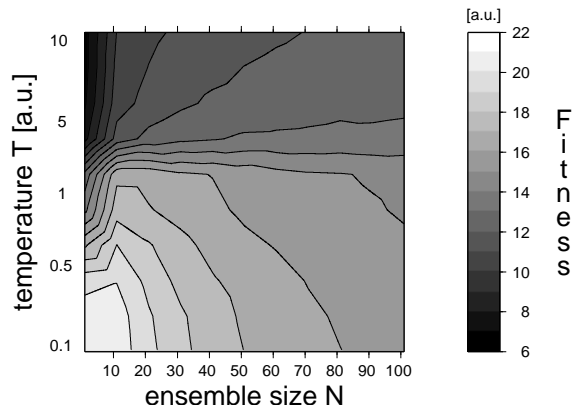


Fig. 2. Boltzmann strategy: Fitness of best seeker; 1000 run average; Engel sequence length  $L = 15$ ;  $t_{max} = 1000$

The temperature parameter was varied between  $T = 0.1$  and  $T = 10$ . The according acceptance probabilities of bad mutations vary between  $\exp[-1/0.1] \approx 45 \cdot 10^{-6}$  and  $\exp[-1/10] \approx 0.9$ .

As expected, the Boltzmann strategy gains no advantages in ensemble based search due to missing coupling between the seekers except that the best seeker of an ensemble has a smaller probability to stick in local optima. The fact that best results are achieved for lower temperatures indicates that some 10 seekers suffice to keep the trapping probability in local optima small.

After all, the Boltzmann strategy succeeds finding the global optimum of 22 in the given time limit for the parameters  $N < 15$  and  $T < 0.4$ . Figure 3 shows simulation results for a *mixed Boltzmann-Darwin strategy* using a tournament selection involving 4 seekers respectively. In order to control the coupling (selection strength) the frequency of selection operations as compared to mutation steps was varied. The corresponding probability  $P_{mut}$  is plotted along the ordinate axes. A value of 0% means that mutation steps were never carried out. Accordingly, 50% means that *in average* every mutation step was followed by a selection step; whereas 100% indicates that there were mutation steps only. Visibly, the optimal number of seekers depends on the chosen mutation/selection ratio. The best results, however, are achieved with relatively small ensembles  $N \approx 20$ .

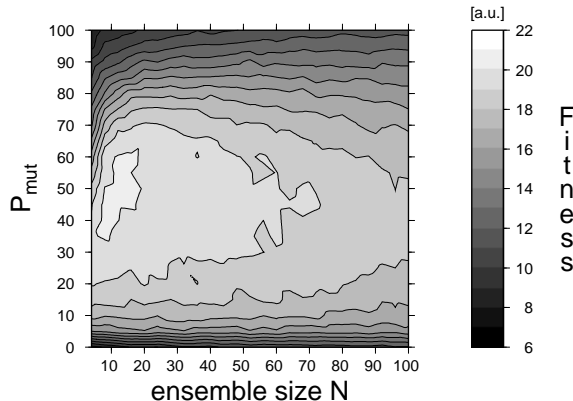


Fig. 3. Boltzmann-Darwin strategy: Fitness of best seeker, 1000 run average; 4 seekers tournament, Engel sequence length  $L = 15$ , temp.  $T = 1$ , time  $t_{max} = 1000$

## 5 Search parameter adaption by variability control

The main result of the simulations represented in the previous section is that the effectivity of a strategy critically depends on the search parameters. A good choice may easily improve the result by a factor 2. Here, we will develop adaptive strategies which bring the system to a favorable parameter region by a control based upon the ensemble's statistical parameters (like dispersion). The idea is to change the search parameters (e.g. temperature or mutation probability) in such a way that the seeker ensemble always has a good (not too big, not too small) dispersion on the fitness landscape. Simi-

lar approaches have been developed for  $(\mu, \lambda)$  strategies (cf. RECHENBERG's '1/5 success rule') [7,9]. In the case of Boltzmann search temperature adaptation was studied already by Andresen et. al. [25,26]. Our method is closely related to the cited method; however it is much simpler with respect to the implementation. We calculated the ensemble's fitness dispersion

$$\sigma^2(fitness) = \langle U^2 \rangle - \langle U \rangle^2 \quad (20)$$

and controlled the temperature according to

$$\dot{T} = -\frac{\nu T^2}{\sigma(fitness)} \quad (21)$$

where  $\nu$  is a small number compared to the inverse of the smallest potential change during a mutation:  $\nu < 1/|\Delta U_{min}|$ . The results obtained this way equal or even surpass those of a pure Boltzmann strategy. Figure 4 shows that for the Engel sequences tested only the best results of a Boltzmann search match those of the controlled annealing runs.

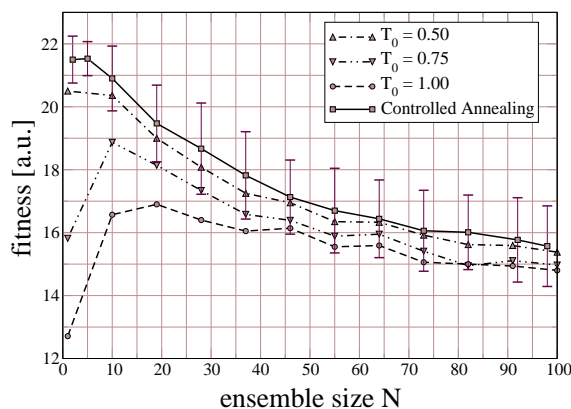


Fig. 4. Boltzmann strategy vs. controlled annealing; Engel sequence length  $L = 15$ , time  $t_{max} = 1000$

While the Boltzmann strategy basically depends on temperature settings and ensemble size only, the mixed strategies' effectiveness also crucially depends on the mutation vs. selection ratio as seen in Fig. 3. Very low mutation rates cause the strategy to resemble gradient search methods with higher likelihood of getting stuck. Too high mutation rates on the other hand basically turn the search into a random walk, destroying information already gathered while searching. Measuring the seekers' Hamming distance to the (nearest degenerated) global optimum and plotting the according histogram clearly demonstrates this behavior for very large times. Figure 5 shows a phase transition distinctly marking the point when the increasing mutation rate starts destroying information. This fact was already pointed out by EIGEN, SCHUSTER et al. [27,28]. An optimal mutation rate lies in the proximity of the phase transition, being just a little bit smaller – thus granting a quick search without information destruction. In order to automatically control the mutation / selection rates one only needs to know where the phase transition occurs. Disappointingly, one can only observe sharp transitions for *very long* simulation times, i.e. when the search approaches a stationary state.

Therefore one needs a different, truly dynamic indicator. An easily obtainable value is the seeker ensembles' variability, both in terms of fitness and genotype: In an ensemble of  $N$  seekers one can observe any number from 1 to  $N$  different seekers  $N_{diff}$ . Normalizing the variability one obtains an indicator  $v$  varying from  $1/N$  to 1:

$$v = \frac{N_{diff}}{N} \quad (22)$$

In extensive investigations, EIGEN, SCHUSTER [27] et. al have shown that for  $L \sim 50$  and in the limit  $L \rightarrow \infty$  the dynamics of eq. (14) shows for a similar optimization problem some sort of 'phase transition' at certain mutation rates. Furthermore, these authors have shown that operating near this phase transition enhances the search effectivity. Later it was shown that in the limit  $L \rightarrow \infty$  the observed transition corresponds to a second order phase transition [28]. Here, we will use these results in order to find optimal regimes for our search parameters. We have to take into account, however, that in finite systems (here:  $N \sim 100$ ,  $L \sim 15 - 30$ ) the expected phase transitions are rather smooth transitions. Nevertheless, we are still able to identify these transitions in our simulations. Figure 5 (solid line) shows that this new indicator indeed redraws the 'phase transition'. Measuring the variability for *optimal* mutation probabilities with respect to the ensemble size  $N$ , one observes a standard deviation like behavior for the optimal variability  $v_{opt}$ :

$$v_{opt} \simeq \frac{1}{\sqrt{N}} \quad (23)$$

Thus, for an ensemble consisting of e.g.  $N = 50$  seekers the variability control has to aim for an optimal ensemble variability:

$$v_{opt} \simeq \frac{1}{\sqrt{50}} = 0.14 \quad (24)$$

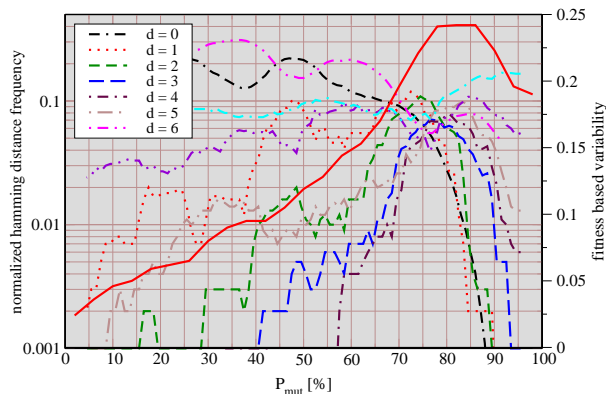


Fig. 5. Boltzmann-Darwin strategy: 4 seekers tournament, Engel sequence length  $L = 15$ ,  $N = 20$ , temp.  $T = 1$ ; Hamming distance distribution:  $H(\text{seeker}, \text{nearest global optimum})$ ; time  $t_{max} = 10^4$ ; right ordinate axis and solid line: ensemble fitness dispersion

corresponding to a mutation probability (reading Fig. 5) of  $P_{mut} \approx 55\%$ . This is indeed a good choice for the mutation rate as can be seen in Fig. 3 for  $N = 50$  and  $P_{mut} = 55\%$ . The fitness variability displayed in Fig. 5 shows that the adapted mutation rate is in fact close to the phase transition point, as initially demanded. A major drawback is, however, that the variability's variance  $\sigma_v^2$  also scales with the ensemble size

$$\sigma_v^2 = \mathcal{O}\left(\frac{1}{N}\right), \quad (25)$$

rendering it useless as control parameter for small ensembles  $N \leq 10$ . To circumvent this problem we can define some 'control free'  $\varepsilon \simeq \sigma_v^2$  around the optimum variability we are aiming for, thus softly blending in control with increasing ensemble size. The adaption actually implemented works as follows: The  $\varepsilon$  - interval around the optimal variability value was set to  $\varepsilon = 1/N$ . The variability's deviation

$$\Delta v = v - v_{opt} \quad (26)$$

from the optimum was measured in terms of  $\varepsilon$  to exponentially adapt the mutation probability  $\dot{P}_{mut} = \alpha P_{mut}$  with

$$\alpha = \begin{cases} \frac{-\Delta v}{\varepsilon}; & \text{if } |\Delta v| > \varepsilon \\ 0; & \text{otherwise} \end{cases} \quad (27)$$

Mutation probabilities outside the interval  $[0, 1]$  are clamped to the interval's boundaries. It takes this algorithm less than 30 time steps to reach and hold optimal parameter values – in contrast, up to  $\mathcal{O}(10^4)$  simulated time steps were necessary to unveil the phase transition showed in Fig. 5. Simulations using this auto-tuning demonstrate that the mutation rate adaption can effectively boost search results, as demonstrated in Fig 6. For optimal mutation rates between  $P_{mut} = 40\% \dots 50\%$  (cf. Fig. 3), search results coincide with the automatic rate adaption results in Fig. 6 (solid line). Now, let us

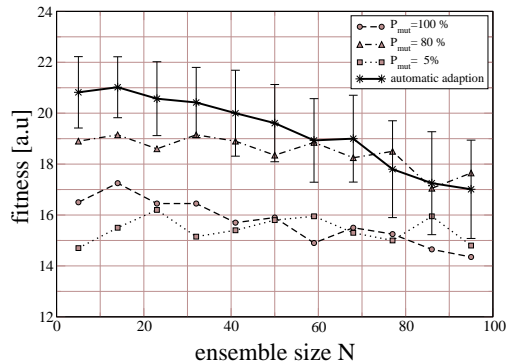


Fig. 6. Boltzmann-Darwin strategy: Fitness of best seeker, Used parameters: 4 seekers tournament, Engel sequence length  $L = 15$ , temperature  $T = 1$ , time  $t_{max} = 1000$ ;  $P_{mut} = 100\%$  represents a random search process; mutation rate adaption by variability control yields almost always best results

put our adaptive strategies to the test on a more difficult task: the Merit problem. As already mentioned in section 4, we were using sequences of length  $L = 29$  which have a maximum Merit factor of  $F = 6.78226$ . For this particular length it is advantageous to reduce the search to the subspace of skew symmetric sequences. Skew symmetric sequences are defined by the following equation for corresponding bits having equal distance  $i$  with regard to the string center labeled  $n$ :

$$S_{n+i} = (-1)^i S_{n-i}; \quad n = \frac{L+1}{2} \quad (28)$$

Figure 7 clearly shows the difference in the non-controlled search results using either normal or skew symmetric mutation operators. It also becomes clear that – in contrast to our Engel problem – the optimal temperature is above zero in the case of skew symmetric mutation.

For an optimal temperature of  $T \approx 0.28$ , the Boltzmann strategy yields a mean ensemble fitness of about 4. This is the corner stone to which our adaptive strategies have to compare in this case. The results of both the controlled annealing and the mixed Boltzmann-Darwin strategy with mutation rate adaption are plotted in Fig. 8. As ensemble based strategies they naturally show a distinct dependence on the ensemble size  $N$ . While the controlled annealing strategy yields best results for some  $N \approx 10$  seekers, the adaptive Boltzmann-Darwin strategy excels for ensembles sized  $N \approx 35$ . Besides that, both strategies do a fairly good job tuning their respective parameters (controlled annealing: temperature  $T$ , adaptive Boltzmann - Darwin strategy: mutation rate  $P_{mut}$ ).

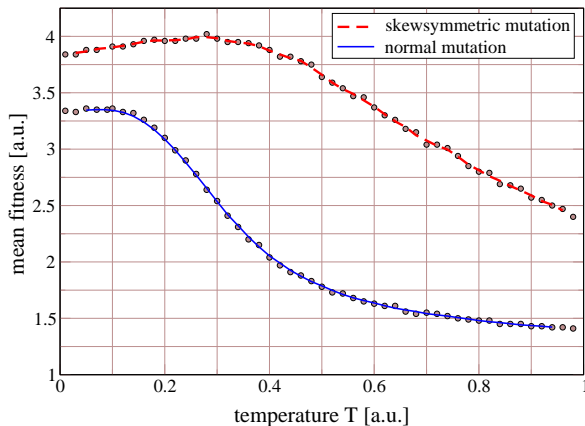


Fig. 7. Boltzmann strategy: mean ensemble fitness; Used parameters: Merit sequence length  $L = 29$ , time  $t_{max} = 10.000$ ; solid line: normal mutation operator; dashed line: skew symmetric mutation operator

The search results are well comparable to the best search results using a pure Boltzmann strategy. For too small ensemble sizes both strategies have a hard time to extract decent statistical measures. A good choice for the ensemble size is  $N > 30$  for both the controlled annealing and the adaptive Boltzmann - Darwin strategy. The latter displays a much higher mean fitness standard deviation ( $\sigma \approx 0.8$ ) than the controlled annealing strategy ( $\sigma \approx 0.2$ ) for all given ensemble sizes. In general, the adaptive strategies have proven to be successful even for this particular difficult problem.

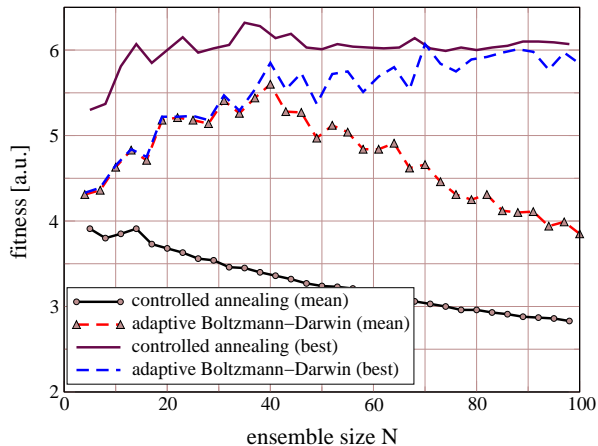


Fig. 8. Merit ensemble fitness; parameters: skew symmetric mutation operator, sequence length  $L = 29$ , time  $t_{max} = 1000$ ; solid line: controlled annealing, initial temp.  $T = 2$ ; dashed line: adaptive Boltzmann - Darwin strategy, tournament of 4 seekers;  $T = 0.3$

## 6 Discussion

In this work, the power of ensemble based optimization strategies was investigated. We have shown that Boltzmann-type search with ensemble based annealing control leads to very good results (cf. Fig 8). On the other hand, mixed strategies with appropriately set parameters do not always give the best results with respect to fitness mean and best fitness value. A certain amount of parallelism introduced by a Darwin term ( $\gamma > 0$ ) is a useful element in good search strategies. The computation time which is lost simulating a number of parallel seekers is gained by certain advantages of parallel search as e.g. the possibility of experience exchange. In our mixed strategy this is modeled by the Darwin elements: competition between seekers, survival of the fittest. However, when the number of seekers working in parallel is too high the parallelism costs more than one gains from it. Theoretical investigations regarding this subject have been carried out by BARKEMA, GOLDBERG, NAKANO et. al. It is easy to see that the computation overhead necessary for parallel strategies also cannot pay off if the fitness landscape is rather simple. According to our experience, a successful search (optimization) requires to choose the proper strategy and a fine balance between parallelism and individualism, that is between Darwin and Boltzmann elements. Similar results were obtained in another work analyzing street network problems [20]. The results summarized

here show a certain advantage of including Darwin elements into the search strategy – especially if these elements get tuned appropriately. As far as the tuning parameter is concerned, it is still an open question whether the variability's optimum value of order  $1/\sqrt{N}$  is truly model independent or has to be adapted to the particular problem.

Possibly, the real power of mixed strategies including Darwin elements will show up on parallel computers with more than some 4 processors. In this work, simultaneous search was always simulated on a single sequential computer. Of course, it could in principle also be carried out on a net of parallel processors. Since coupling (selection process) between the elements of a Darwin ensemble is a rather seldom event, the speed up by using such an  $N$ -processor system is expected to be higher than of order  $\mathcal{O}(\log N)$ , thus scaling better than simulated annealing techniques.

## References

- [1] E. Bruckner, W. Ebeling, M. A. Jimenez-Montano, and A. Scharnhorst. Nonlinear stochastic effects of substitution. *J. Evol. Econ.*, 6:1–30, 1996.
- [2] W. Ebeling, A. Engel, and R. Feistel. *Physik der Evolutionsprozesse*. Akademie-Verlag Berlin, 1990.
- [3] R. Feistel and W. Ebeling. *Evolution of Complex Systems*. Kluwer Publ. Dordrecht, 1989.
- [4] H. J. Bremermann. A method of unconstrained global optimization. *Math. Biosci.*, 9:1–15, 1970.
- [5] J. H. Holland. *Adaption in natural and artificial systems*. MIT Press, Cambridge MA, 2 edition, 1975.
- [6] S. Kirkpatrick, C. D. Gelatt, and M. P. Vecchi. Optimization by simulated annealing. *Science*, 220:671–680, 1983.
- [7] L. Rechenberg. *Evolutionsstrategie 94. Optimierung technischer Systeme nach Prinzipien der biologischen Evolution*. Fromman-Verlag, Stuttgart, 1994.
- [8] Hans P. Schwefel. *Numerical Optimization of Computer Models*. Wiley, New York, 1981.
- [9] Hans P. Schwefel. *Evolution and Optimum Seeking*. Wiley, New York, 1994.
- [10] D. Helbing. *Quantitative sociodynamics. Stochastic methods and models of social interaction processes*. Kluwer Academic, Dordrecht, 1995.
- [11] W. Weidlich and G. Haag. *Concepts and Models of quantitative sociology*. Springer, Berlin, 1983.
- [12] Manfred Eigen and Peter Schuster. The hypercycle. *Naturwiss.*, 64, 65:541, 341, 1977, 1978.
- [13] P. Allen and W. Ebeling. Evolution and the stochastic description of simple ecosystems. *Biosystems*, 16:113–126, 1983.
- [14] W. Ebeling, I. Sonntag, and L. Schimansky-Geier. On the evolution of biological macromolecules. *Studia biophysica*, 84:87–97, 1981.
- [15] S. Galam. A random field model for rational group decision making. contribution to this Heraeus-Seminar, March 2000.
- [16] T. Boseniuk and W. Ebeling. Optimization of np-complete problems. *Europhys. Lett.*, 6:107, 1988.
- [17] T. Boseniuk, W. Ebeling, and A. Engel. Boltzmann and darwin strategies in complex optimization. *Phys. Lett.*, 125:307, 1987.
- [18] D. T. Gillespie. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. *J. Comp. Phys.*, 22:403–434, 1976.

- [19] R. Feistel and W. Ebeling. Deterministic and stochastic theory of sustained oscillations in autocatalytic reaction systems. *Physica*, 43:114–124, 1978.
- [20] F. Schweitzer, W. Ebeling, H. Rosé, and O. Weiss. Optimization of road networks using evolutionary strategies. *Evol. Computation*, 5:419–438, 1998.
- [21] J. Bernasconi. Low autocorrelated binary sequences: statistical mechanics and configuration space analysis. *J. Physique*, 48:559–567, April 1987.
- [22] M. J. E. Golay. A class of finite binary sequences with alternative autocorrelation values equal to zero. *IEEE Trans. on Information Theory*, 18:449–450, May 1972.
- [23] Burkhard Militzer, Michele Zamparelli, and Dieter Beule. Evolutionary search for low autocorrelated binary sequences. *IEEE Trans. Evol. Comput.*, 2:34–39, 1998.
- [24] M. J. E. Golay. Sieves for low autocorrelation binary sequences. *IEEE Trans. on Information Theory*, 23:43–51, January 1977.
- [25] B. Andresen and J. M. Gordon. Constant thermodynamic speed for minimizing entropy production thermodynamic processes and simulated annealing. *Physical Review E*, 6:4346–4351, 1994.
- [26] George Ruppeiner, Jacob Mørch Pedersen, and Peter Salamon. Ensemble approach to simulated annealing. *J. Phys. I*, 1:455–470, 1991.
- [27] Manfred Eigen, John McCaskill, and Peter Schuster. Molecular quasi-species. *J. Phys. Chem*, 92:6881–6891, 1988.
- [28] Ira Leuthäusser. Statistical mechanics of eigen’s evolution model. *Journal of Statistical Physics*, 48(1/2):343–360, 1987.