Multimodal Optimization, Dynamic Optimization

- Particularities of multi-modal optimization
- Mechanisms for dealing with multiple optima
- Particularities of dynamic optimization
- Mechanisms for dealing with dynamic objective functions

Particularities of multi-modal optimization

- Aim: find all optima (global and/or local) of the objective function

- Motivation:
  - give to the decision maker not a single optimal solution but a set of good solutions
  - find all solutions with local optimal behavior

- Similar (in some sense) with: multiobjective optimization

- Applications:
  - Optimal design in engineering (e.g. induction motors)
  - Data analysis (e.g. clustering)
  - Protein structure prediction

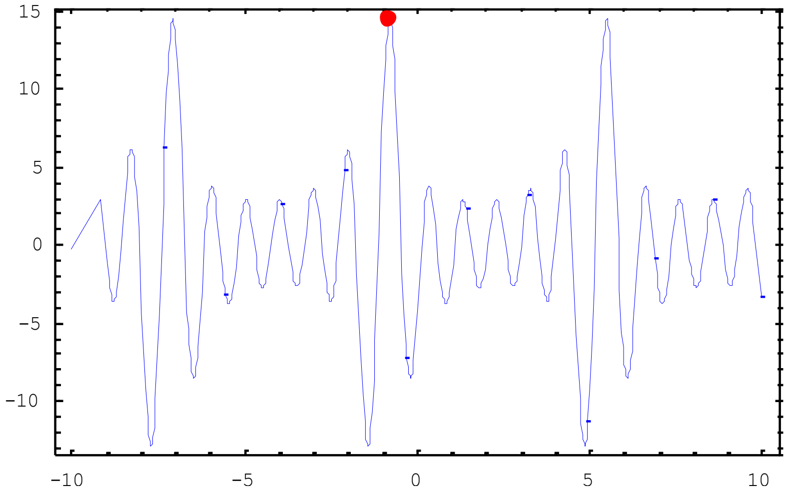
Global vs multi-modal optimization

Global optimization

Aim: find a global optimum

Evolutionary approach: population concentrates on the global optima (a single powerful species)

Premature convergence: bad

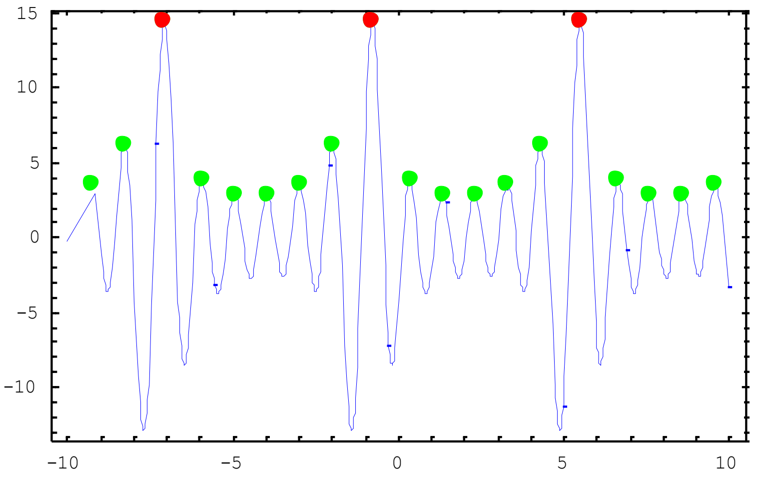


Multimodal optimization

Aim: find all (global/ local) optima

Evolutionary approach: multiple species are formed each one identifying an optimum

Premature convergence: not so bad



Approaches in multi-modal optimization

Implicit speciation (also called niching)

- species emerges in the population
- niching = finding and preserving multiple stables niches in the search space (around the potential optima)
  - Sequential – the niches are identified in several stages characterized by different mechanisms or values of some control parameters (e.g. resolution of the search)
  - Parallel – the niches are identified during one evolutionary process (based on specific selection mechanisms)
- Mechanisms which induce the formation of species:
  - fitness sharing
  - crowding
- Remark: it is usual combined with archiving (the “good” configurations are collected in the population or in an additional archive)

Approaches in multi-modal optimization

Explicit speciation

- The population is divided into communicating subpopulations which evolves in parallel; the communication should be rare in order to allow the preservation of the species (and avoid the full mixing of the population elements)
- The division of a population in subpopulations can be based on a clustering process (based on a specific similarity measure)
- Each subpopulation corresponds to a species whose aim is to populate a niche in the fitness landscape and to identify an optimum

Implicit speciation: fitness sharing

Fitness sharing:

- Discourage the agglomeration of many elements in the same search region by penalizing the fitness function based on a distance related factor

[see lecture 10 – fitness sharing is used also for multiobjective optimization]

Related idea: clearing [Petrowski, 1997] = the elements in the same niche are eliminated (by setting their fitness to 0)

Example for a maximization problem

Pseudo-code for Clearing technique

```
function Clearing(σ , C) /∗ Function definition∗/
begin
Sort_Fitness (P) /∗ Sort the population in descending or-
der according to their fitness∗/
 fori = 0 to (Np − 1)
      if (Fitness(P[i]) > 0)
       n_Win = 1
       forj = (i + 1) to(Np − 1)
        if (Fitness(P[j]) > 0 AND Distance(P[i], P[j]) < σ)
          ifn_Win < C
          n_Win = n_Win + 1
          else
          Fitness (P[j]) = 0
          endif
        endif
       endfor
      endif
 endfor
end
```

Niche radius (σ)

Niche size (C)

[S. Das et al. Real-parameter evolutionary multimodal optimization – a survey, SWEVO, 2011]

Implicit speciation: crowding

Main idea:

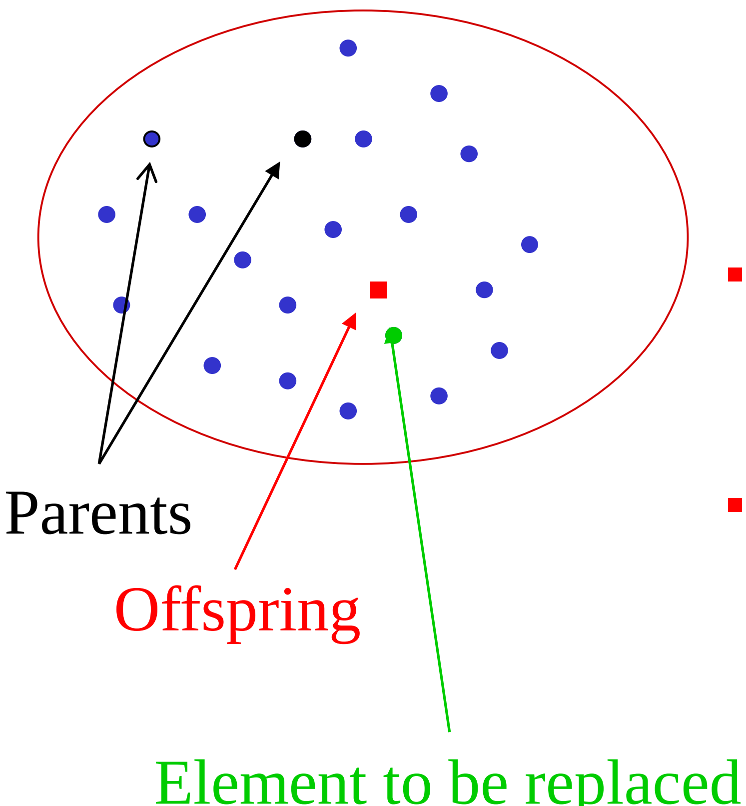
- Elements belonging to different species should not compete between them
- The selection process involve similar elements → this allows to maintain the existing diversity in the population

Implementation:

- A new element replaces the closest old element in a population sample (the ratio between the population sample and the population size is called crowding factor, CF); CF=1 → global crowding
- The distance between elements is computed in the search space

Implicit speciation: crowding

- Idea of crowding



- Particularities
  - preserves the population diversity
  - encourages the species formation
  - the final population consists of elements concentrated around the optima
- Advantage:
  - simple
- Disadvantage:
  - global character of crowding ($O(m^2)$)

Implicit speciation: crowding

- Implementation of deterministic crowding [Mahfoud, 1995]

- to be applied for m/2 times (m=pop size)

Pseudo-code for deterministic crowding

1. Select two parents p_1, p_2 randomly with no replacement.
2. Perform a crossover between them yielding offspring c_1, c_2.
3. Apply mutation operator to generate c'_1, c'_2.
4. if $[d(p_1, c'_1) + d(p_2, c'_2) \leq d(p_1, c'_2) + d(p_2, c'_1)]$
   - if $f(c'_1) \geq f(p_1)$ replace p_1 with c'_1
   - if $f(c'_2) \geq f(p_2)$ replace p_2 with c'_2

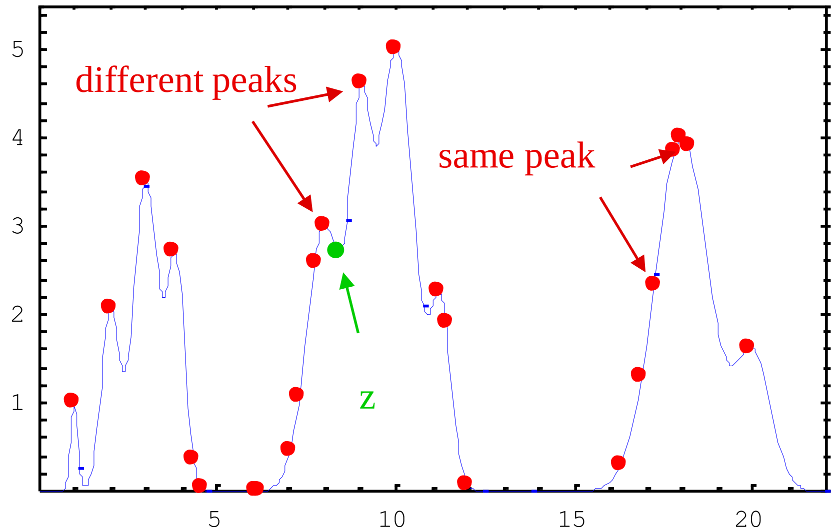
   else
   - if $f(c'_2) \geq f(p_1)$ replace p_1 with c'_2
   - if $f(c'_1) \geq f(p_2)$ replace p_2 with c'_1

[S. Das et al. Real-parameter evolutionary multimodal optimization – a survey, SWEVO, 2011]

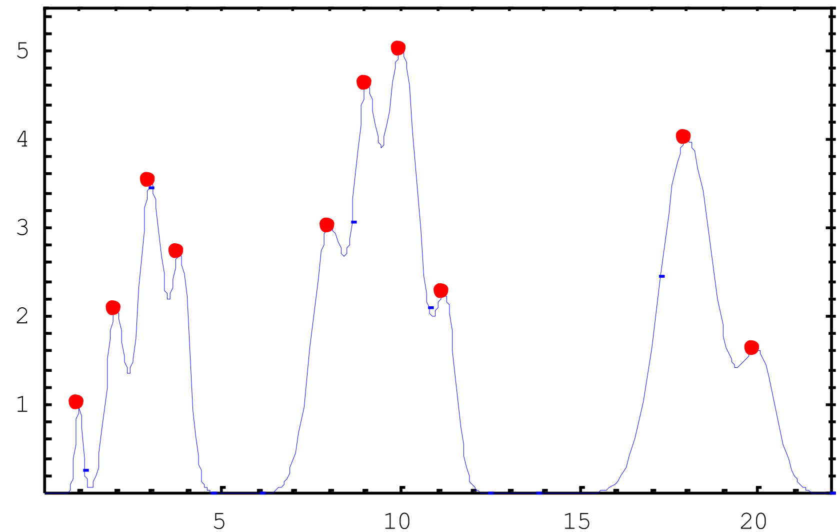
Archiving

- Hill-valley function [Ursem, Multinational Evolutionary Algorithms, 1999]:
  - If there exists c in (0,1) such that

    z=cx+(1-c)y implies f(z)<f(x) and f(z)<f(y)
  - then there exists a valley between x and y → they belong to different "peaks" → both of them may remain in the archive
  - The decision is based on computing z for some values of c



Archive without valley detection



Archive with valley detection

Explicit speciation

Main idea:

- the population is explicitly divided into sub-populations
- each sub-population aims to identify one or several optima

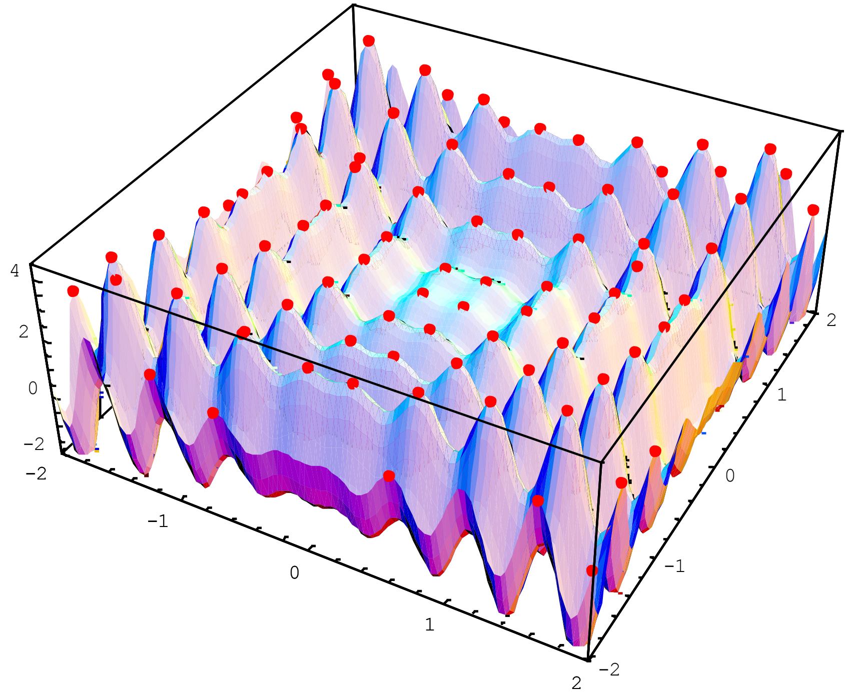
- Question: which is the influence of the communication between sub-populations?

  - Example: Multipopulation Multiresolution DE (2004)
  - Notations:
    - m = number of elements in the subpopulation
    - s = number of subpopulations
    - g = Number of generations
    - f = frequency of migration (generations between 2 migration stages)
    - p_m= migration probability

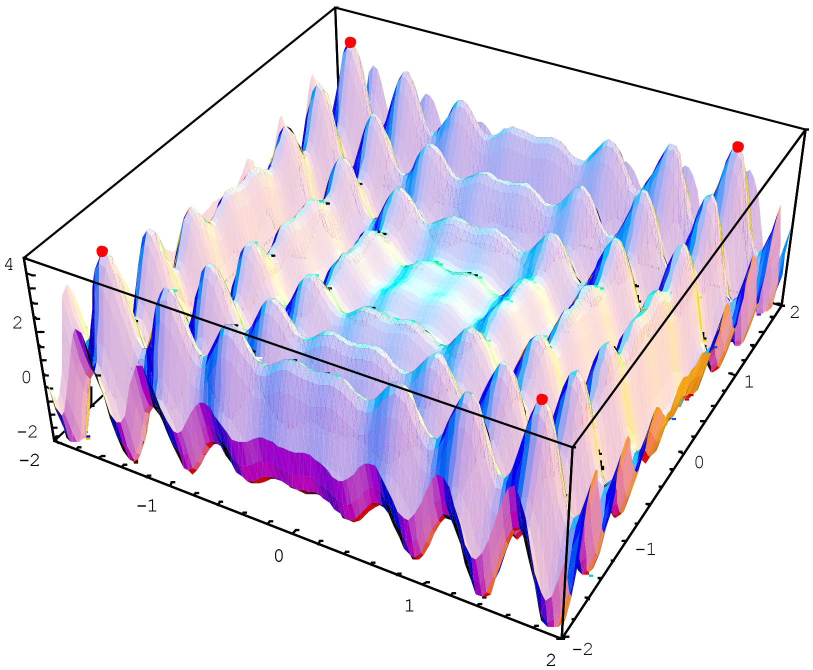
Explicit speciation

- Test function: multi-peaks

$$f(x,y) = x\sin(4\pi x) - y\sin(4\pi y + \pi),\ \ x,y \in [-2,2]$$



m=5, s=50, g=50, e=20, f=0, p_m=0
(91 elements in the archive)



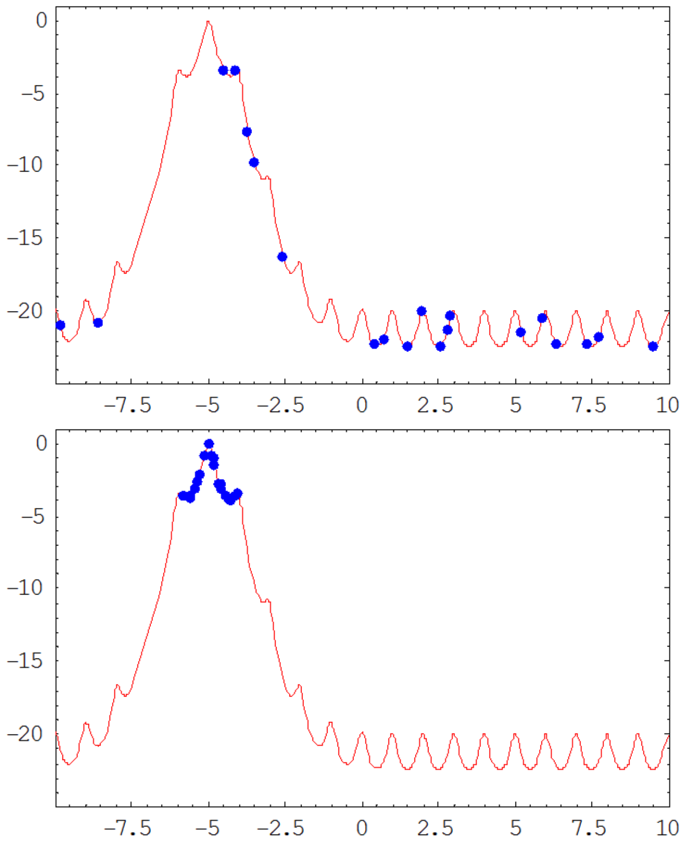
m=5, s=50, g=100, f=10, e=20, p_m=0.5
(4 elements in the archive)

Dynamic optimization

- Dynamic optimization:

$$f(t, x^*(t)) \geq f(t, x), \;\; x \in D^n$$

- Aim: track a changing optimum
- Difficulty: inability to track the optimum
- Cause: the population lost its diversity
- Solution: stimulating the population diversity



Dynamic optimization

- Dynamic optimization:

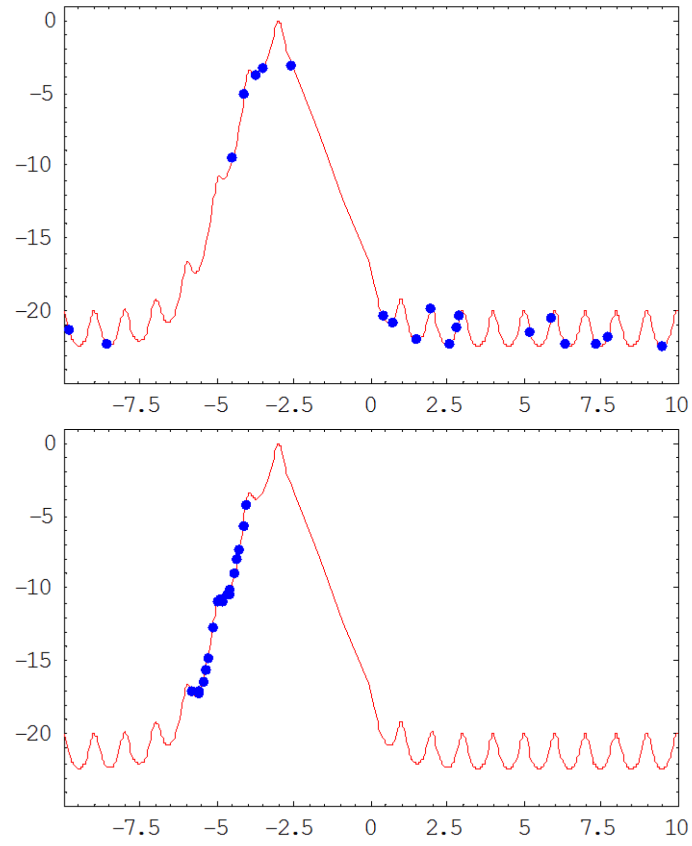
$$f(t, x^*(t)) \geq f(t, x), \;\; x \in D^n$$

- Aim: track a changing optimum

- Difficulty: inability to track the optimum

- Cause: the population lost its diversity

- Solution: stimulating the population diversity



Dynamic optimization

- Dynamic optimization:

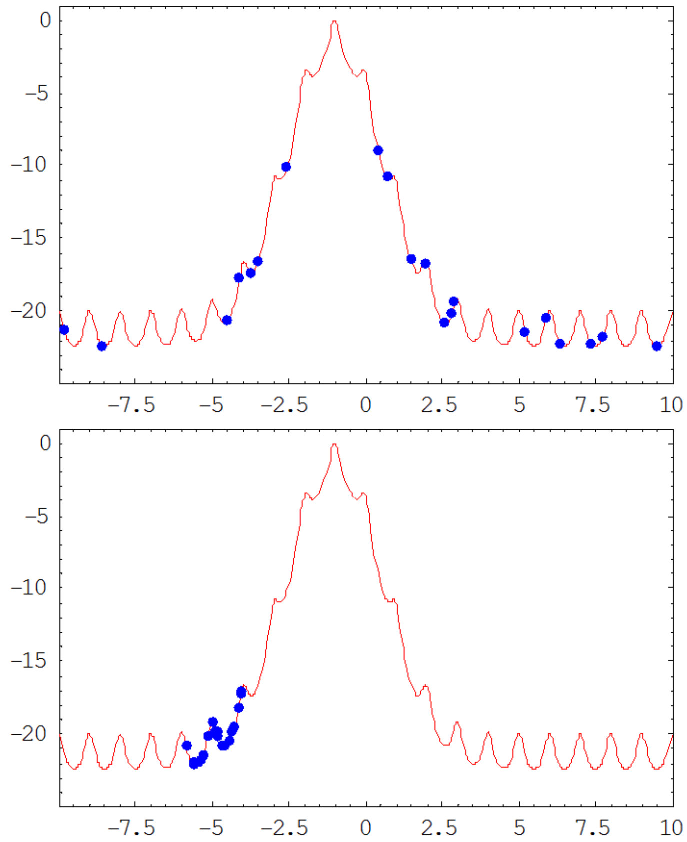
$$f(t, x^*(t)) \geq f(t, x), \;\; x \in D^n$$

- Aim: track a changing optimum

- Difficulty: inability to track the optimum

- Cause: the population lost its diversity

- Solution: stimulating the population diversity

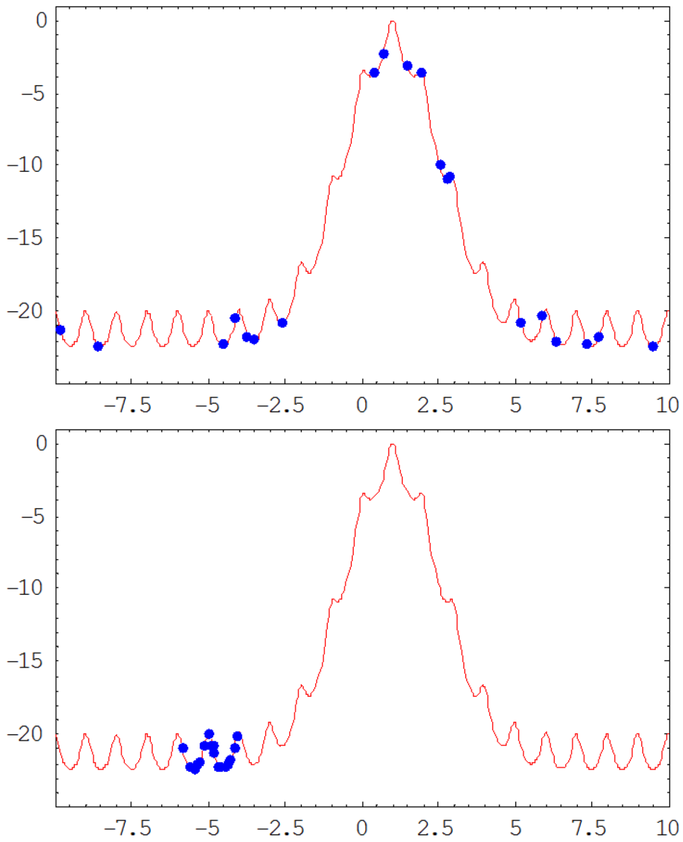


Dynamic optimization

- Dynamic optimization:

$$f(t, x^*(t)) \geq f(t, x), \quad x \in D^n$$

- Aim: track a changing optimum
- Difficulty: inability to track the optimum
- Cause: the population lost its diversity
- Solution: stimulating the population diversity

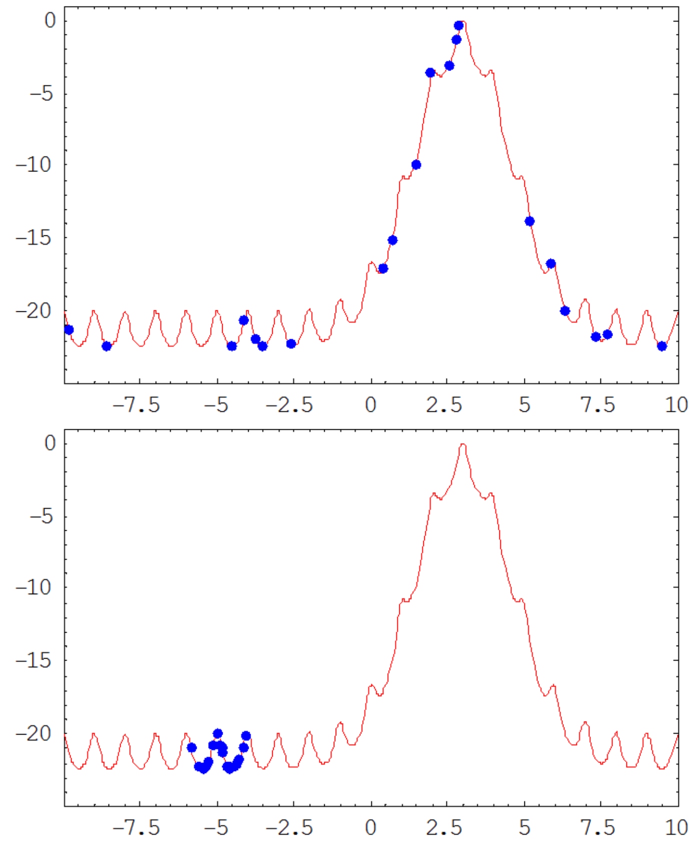


Dynamic optimization

- Dynamic optimization:

$$f(t, x^*(t)) \geq f(t, x), \;\; x \in D^n$$

- Aim: track a changing optimum
- Difficulty: inability to track the optimum
- Cause: the population lost its diversity
- Solution: stimulating the population diversity



Dynamic optimization - characteristics

- Types of optimum changes:
  - continuous trajectory of the optimum (smooth changes)
  - discontinuous trajectory of the optimum (jumps of the optimum position)

Main questions:

- Time-linkage: the future behaviour depends on the current state?
- Detectability: can be the changes detected?
- Predictability: can be the changes predicted? are the changes periodic?

Dynamic optimization – problems

- Practical applications:
  - dynamic scheduling
  - dynamic resource allocation
  - dynamic risk minimization
  - other problems with parameters changing in time

- Test problems
  - Moving Peaks Benchmark - MPB [Branke, 1999]
  - Dynamic Knapsack Problem - time dependent weights and profits
  - Dynamic TSP – time dependent costs

Dynamic optimization – specific mechanisms

- Reactive approaches: trigger a diversity increasing mechanism when a change is detected
  - hypermutation (high mutation rate)
  - random immigrants (random elements are introduced in the population)
  - memory – useful in the case of recurrent or periodic changes
    - Implicit (diploid elements)
    - Explicit (store the best elements at various stages of the evolution and re-use them when a change is detected) – similar to archiving

Remarks:

- Appropriate for discontinuous changes of the optima
- How can be detected a change?
  - Use of some detectors (elements of the population which are periodically re-evaluated) - if their current fitness value is different from the previous one then a changed occurred

Dynamic optimization – specific mechanisms

Diploid elements:

- usually applied for genetic algorithms with binary encoding
- each element ,X_i, from the population consists of two chromosomes C1(i) and C2(i)
- when the fitness of X_i is evaluated only the dominant components are taken into account

If C1(i,j)= C2(i,j)

then X_i (j)=C1(i,j) // same value of component j in both chromosomes

Else

if rand(0,1)<Pd(j,g) then X_i (j)=C1(i,j)

else X_i (j)=C2(i,j)

Dynamic optimization – specific mechanisms

Diploid elements:

- The selection of the dominant component is based on the dominance probability vector Pd(j,g) which is adjusted at each generation:

  Pd(j,g+1)=(1-alpha)*Pd(j,g)+alpha*Pbest(j,g)

- Variants:

  - Pbest(g) = mask corresponding to the best element in the population

  i.e. Pbest(j,g)=1 if the component j from the best element is taken from C1

  [S. Yang, Dominance Learning in Diploid GA for DOP, 2006]

  - Pbest(g) = frequency vector computed based on the best vectors identified during the evolutionary process

Dynamic optimization – specific mechanisms

- Proactive approaches: maintain the population diversity by:
  - diminishing the selection pressure (sharing, crowding)
  - directly stimulating the diversity (periodically inserting random elements)

Remarks

- the proactive approaches are appropriate for continuously changing optima