Fitness Uniform Selection to Preserve Genetic Diversity

Marcus Hutter

Istituto Dalle Molle di Studi sull'Intelligenza Artificiale
IDSIA, Galleria 2, CH-6928 Manno-Lugano, Switzerland
marcus@idsia.ch, http://www.idsia.ch/~marcus

ECCO XV, 30 May 2002

Contents

Evolutionary algorithms (EA)

- EAs are capable of solving complicated optimization tasks in which an objective/fitness function $f : I \to \mathbb{R}$ shall be maximized.
- $i \in I$ is an individual from the set I of feasible solutions.
- A population $P \subseteq I$ of individuals is maintained.
- In steady state EAs P is updated by selecting (and possibly deleting) a few individuals from the current population.
- The individuals are mutated and recombined and added to the population.
- We are interested in finding a single individual of maximal fitness f for difficult multimodal and deceptive problems.

Problem: Local Optima & Selection Intensity

Standard selection schemes (STD)

- Proportionate [Holland:75],
- Tournament [Baker:85].
- Ranking [Whitley:89],
- Truncation [Muehlenbein&Schlierkamp-Voosen:94],

Further selection schemes:

- Boltzmann [Maza&Tidor:93]

The right selection pressure is critical in ensuring sufficient optimization progress on the one hand ...

and in preserving genetic diversity to be able to escape from local optima on the other hand.

Solutions: Local Optima & Selection Intensity

Dynamically determine and adapt the selection pressure parameters. [Eshelman:91,Baeck:91,Herdy:92,Schlierkamp-Voosen&Muehlenbein:94,...]

Preserve genetic diversity by

- crowding [DeJong&:75]
- fitness sharing [Goldberg&Richardson:87]
- local mating [Collins&Jefferson:91]

Crucial: Proper design of a neighborhood function based on the specific problem structure and/or coding.

In the following we suggest a new selection scheme, which automatically generates a suitably adapting selection pressure and which does not need special problem insight.

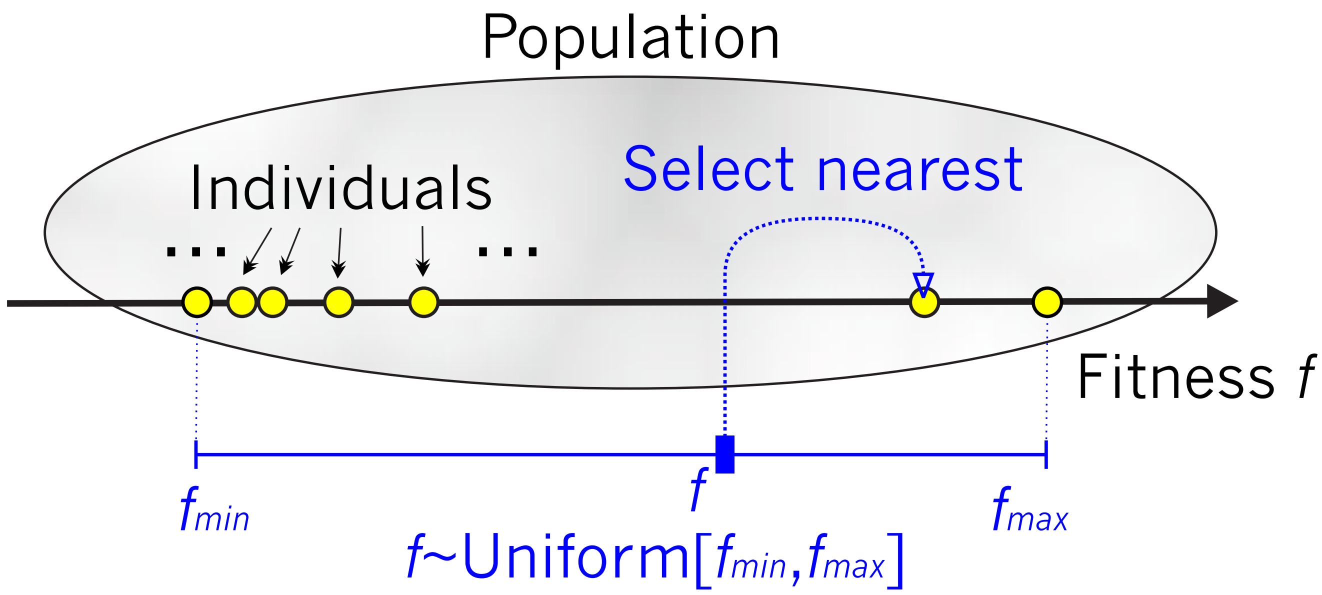
The Fitness uniform selection scheme (FUSS)

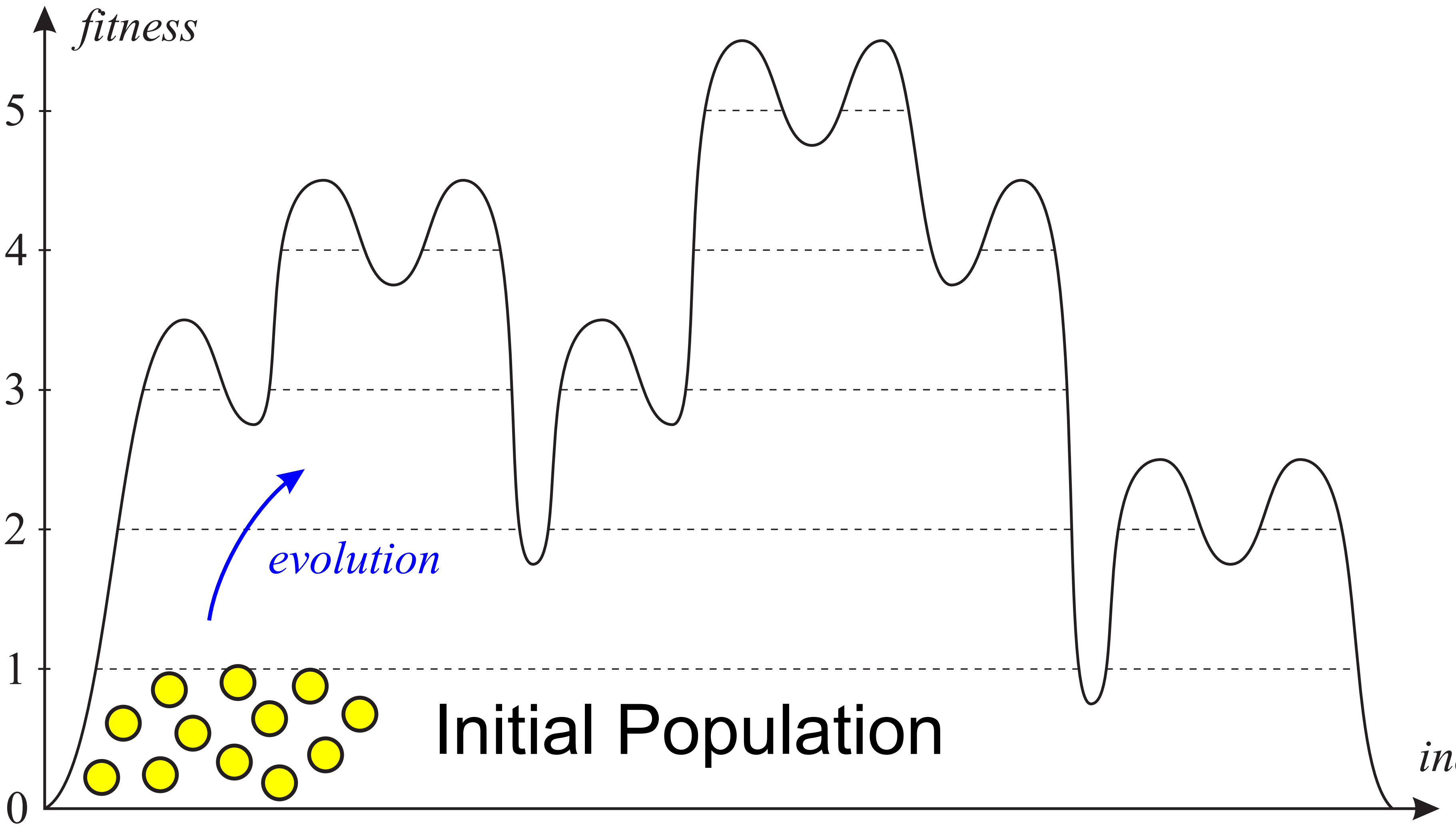
If the lowest/highest fitness values in the current population P are $f_{min/max}$,

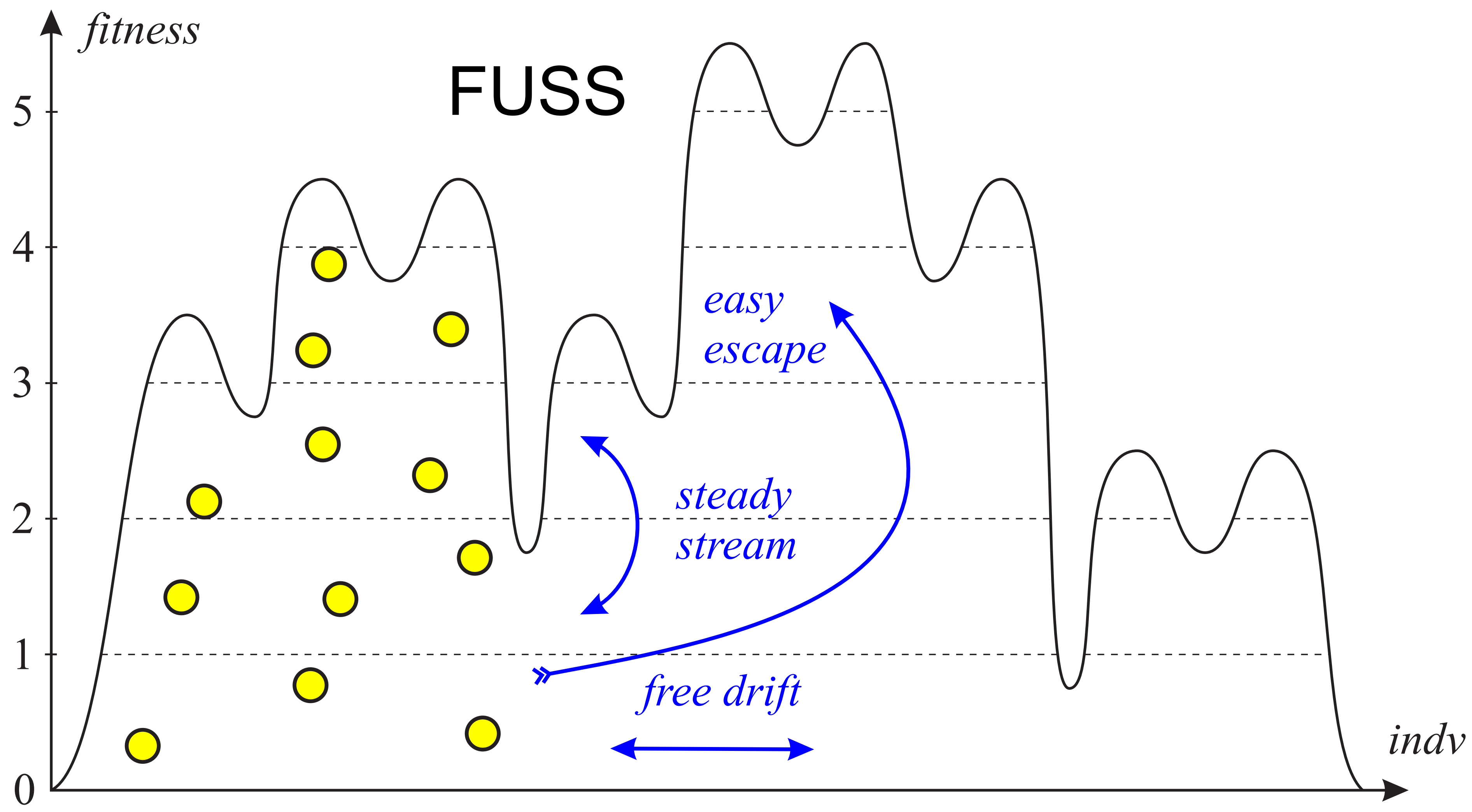
we select a fitness value f uniformly in the interval $[f_{min}, f_{max}]$,

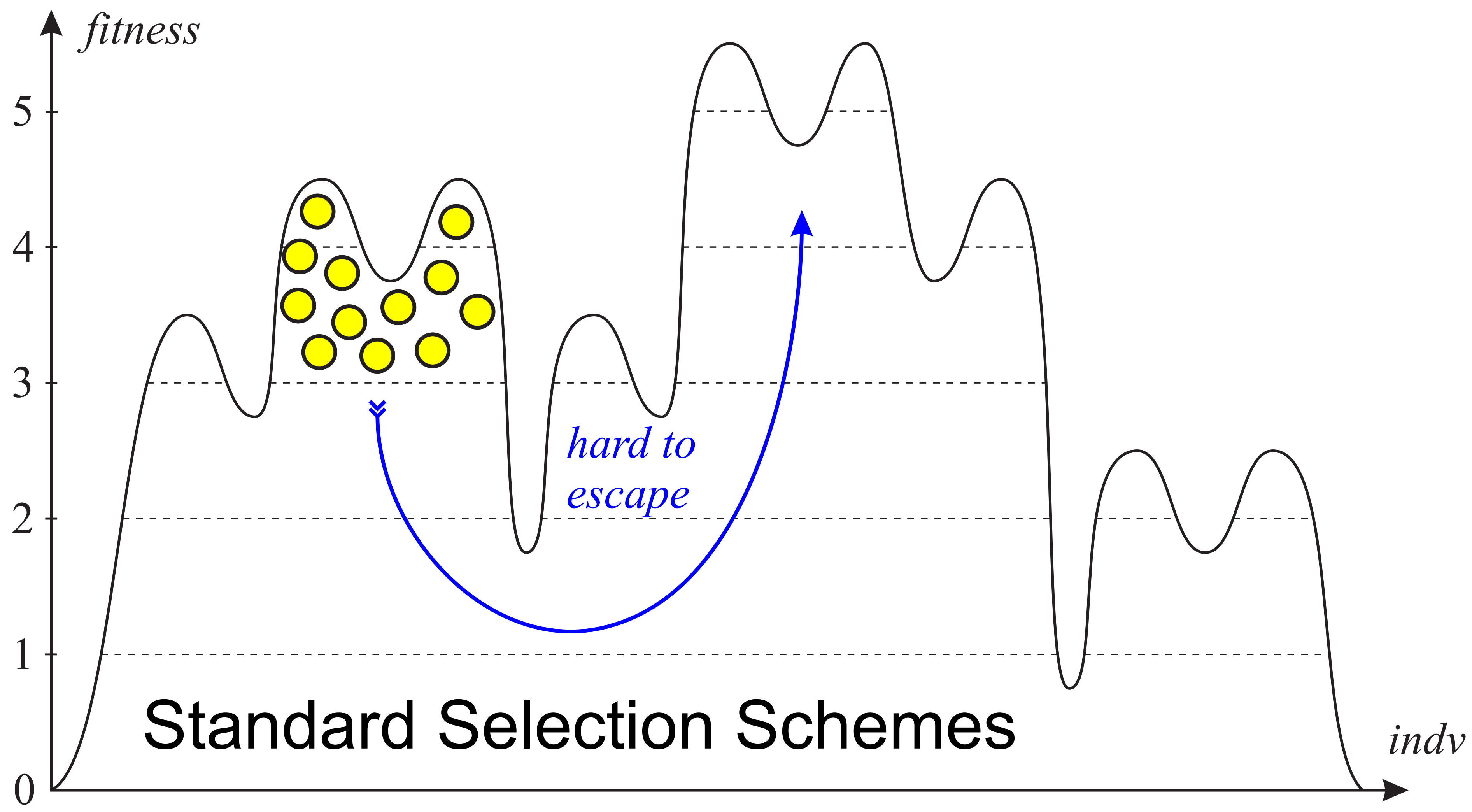
then, the individual $i \in P$ with fitness nearest to f is selected

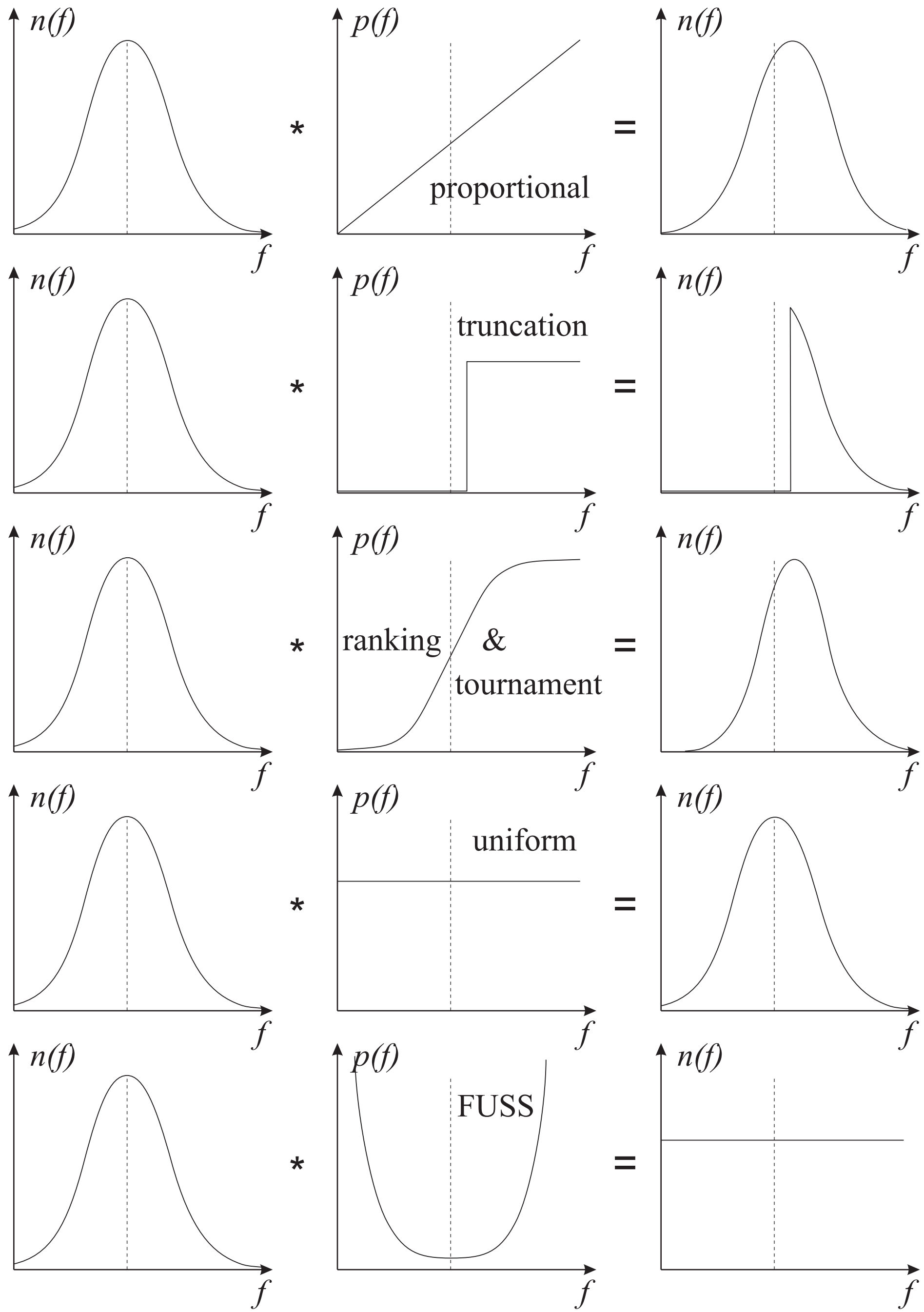
and a copy is added to P, possibly after mutation and recombination.




fitness
5
4
3
2
1
0
evolution
Initial Population
indv


fitness
FUSS
5
4
3
2
1
0
easy
escape
steady
stream
free drift
indv


fitness
5
4
3
2
1
0
indv
hard to
escape
Standard Selection Schemes



tournament ≡ ranking
[First proof in
Hutter, 1991]

Selection Pressure in FUSS

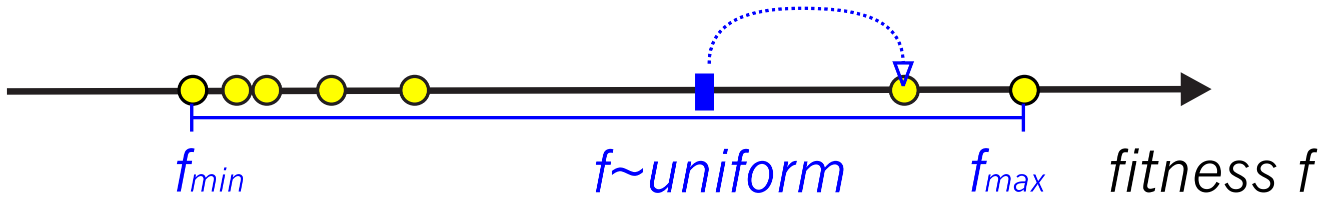
A probability distribution over the fitness values is used, unlike STD, which all use a distribution over individuals.

Premature convergence is avoided in FUSS by abandoning convergence at all (No takeover of fittest individual).

Nevertheless there is a selection pressure in FUSS towards higher fitness:

The probability of selecting a specific individual is proportional to the distance to its nearest fitness neighbor.

In a population with a high density of unfit and low density of fit individuals, the fitter ones are effectively favored.



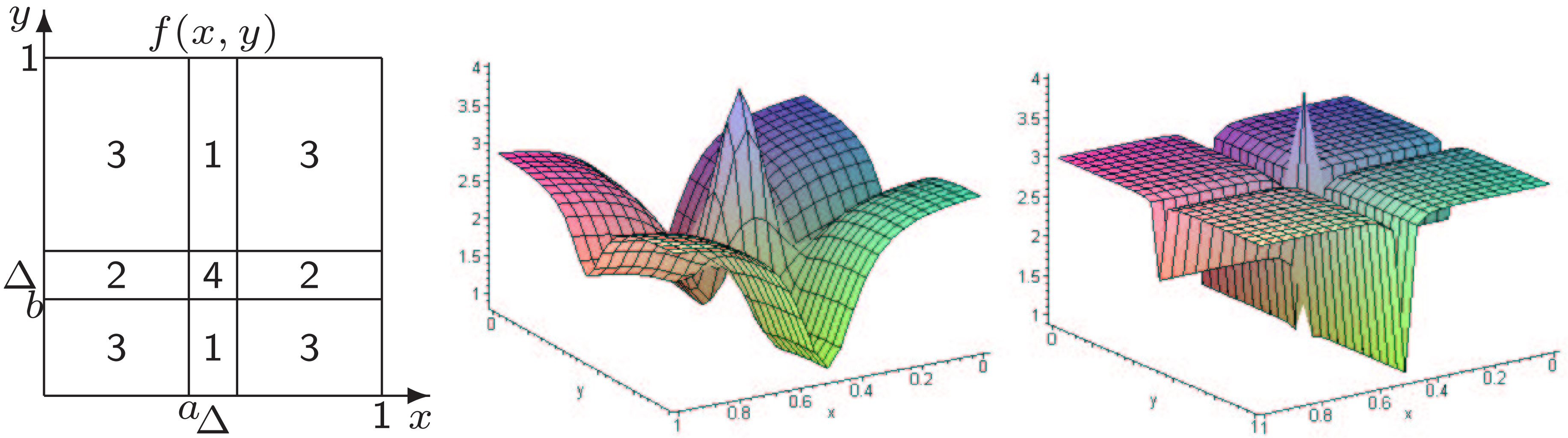
Dynamics and Equilibrium Distribution

- Within a fitness level the individuals freely drift around (by mutation).
- There is a steady stream of individuals into and out of a fitness level by (d)evolution from (higher)lower levels.
- All fitness levels remain occupied from which new mutants are steadily created, occasionally one leading to further evolution in a more promising direction.
- The equilibrium distribution is uniform in the fitness values, i.e. the number of individuals in a fitness interval $[f, f + \Delta f]$ is independent f.

Transformation properties of FUSS

- FUSS is independent of a scaling and a shift of the fitness function: FUSS($\tilde{f}$) with $\tilde{f}(i) := a \cdot f(i) + b$ is identical to FUSS(f).
- This is true even for $a < 0$, since FUSS searches for maxima and minima.
- FUSS is not independent of a non-linear (monotone) transformation unlike tournament, ranking and truncation selection.
- The non-linear transformation properties are more like the ones of proportionate selection.

Simple Deceptive Multimodal 2D Examples

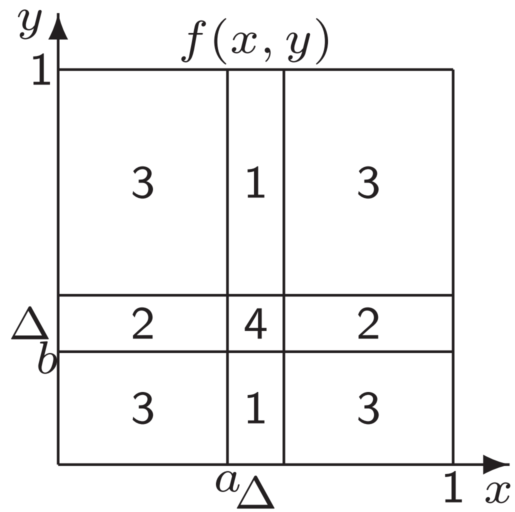


- The examples model individuals possessing up to 2 “features”.
- Possessing both features is better than none is better than one.
- Sort of an XOR structure, which is hard for most optimizers.

Analytical Results for a Similar 3D Example

Average search times for the global optimum ($\Delta \ll 1$):

Random search:	T_{RAND}	$\sim \Delta^{-3}$
Standard selection:	T_{STD}	$\sim \Delta^{-3}$
FUSS (no crossover):	T_{FUSS}	$\sim \Delta^{-2}$
FUSS (with crossover):	T_{FUSSX}	$\sim \Delta^{-1}$



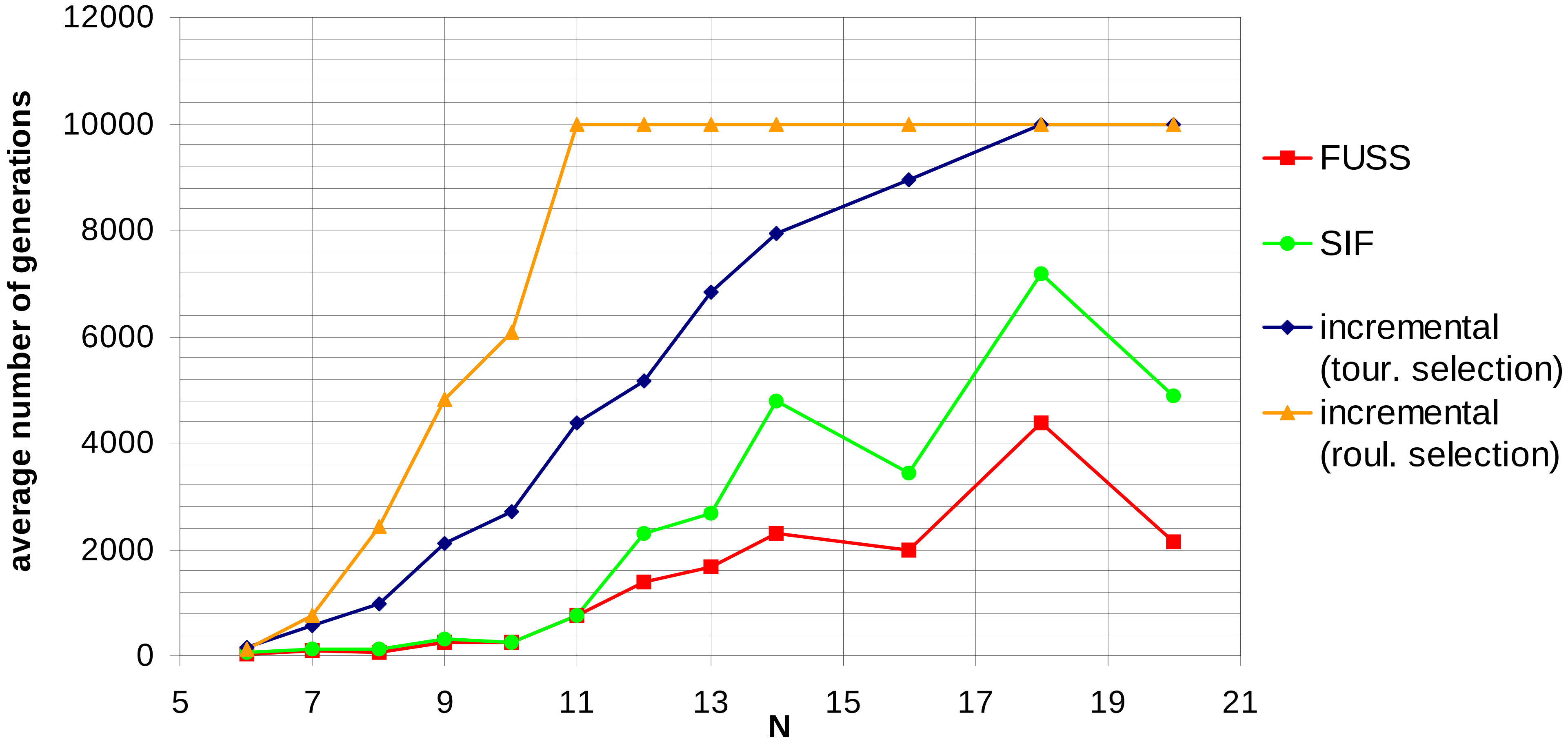
$\Longrightarrow$ FUSS can be faster than RAND and STD.

$\Longrightarrow$ Crossover can give a further boost in FUSS, even when ineffective in combination with STD.

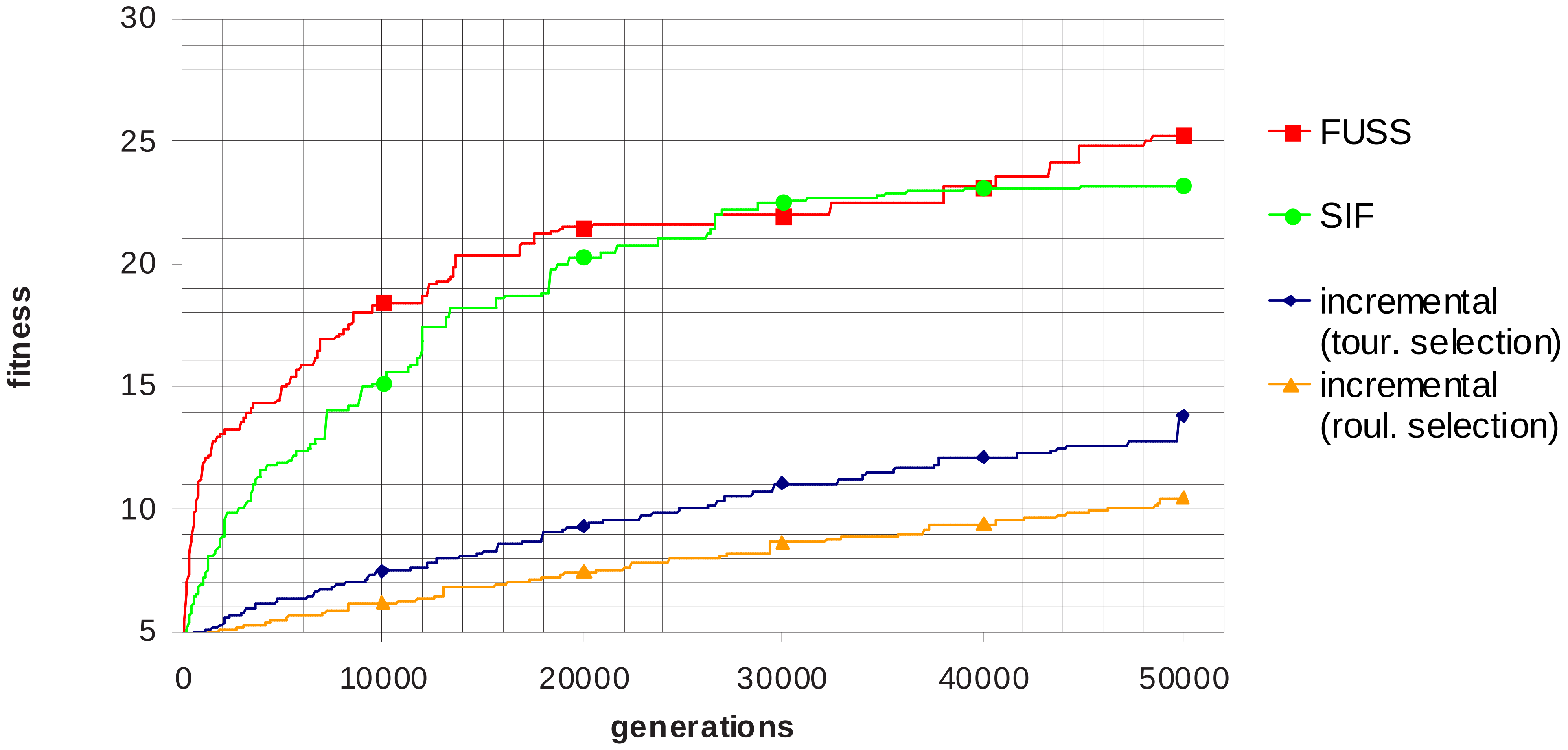
Traveling Salesman Problem

- We considered $10^{1..3}$ cities with random matrix distances and random 2D Euclidian distances,
- random initial paths,
- random 1-Opt and 2.5-Opt mutation operators,
- inverse path length as fitness,
- The solutions found by FUSS are consistently and significantly better than those found by STD (in the range of 20-50% given same number of selections and comparable parameter settings).
- The current implementation can in no way compete with up-to-date TSP-solvers, but this was not the intention of the comparison.

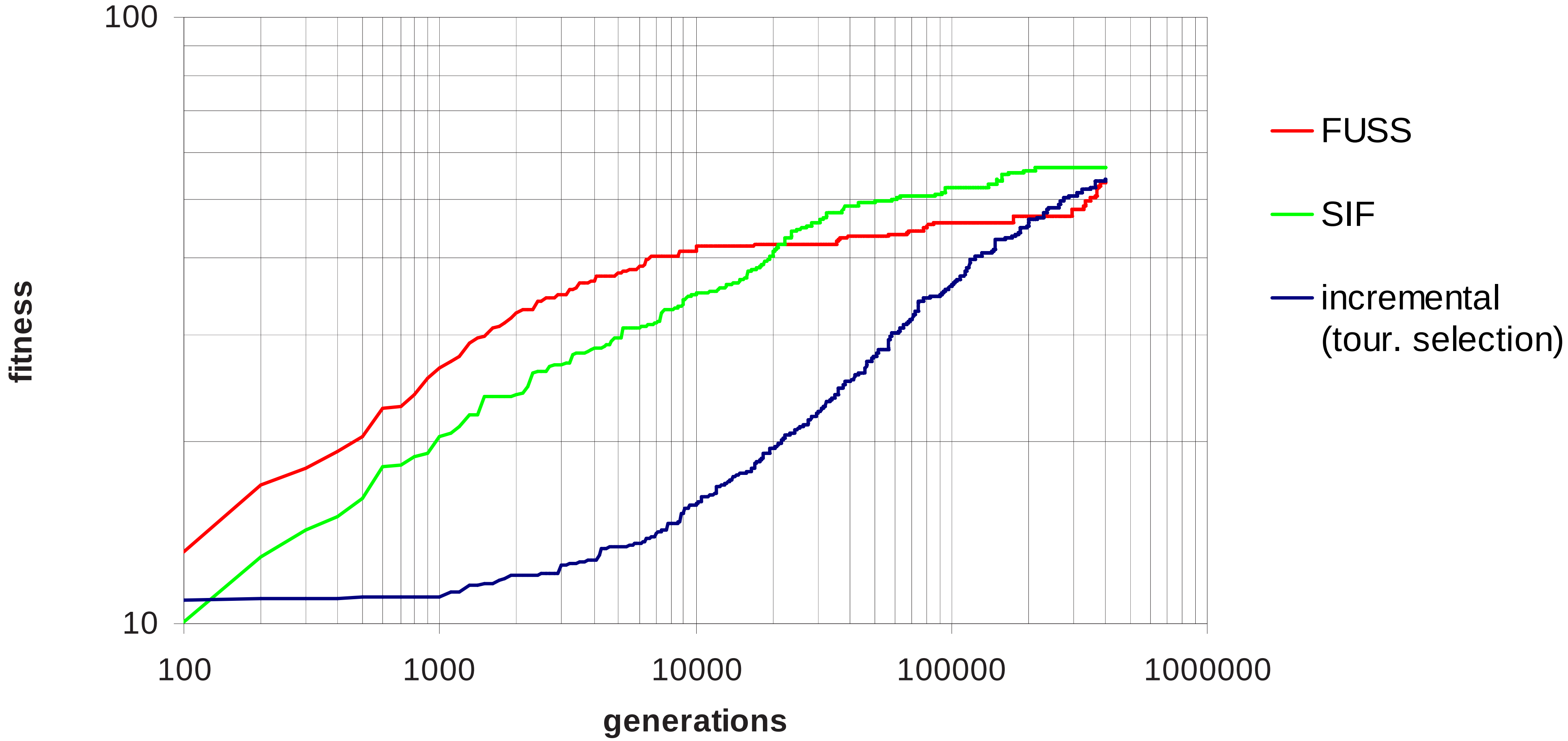
Average number of generations needed for finding exact solution of TSP with N cities.



Evolution of the best fitness in population for a random-distance TSP with 100 cities



Evolution of the best fitness in population for a random-distance TSP with 50 cities



Improvements

- There is a possible slowdown of FUSS as compared to STD on simple unimodal problems.

- Solution: A scale independent selection scheme, which is sort of a "best" compromise between greedy hill climbing and FUSS.

Worst case Analysis

- STD with extremely high selection pressure is equivalent to gradient ascent.

- Once population is fitness uniform, FUSS selects fittest individual only with probability $1/popsize$.

- For unimodal function, local mutation $x \rightarrow x \pm \varepsilon$, no crossover, most favorite STD against worst case for FUSS $\Longrightarrow$ FUSS is factor $popsize$ slower (with crossover even $popsize^2$).
- Similar problem as slowdown of proportionate selection in later optimization stages.
- Slowdown not observed in our simple 2D/3D examples and the TSP experiments (FUSS outperformed STD).
- Since real world problems often lie in between these extreme cases it is desirable to modify FUSS to cope with simple problems as well, without destroying its advantages for complex objective functions.

Scale Independent Selection (SIF)

- Goal: Optimal compromise: a high selection probability $p(f) \sim 1$ if $f \approx f_{max}$ and $p(f) \sim \frac{1}{|F|}$ otherwise.
- Approximate Solution: A "scale independent" probability distribution $p(f) \sim \frac{1}{|f_{max} - f|}$.
- Maximal slowdown: log($popsize$) - With crossover: log($popsize$)^2.

Summary & Conclusions

- Addressed problem: Balancing selection intensity in EAs, which determines speed versus quality of a solution.
- Proposed solution: Fitness uniform selection scheme FUSS.
- FUSS generates selection pressure towards sparsely populated fitness levels.
- This property is unique to FUSS as compared to other standard selection schemes (STD).

\+ The selection pressure is automatically reduced when the number of fit individuals increases.

\+ Exponential takeover and the resulting loss of genetic diversity is avoided while still generating enough selection pressure.

– It does not help in getting a more uniform distribution within a fitness level.

\+ We showed analytically by way of a simple example and numerically for the TSP that FUSS can be much more effective than STD.