Recent Advances in Multi-Modal Optimization using niching methods

Michael G. Epitropakis,
Data Science Institute
Department of Management Science,
Lancaster University, Lancaster, UK
Email: m.epitropakis@lancaster.ac.uk

Xiaodong Li,
School of Science,
RMIT University, Melbourne, Australia
Email: xiaodong.li@rmit.edu.au

Andries Engelbrecht
Department of Computer Science,
School of Information Technology,
University of Pretoria, Pretoria, South Africa
Email: engel@driesie.cs.up.ac.za

IEEE CEC'2017, Donostia - San Sebastián, Spain, June, 2017

Instructors

Xiaodong Li,

School of Science,
RMIT University, Melbourne,
Australia
xiaodong.li@rmit.edu.au

Andries Engelbrecht

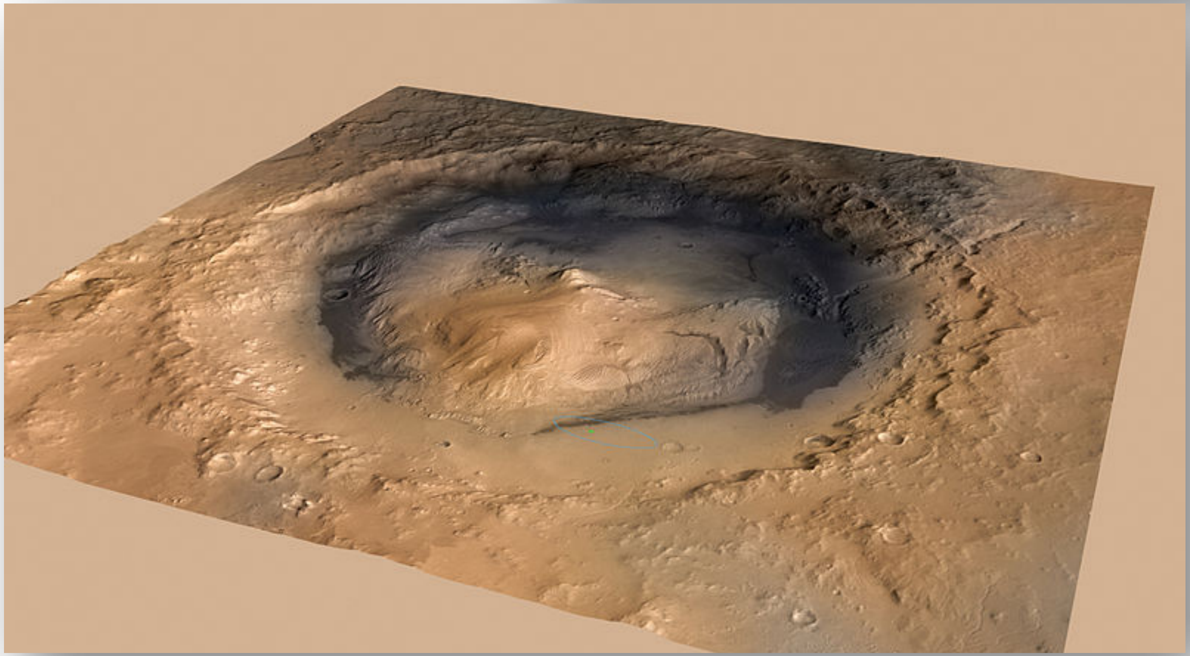
Department of Computer
Science, School of Information
Technology,
University of Pretoria, Pretoria
South Africa
engel@driesie.cs.up.ac.za

Michael G. Epitropakis,

Data Science Institute,
Department of Management
Science,
Lancaster University,
Lancaster, United Kingdom
m.epitropakis@lancaster.ac.uk

Multi-Modal Optimization?

Curiosity's view of "Mount Sharp" (September 9, 2015)



Mount Sharp rises from the middle of Gale Crater; the green dot marks Curiosity's landing site (north is down).

Source: https://en.wikipedia.org/wiki/Curiosity_(rover)

General Notes

- More questions than answers in **Multi-Modal Optimization** (**MMO**)
- Limited theoretical advances/strict formulations
- Huge amount of literature (from the 80s onwards)
- Tutorial: a “short” presentation on advances in the field.
- **Stay connected**… More to come in the near future…

- Please **interrupt** for questions/comments
- **Suggestions** for future work are more than welcome

Outline

- Background on multimodal optimization
- What are the benefits for studying MMO?
- What are niching methods?
- Some real-world examples
- Classic niching methods
- Niching methods derived from PSO and DE
- Other state-of-the-art niching methods
- Niching benchmark suites and performance measures
- Niching in specialized tasks
  - Dynamic and multi-objective optimization
  - Clustering, feature selection, and machine learning
- IEEE CIS Taskforce on Multi-Modal Optimization
- Discussion and summary
- References

What is multi-modal optimization?

- **Multi-modal Optimization (MMO)**: to locate multiple optimal (or close to optimal) solutions in the search space
  - This is different from a conventional optimization method which has a common goal of seeking to locate a single global optimum

- A rough **definition**:

*In a multimodal optimization task, the **main purpose** is **to find multiple optimal solutions** (global and local), so that the user can have a **better knowledge** about different optimal solutions in the search space and as and when needed, the current solution may be switched to another suitable optimum solution*

Deb, Saha: [Multimodal Optimization Using a Bi-Objective Evolutionary Algorithm](), ECJ, 2012

What is multi-modal optimization? (II)

- **MMO** problems represent an important class of optimization problems
- Many **real-world optimization problems**:
  - multimodal **by nature**
  - **multiple satisfactory solutions exist** (several real-world examples of MMO problems are provided in subsequent slides)
- From a **decision maker's** point of view:
  - it might be desirable to locate all global optima and/or some local optima that are also considered as being satisfactory
  - Better knowledge of alternative solutions

Methods for MMO

- Optimization methods specifically designed for solving MMO problems:
  - often called **multimodal optimization** or **niching methods**
  - predominately developed from the field of **meta-heuristic algorithms**
  - Covers the family of **population-based stochastic optimization algorithms**, including evolutionary algorithms, evolutionary strategies, particle swarm optimization, differential evolution, and so on
- These meta-heuristic algorithms are shown ***particularly effective*** in solving multimodal optimization problems, if equipped with specifically designed ***diversity preserving mechanisms***, commonly referred to as **niching methods**
- Two-fold aim: accurately **locate** and robustly **maintain** multiple optima

What are the benefits?

- A decision maker may be interested to know whether there exist ***multiple equally good solutions*** before making a final decision
- Important for a **sensitivity** study of a problem, and helps develop more **robust solutions** to the problem
- Plays an important role in **keeping a diverse population** of candidate solutions, hence helps **prevent** the population from **converging prematurely** to a sub-optimum
- May increase the probability of finding the global optimum

Different scenarios

- One-global optimum:
  - Looking for the global optimum solution only (not MMO)
- All-global optima:
  - Find all the global optimum solutions
  - Benchmark problems of the CEC 2013/2015/2016 niching competition series belong here
- All-known optima:
  - Find all local and/or global optimal solutions
- Approximate set of solutions:
  - Locate as many as possible (subset) optimal solutions (global/local) that are well distributed over the search space

Mike Preuss. 2015. Multimodal Optimization. GECCO Companion '15, http://dx.doi.org/10.1145/2739482.2756572

Ecological inspiration

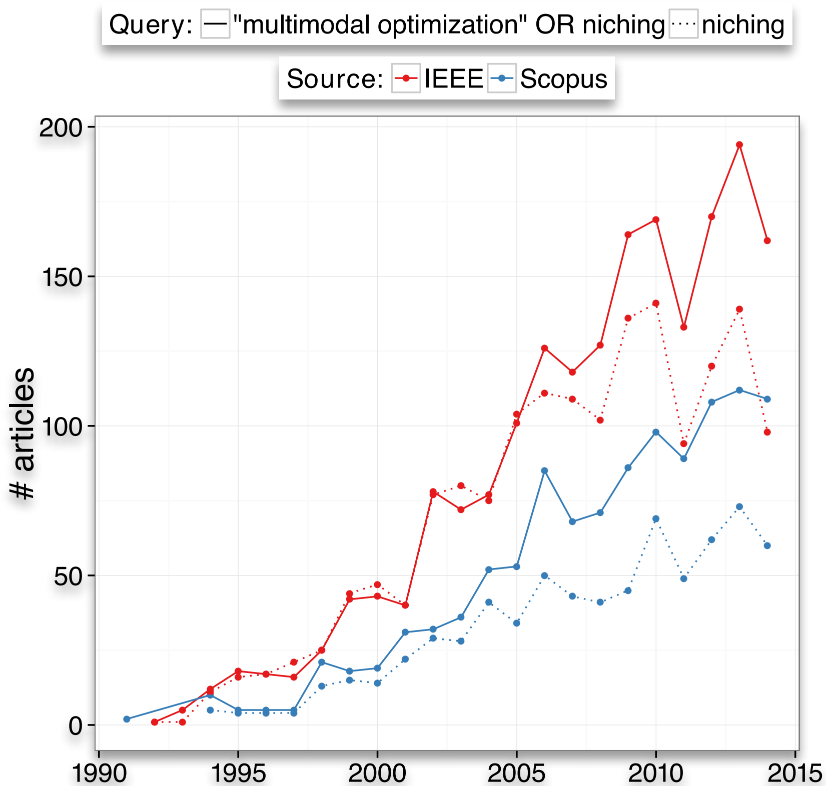
- In natural ecosystems, individual species must compete to survive by taking on different roles. Different species evolve to fill different niches (or subspaces) in the environment that can support different types of life

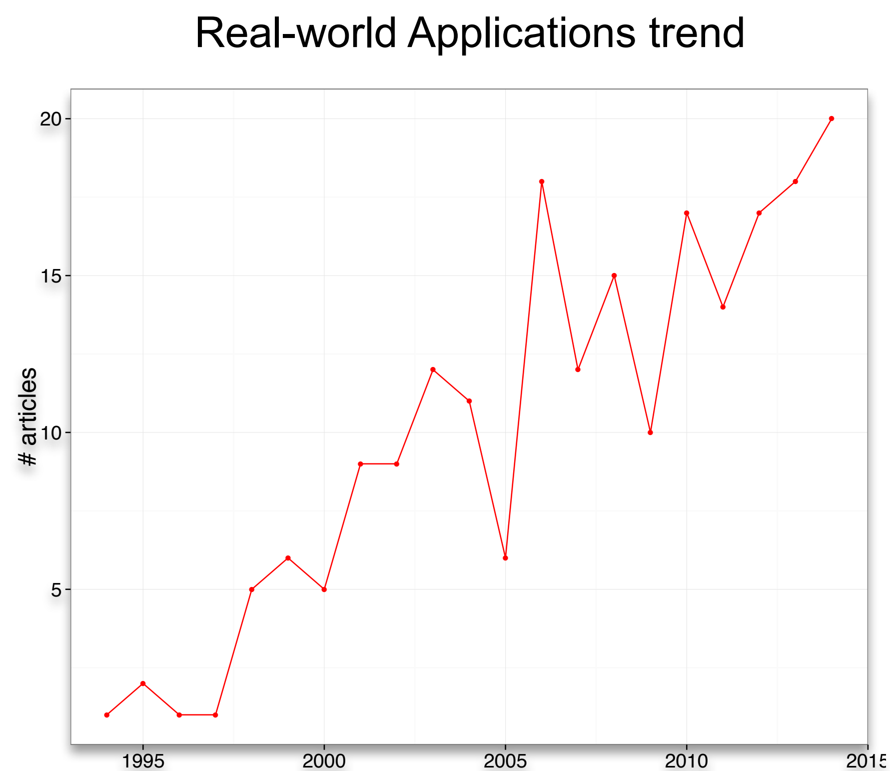
What are niching methods?

- According to the Oxford Dictionary, a **niche** refers to "*a role taken by a type of organism within its community*"; and a **species** refers to "*a group of living organisms consisting of similar individuals capable of exchanging genes or interbreeding*"
- These concepts of niches, species and speciation can be adopted in an EA to encourage an EA population to evolve different species targeting different optimal solutions in the search space

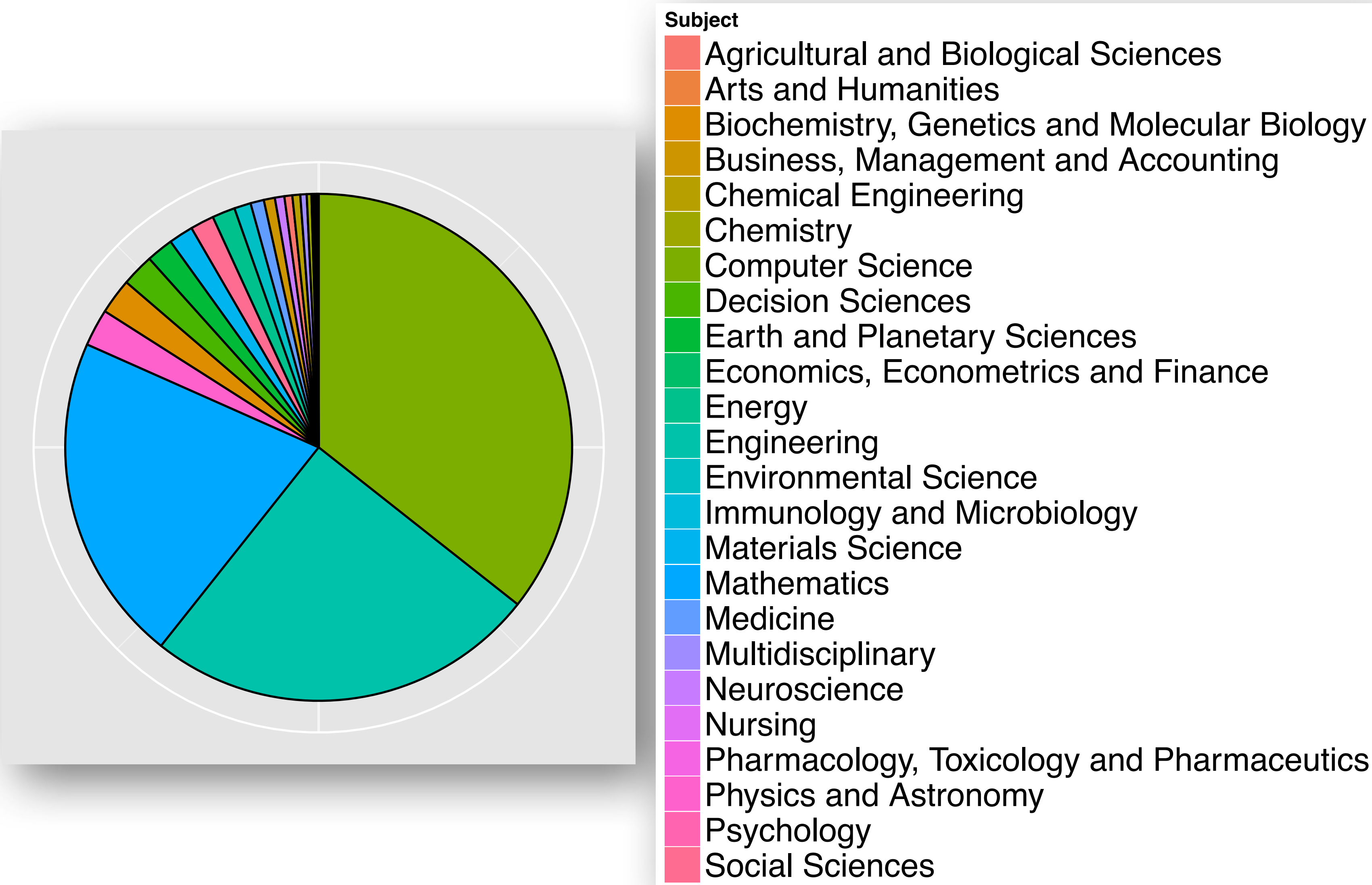
MMO Publication trends

- Despite niching methods first appeared more than 30 years ago, currently niching techniques are experiencing a revival, attracting researchers from across a wide range of research fields





MMO Application areas



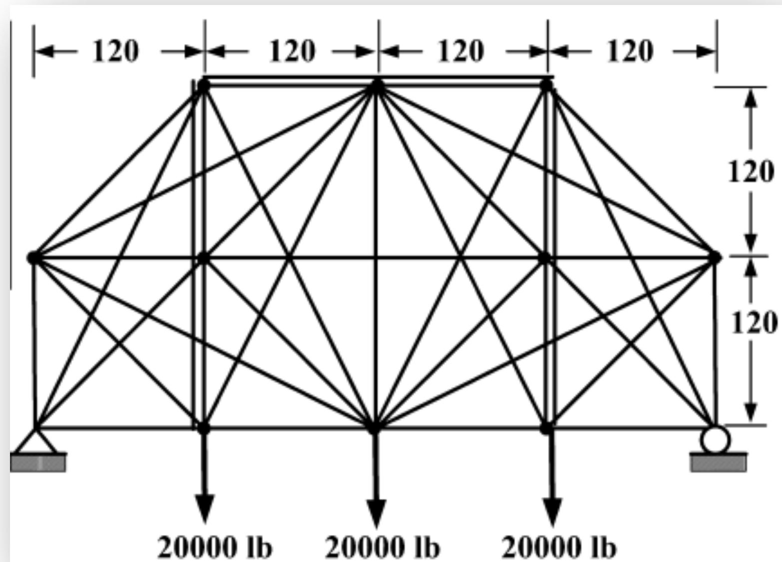
Multi-Modal Optimization

REAL-WORLD APPLICATIONS

Engineering example: truss topology design

- In **topology optimization**, the connectivity of members in a truss is to be determined. There exist ***multiple different topologies*** with almost equal overall weight in truss-structure design problems as the members in the ground structure increase
- The resulting solution of truss-structure optimization design problems becomes ‘‘**multi-modal**’’ with large number of truss members

Some nodes in the **ground structure** may or may not be removed. The optimal structure is found as a subset of the ground structure



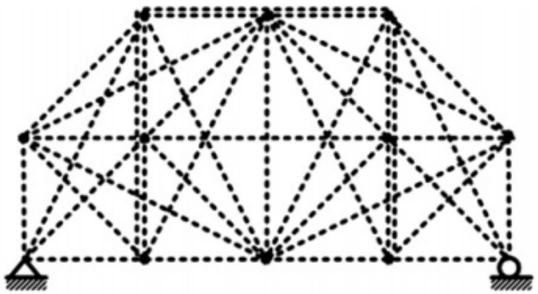
Deb K, Gulati S. “Design of truss-structures for minimum weight using genetic algorithms,” *Finite Elements Anal Des* 2001; 37: 447–65.

G.-C. Luh and C.-Y. Lin, “Optimal design of truss-structures using particle swarm optimization,” *Computers and Structures*, vol. 89, no.23-24, pp. 2221 – 2232, Dec. 2011.

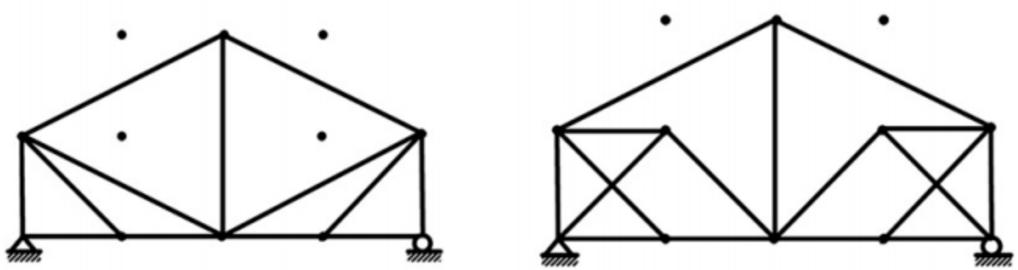
Truss topology design

- **Sharing scheme** is used to compute the similarity between different topology design solutions
- The **sharing fitness** is a reduced one from the original fitness, in order to discourage solutions in the vicinity

Binary PSO is run based on the sharing fitness values, and multiple dissimilar truss topologies are derived and saved

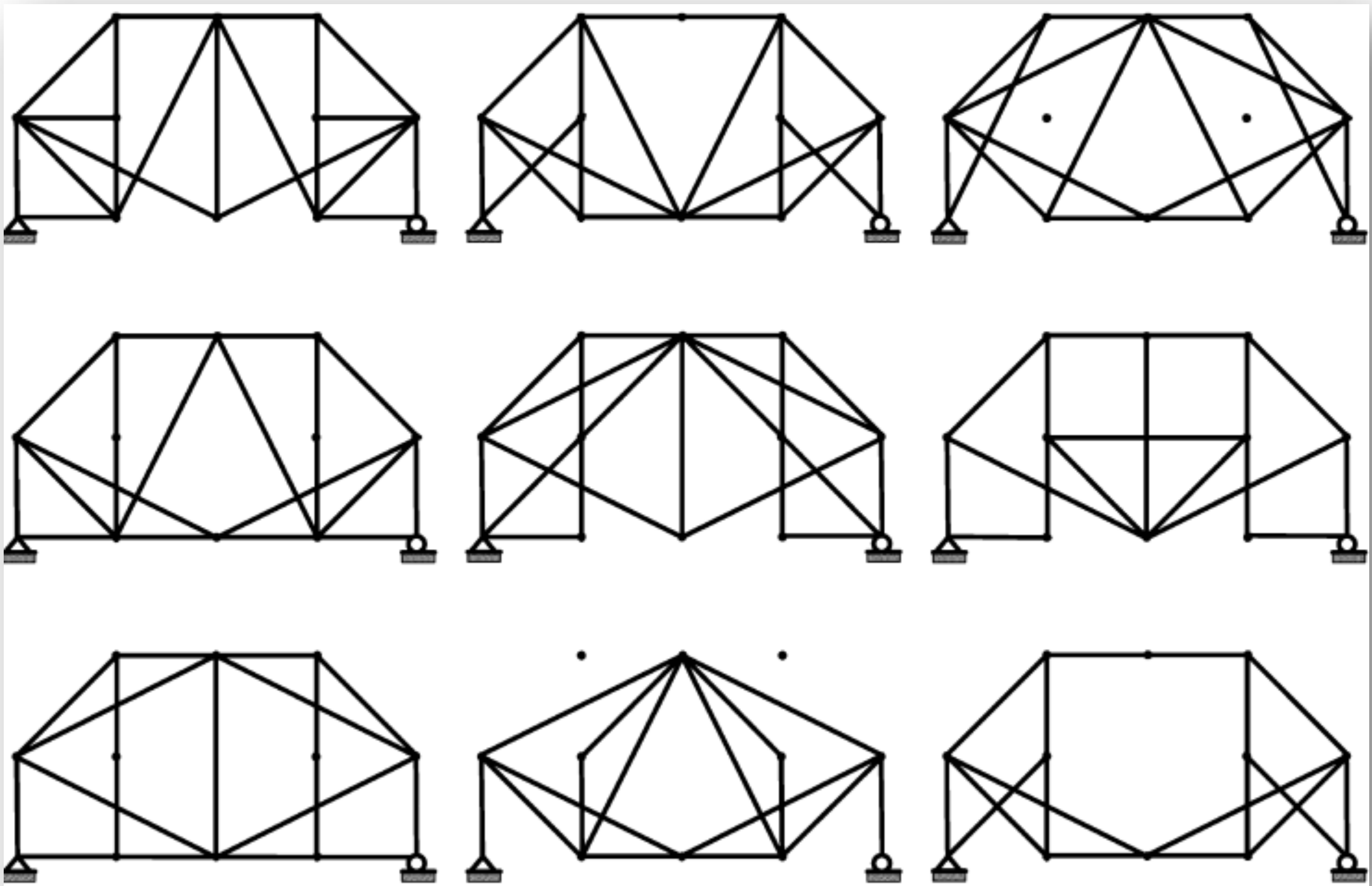


(*a*) Two-tier, 39-member, 12-node ground structure



(*b*) Two-tier, 13-member, 8-node truss (*c*) Two-tier, 17-member, 10-node truss

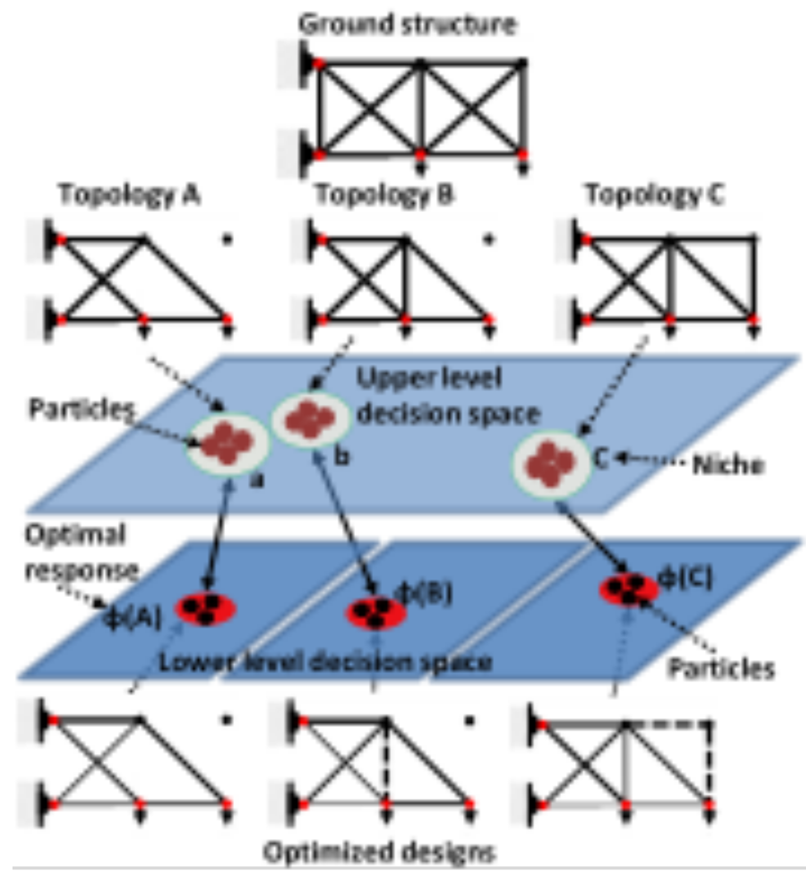
Truss topology design



Multiple optimal truss topologies found by BPSO with niching.

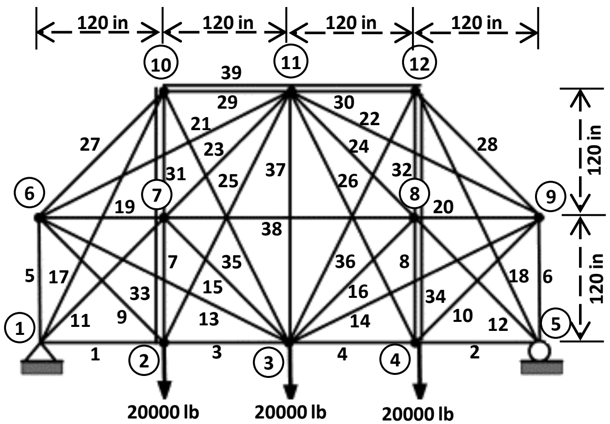
Trust structure design using Bilevel and niching aspects

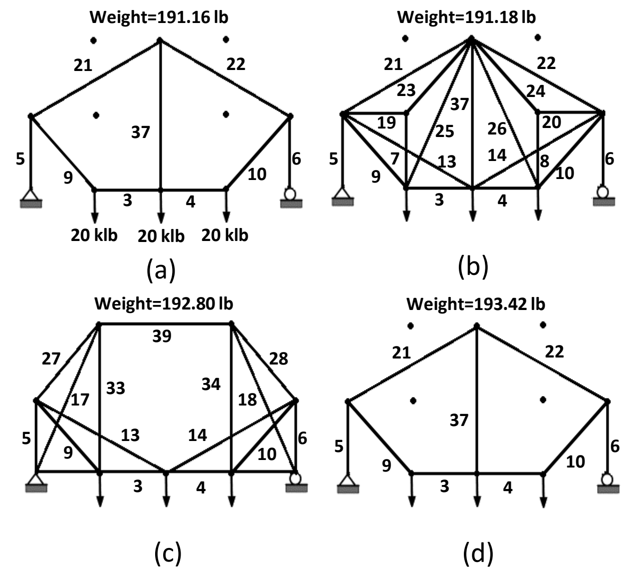
- Formulate the truss problem as a **bilevel** optimization problem
- A new bilevel PSO niching method locates multiple optimal solutions
- **Stable topologies** can be found in the **upper level**
- The **optimized sizes** of the members of these topologies can be found in the **lower level**
- **Niching** at the **upper level**
- **Standard optimizer** is used at the **lower level** to optimize a bilevel truss problem

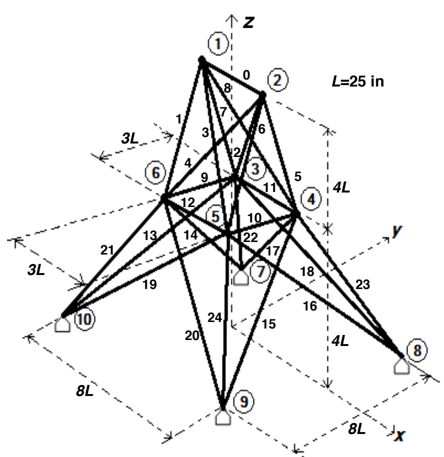


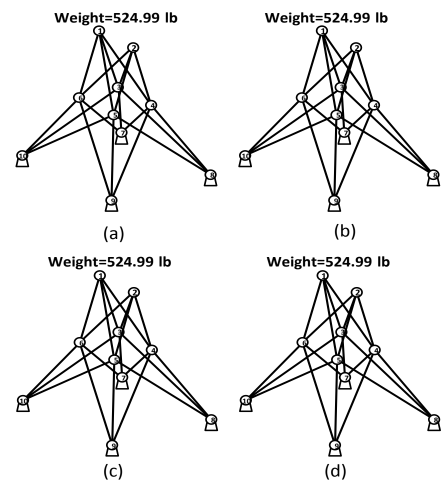
Md. Jakirul Islam, Xiaodong Li, and Kalyanmoy Deb. 2017. Multimodal Truss Structure Design Using Bilevel and Niching Based Evolutionary Algorithms. In Proceedings of GECCO '17, Berlin, Germany, July 15-19, 2017, DOI: h p://dx.doi.org/10.1145/3071178.3071251

Trust structure design examples



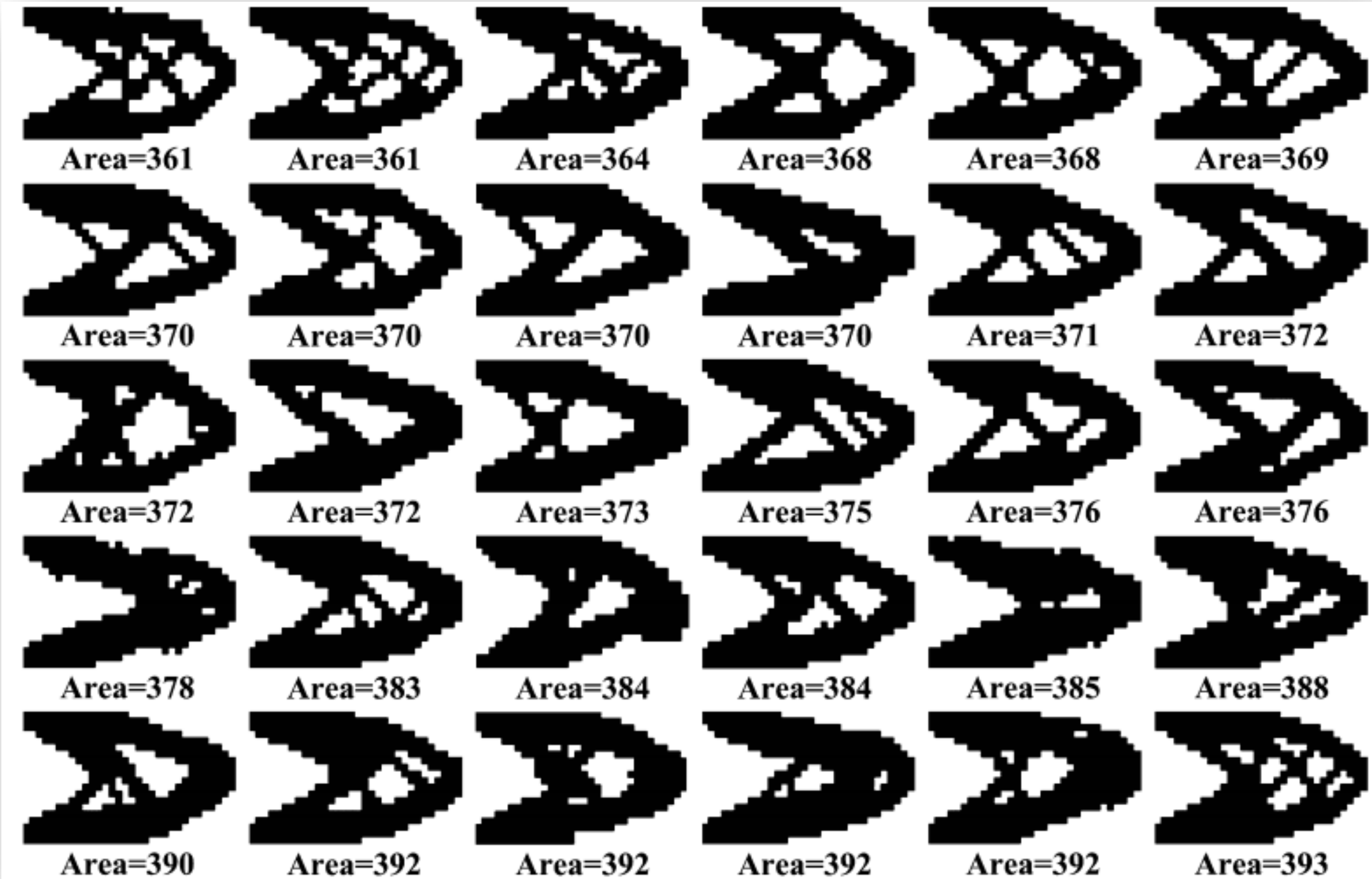






Md. Jakirul Islam, Xiaodong Li, and Kalyanmoy Deb. 2017. Multimodal Truss Structure Design Using Bilevel and Niching Based Evolutionary Algorithms. In Proceedings of GECCO '17, Berlin, Germany, July 15-19, 2017, DOI: h p://dx.doi.org/10.1145/3071178.3071251

Continuum structural topology optimization



G.-C. Luh, C.-Y. Lin, Y.-S. Lin, "A binary particle swarm optimization for continuum structural topology optimization", *Applied Soft Computing*, Volume 11, Issue 2, March 2011, Pages 2833-2844, ISSN 1568-4946,

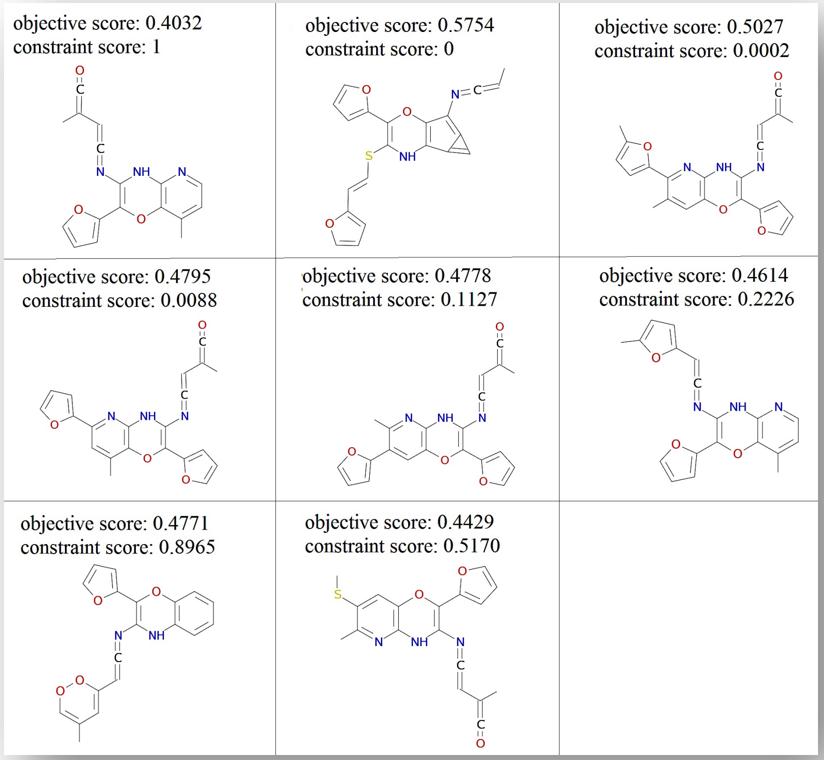
Drug Molecule Design (I)

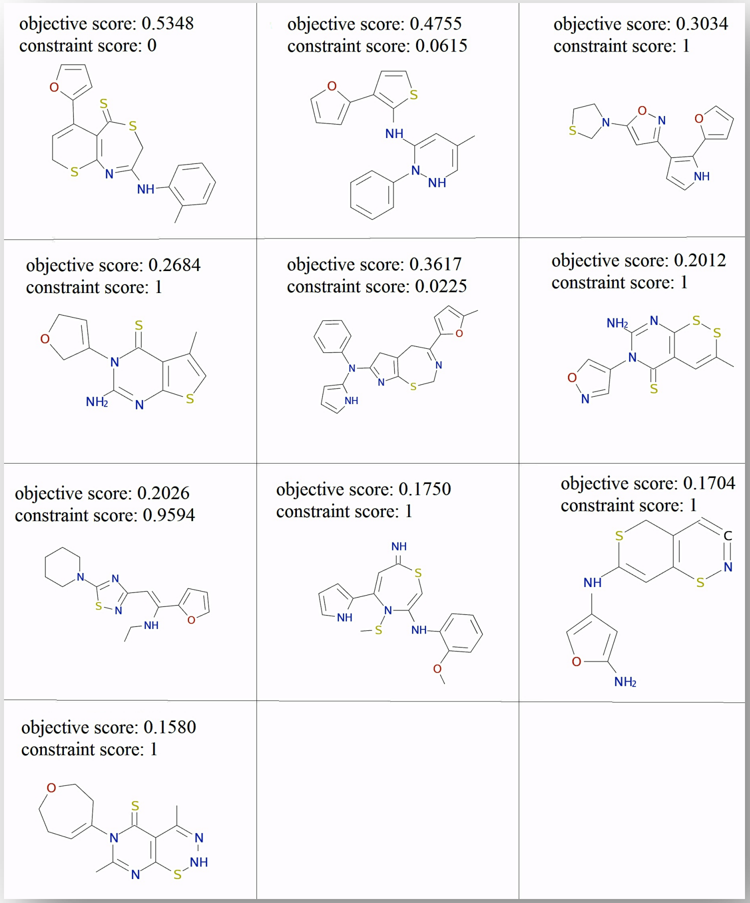
- Search for **molecular structures** with specific pharmacological or biological activity that influence the behavior of certain targeted cells
- **Objectives**: Maximization of potency of drug & Minimization of side-effects
- **Aim:** provide the medicinal chemist a **set of diverse molecular structures** that can be promising candidates for further research
  - **Fit solutions may** result in finding structures that **fail in practice**
  - The chemist **desires a set of promising structures** rather than only one single solution

J. W. Kruisselbrink, A. Aleman, M. T. M. Emmerich, A. P. Ijzerman, A. Bender, T. Baeck, and E. van der Horst, "Enhancing search space diversity in multi-objective evolutionary drug molecule design using niching," GECCO'09, 2009, pp. 217–224.

Drug Molecule Design (II)

- Dynamic Niche Sharing technique incorporated to MOEA





Without (left) and with (right) niching

J. W. Kruisselbrink, A. Aleman, M. T. M. Emmerich, A. P. Ijzerman, A. Bender, T. Baeck, and E. van der Horst, “Enhancing search space diversity in multi-objective evolutionary drug molecule design using niching,” GECCO’09, 2009, pp. 217–224.

Electromagnetics Optimization

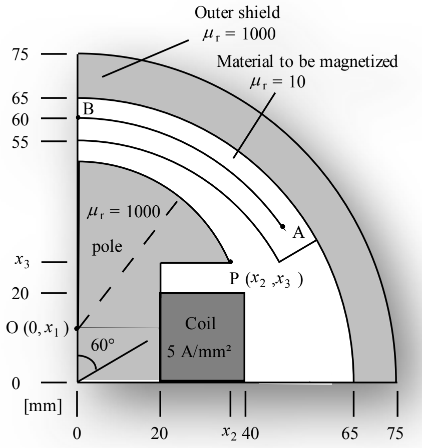


Fig. 2 Geometry of the magnetizer

Many niching techniques have been used to address real-world problems in Electromagnetics:

- Restricted Tournament Selection
- Deterministic Crowding
- Sharing
- Clearing

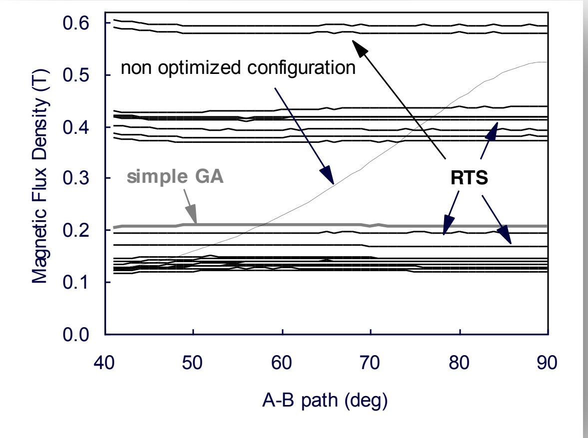


Fig. 3. Examples of uniform induction levels detected by a run of Restricted Tournament Selection (RTS). Optimal solutions are compared with that obtained with a simple GA and a non-optimized configuration

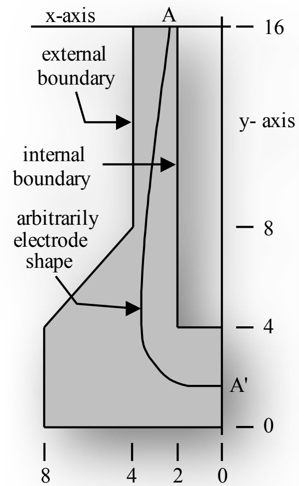


Fig. 1. Electrode template

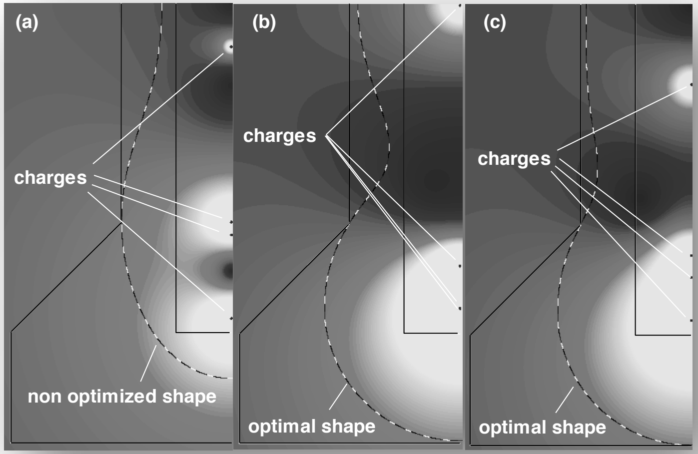
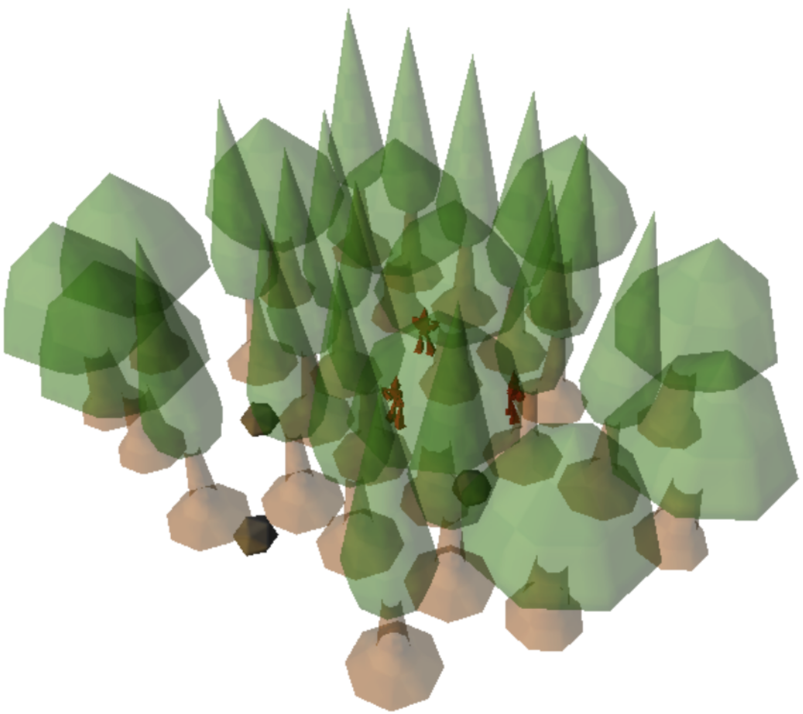


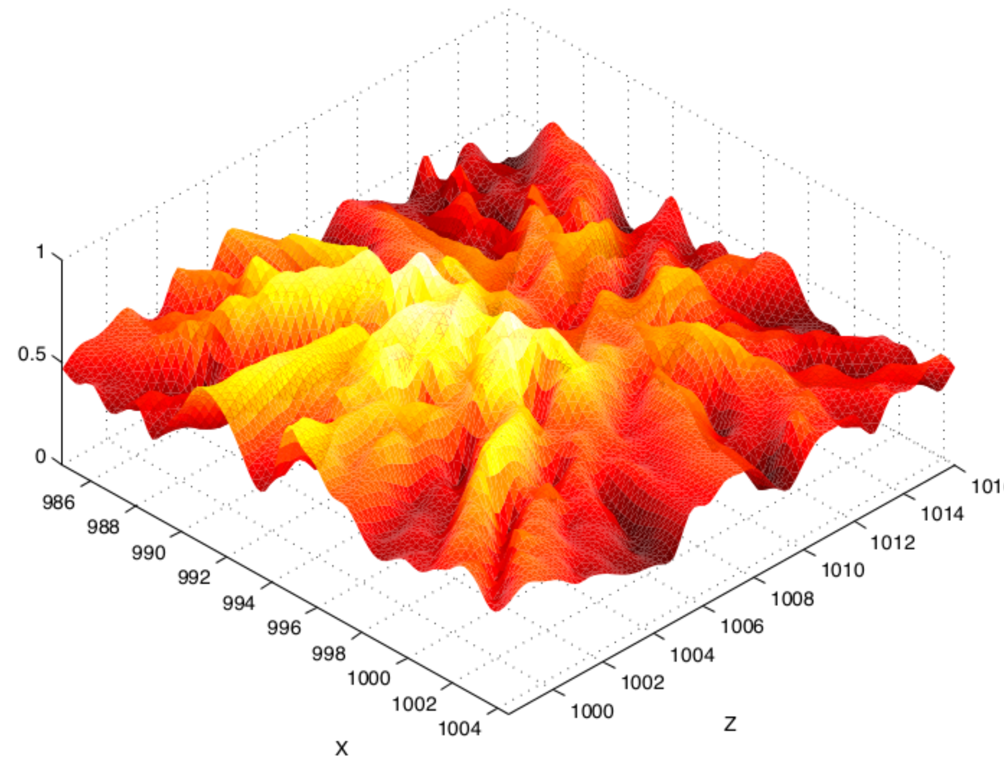
Fig. 3. Examples of profiles with the corresponding electric field

B. Sareni, L. Krahenbuhl, and A. Nicolas, “Niching genetic algorithms for optimization in electromagnetics. i. fundamentals,” IEEE Transactions on Magnetics, vol. 34, no. 5, pp. 2984–2987, 1998.

B. Sareni, L. Krahenbuhl and D. Muller, "Niching genetic algorithms for optimization in electromagnetics. II. Shape optimization of electrodes using the CSM," in IEEE Transactions on Magnetics, vol. 34, no. 5, pp. 2988-2991, 1998.

Camera Positioning in "Virtual" Worlds



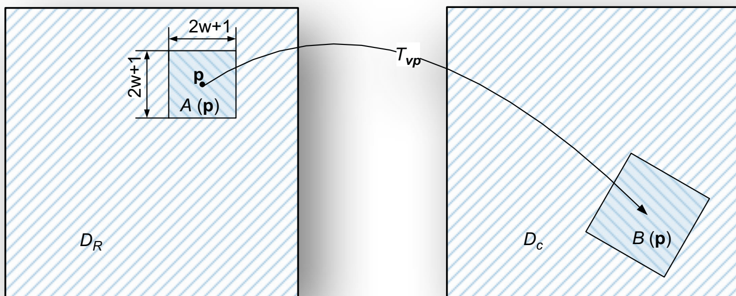


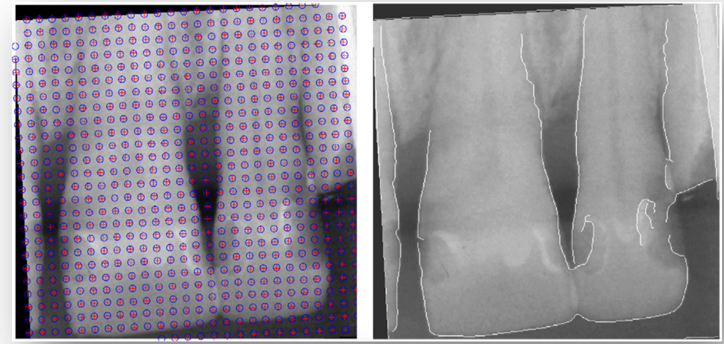
Preuss, Burelli, Yannakakis. Diversified Virtual Camera Composition. In EvoApplications 2012, pp. 265-274. Springer, 2012

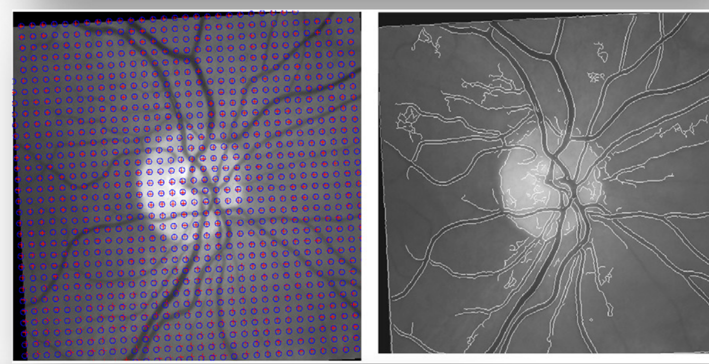
Medical Informatics

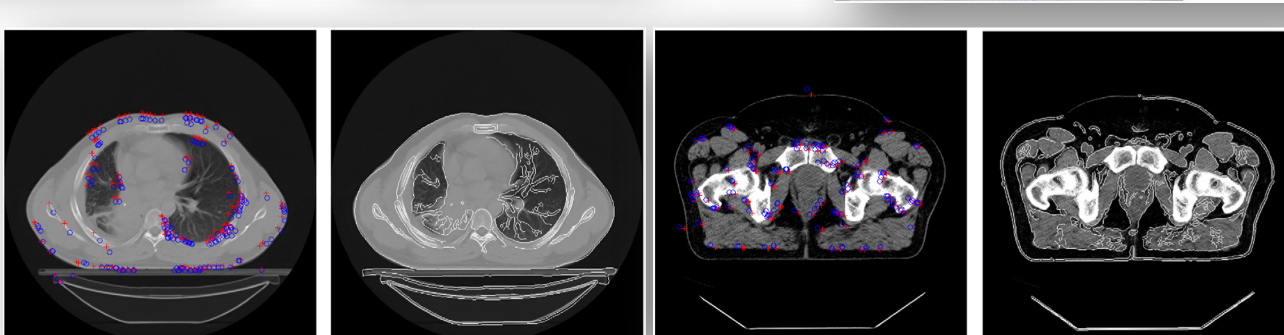
Automatic determination of point correspondence between images

Niching techniques (RTS) successfully discovered optimal solutions that are measured by the similarity between patches of the two images



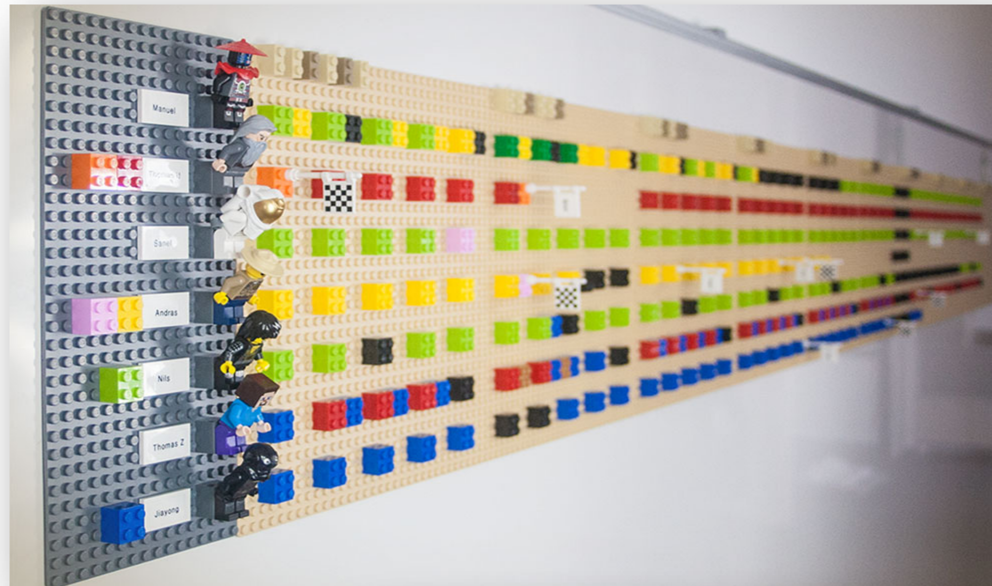






K. Delibasis, P. A. Asvestas, and G. K. Matsopoulos, “Multimodal genetic algorithms-based algorithm for automatic point correspondence,” Pattern Recognition, vol. 43, no. 12, pp. 4011-4027, 2010

Scheduling Problems

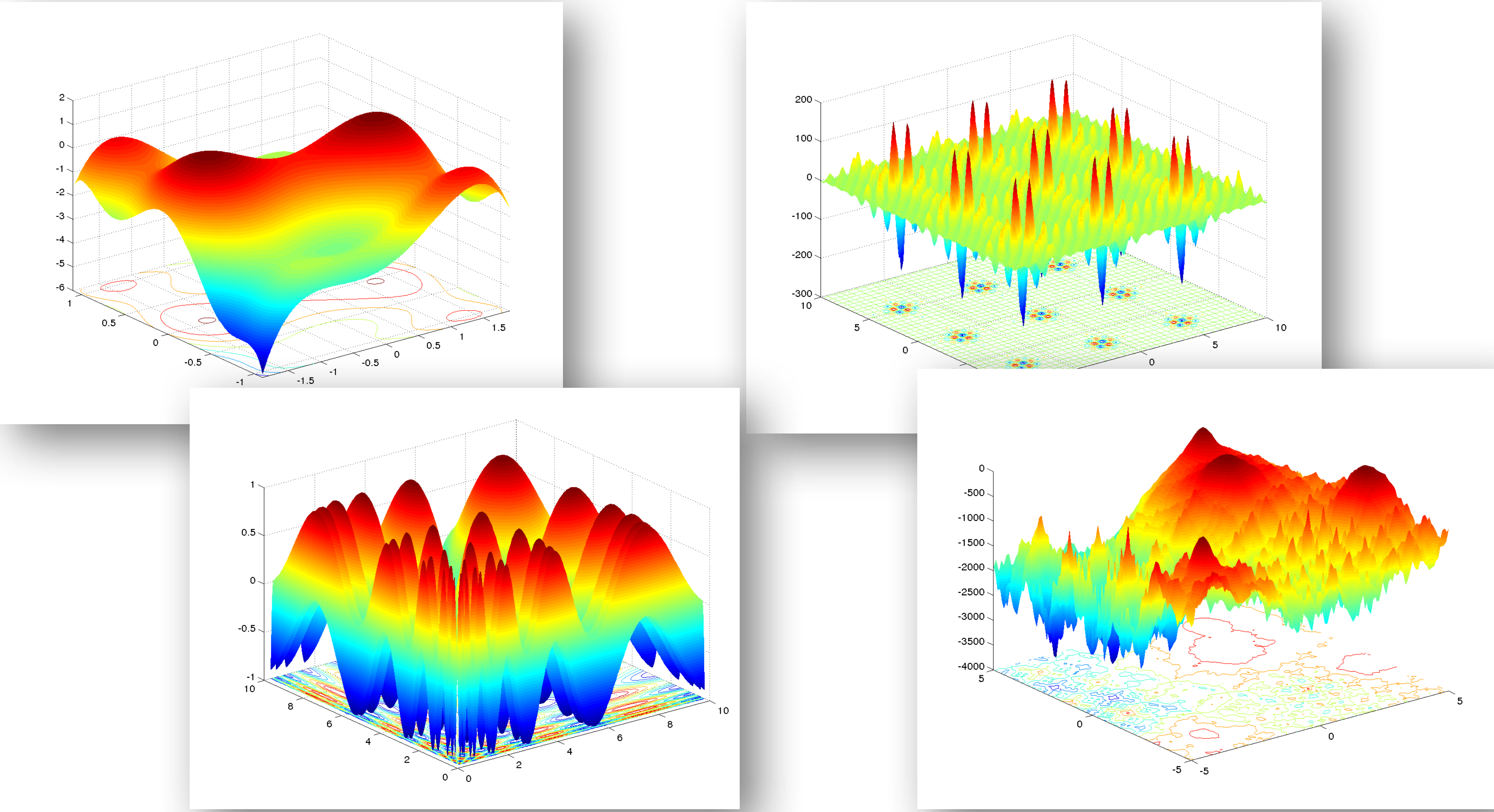


- Project Management
  - Optimize productivity
    - Makespan, Due dates
  - Maximize revenue
  - Minimize delays
- Job shop Scheduling

- Pérez, E., Posada, M. & Lorenzana, A. Taking advantage of solving the resource constrained multi-project scheduling problems using multi-modal genetic algorithms, Soft Comput (2016) 20: 1879. doi:10.1007/s00500-015-1610-z
- E. Prez, F. Herrera, and C. Hernndez, “Finding multiple solutions in job shop scheduling by niching genetic algorithms,” Journal of Intelligent Manufacturing, vol. 14, no. 3-4, pp. 323–339, 2003.
- E. Prez, M. Posada, and F. Herrera, “Analysis of new niching genetic algorithms for finding multiple solutions in the job shop scheduling,” Journal of Intelligent Manufacturing, vol. 23, no. 3, pp. 341–356, 2012.

Pictures from: http://www.ymc.ch/en/lego-resource-scheduling-wall

Artificial examples



X. Li, A. Engelbrecht, and M. Epitropakis, “Benchmark functions for cec’2013 special session and competition on niching methods for multimodal function optimization,” Technical Report, Evolutionary Computation and Machine Learning Group, RMIT University, 2013.

Multi-Modal Optimization

RESEARCH QUESTIONS

Main Research Questions

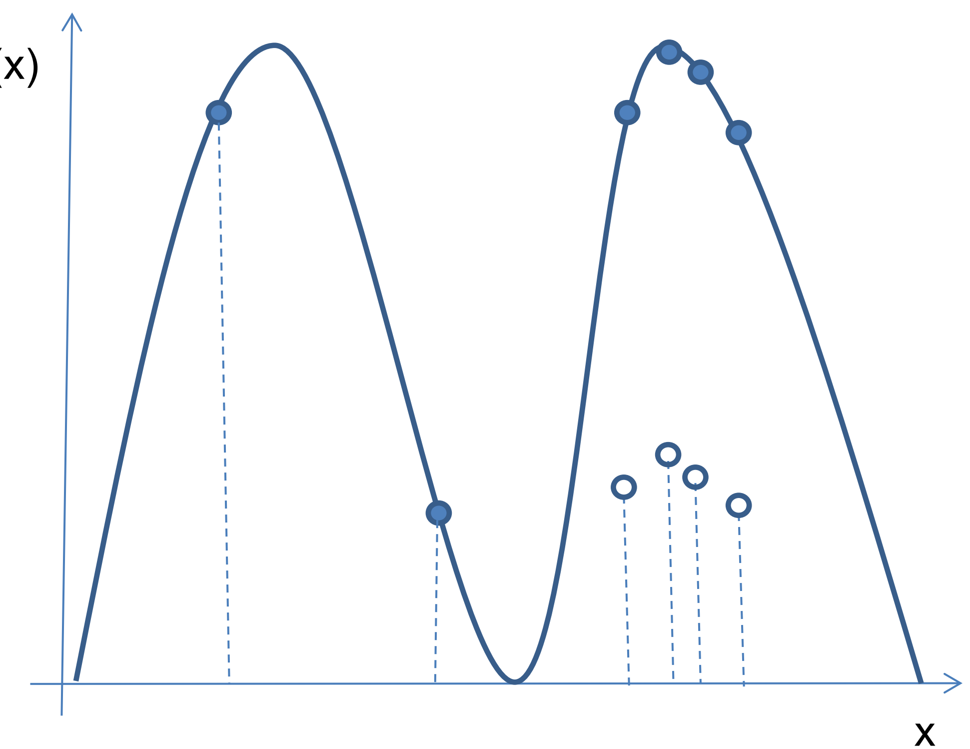
- In which situations are MMO methods actually ***better than*** "usual" EC optimization algorithms?
  - Problems (problem classes)
  - Performance measures
  - Properties, e.g. time/space complexity
- What are the ***advantages/characteristics*** of different MMO methods, which one shall we choose?
- What are the ***limits*** for further improvement?
- How can we ***rigorously*** define the field (Theoretical justifications/analyses)?

Multi-Modal Optimization

CLASSIC NICHING METHODS

Fitness sharing

- A **sharing function** can be used to degrade an individual's fitness based on the presence of other neighbouring individuals
- During selection, many **individuals in the same neighbourhood would degrade each other's fitness**
  - Limiting the number of individuals occupying the same niche



An example to illustrate fitness sharing.

D. E. Goldberg and J. Richardson, "Genetic algorithms with sharing for multimodal function optimization," in Proc. of the Second International Conference on Genetic Algorithms, J. Grefenstette, Ed., 1987, pp. 41–49.

Crowding methods

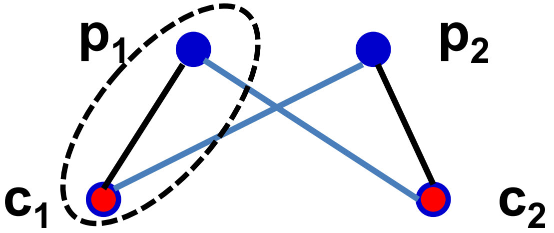
- Originally by De Jong (1975), and later modified by Mahfoud (1995)
- **Crowding** usually consists of two phases:
  - **Pairing phase**: pairing each offspring with a similar individual in the current population
  - **Replacement phase**: which of the two will remain in the population?
- **Deterministic Crowding** selects the fittest individual in each pair in the replacement phase
- **Probabilistic Crowding** selects the surviving individual for each pair based on a probabilistic formula that takes fitness into account

S. W. Mahfoud. *Niching Methods for Genetic Algorithms*. PhD thesis, Department of General Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, 1995.
K. A. de Jong. *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*. PhD thesis, Department of Computer and Communication Sciences, University of Michigan, Ann Arbor, MI, 1975.

Deterministic crowding

Algorithm 1: The pseudocode of deterministic crowding.

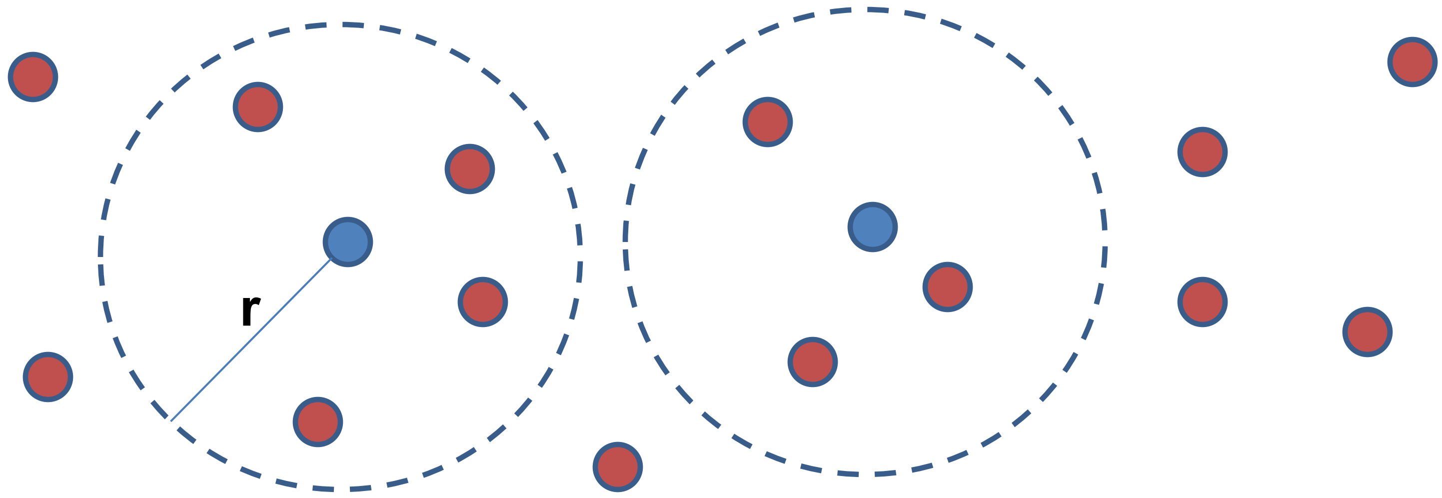
```
Select two parents, p1 and p2 randomly, without replacement
Generate two offspring c1 and c2
if d(p1, c1) + d(p2, c2) ≤ d(p1, c2) + d(p2, c1) then
    if f(c1) > f(p1) then replace p1 with c1
    if f(c2) > f(p2) then replace p2 with c2
else
    if f(c2) > f(p1) then replace p1 with c2
    if f(c1) > f(p2) then replace p2 with c1
end if
```



Each offspring tends to compete for survival with its most similar parent.

Clearing

- Proposed by **Petrowski** (1996); inspired by the principle of ***sharing of limited resources within each subpopulation*** (or species).
- The ***clearing*** procedure only ***supplies the resources*** to the ***best individuals*** in each subpopulation
- All individuals fall within **r** distance from the best **k** individuals (below shows **k** = 2) from the population are **cleared**. This process is repeated until the whole population is considered.



A. Petrowski. A clearing procedure as a niching method for genetic algorithms. In *Proceedings of Third IEEE International Conference on Evolutionary Computation*(ICEC'96), pages 798–803. Piscataway, NJ:IEEE Press, 1996.

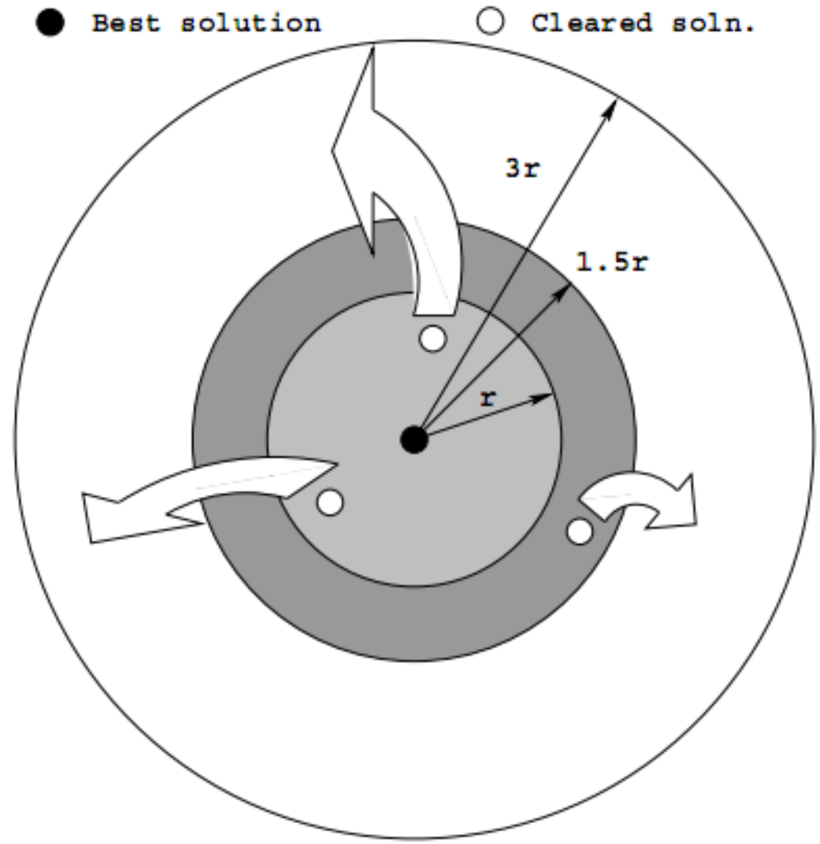
Restricted Tournament Selection

- Proposed by Harik (1997)
- A ***modification of standard tournament selection***, based on local competition
- Two individuals ***x*** and ***y*** are picked, and ***crossover and mutation*** is performed in the standard way, creating new individuals ***x'*** and ***y'***
- Then ***w*** (i.e., window size) individuals are ***randomly chosen*** from the population, and among these the ***closest*** one to ***x'***, namely ***x''***, competes with ***x'*** for a spot in the new population

G. Harick. Finding multi-modal solutions using restricted tournament selection. In *Proceedings of the Sixth International Conference on Genetic Algorithms*(ICGA-95), pages 24–31, 1997.

Other methods

- **Clustering** based methods (Yin and Germay 1991)
- **Species conserving** GA (SCGA) by Li et al. (2002)
- **Modified clearing** by Singh and Deb (2006)
- Also **sequential niching** methods, and so on



G. Singh and K. Deb, "Comparisons of multi-modal optimization algorithms based on evolutionary algorithms," in Proc. of the Genetic and Evolutionary Computation Conference 2006 (GECCO'06), Washington, USA, 2006, pp. 1305 – 1312.

Multi-Modal Optimization

NICHING WITH PSO AND DE

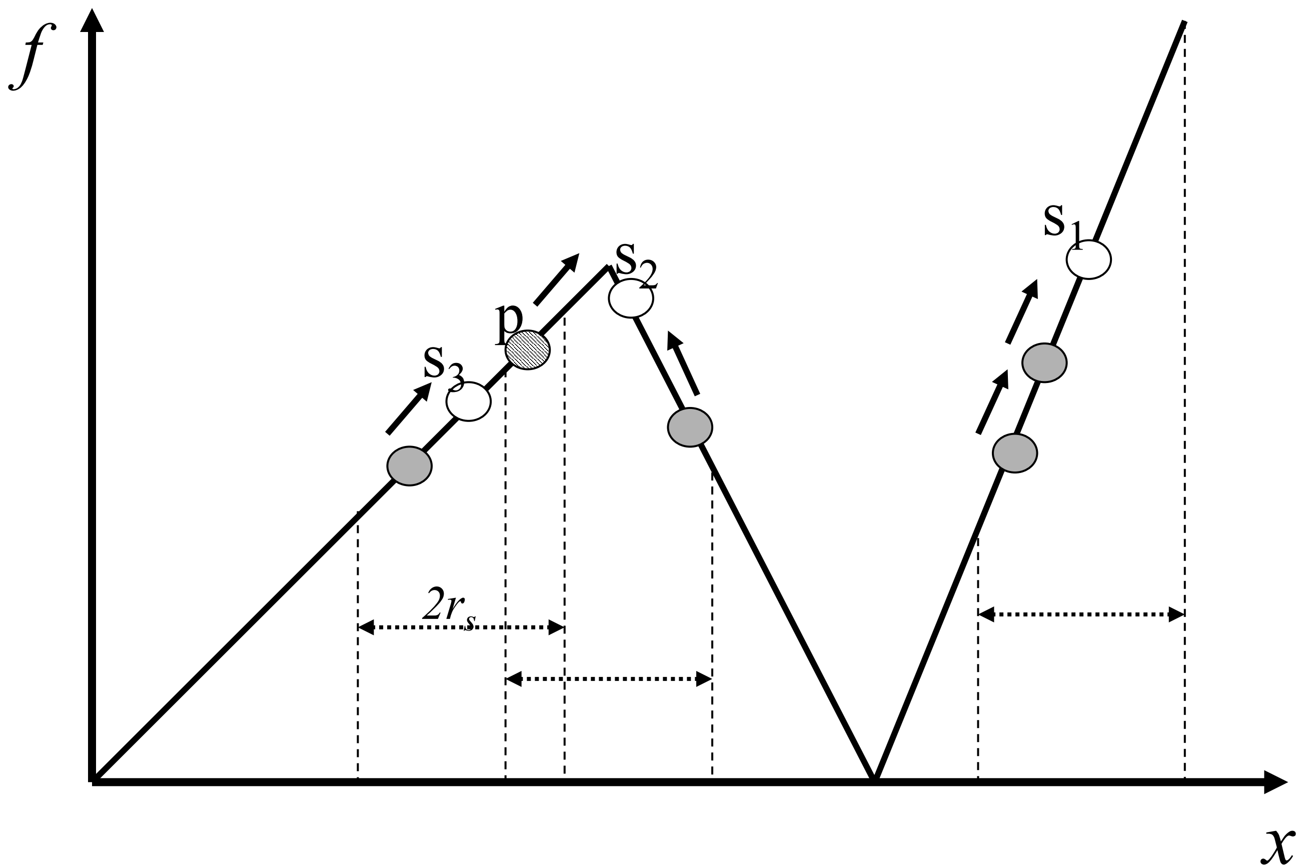
PSO niching methods

- In Particle Swarm Optimization (PSO), each particle has its ***own memory*** remembering its ***best known position*** so far, and ***share*** this information with other particles
- At each iteration, each particle is ***propelled towards*** the area defined by the ***stochastic average of its own known best*** *position and* ***the swarm best*** *position*
- The notion of memory associated with each particle is unique to PSO, and this property can be used to ***induce niching behaviour***
- A swarm can be divided into two parts:
  – an **explorer-swarm** consisting of the current particles
  – a **memory-swarm,** comprising of only best known positions of individual particles

X. Li, “Developing niching algorithms in particle swarm optimization,” in *Handbook of Swarm Intelligence*, ser. Adaptation, Learning, and Optimization, B. Panigrahi, Y. Shi, and M.-H. Lim, Eds. Springer Berlin Heidelberg, 2011, vol. 8, pp. 67–88.

Speciation-based PSO

An example of how to determine the species seeds from the population at each iteration. s_1, s_2, and s_3 are chosen as the species seeds. Note that p follows s_2



D. Parrott and X. Li, "Locating and tracking multiple dynamic optima by a particle swarm model using speciation," IEEE Trans. on Evol. Comput., vol. 10, no. 4, pp. 440–458, August 2006.

Speciation-based PSO

Step 1: **Generate** an initial population with randomly generated particles;

Step 2: **Evaluate** all particle individuals in the population;

Step 3: **Sort** all particles in descending order of their fitness values (i.e., from the best-fit to least-fit ones);

Step 4: **Determine the species** seeds for the current population;

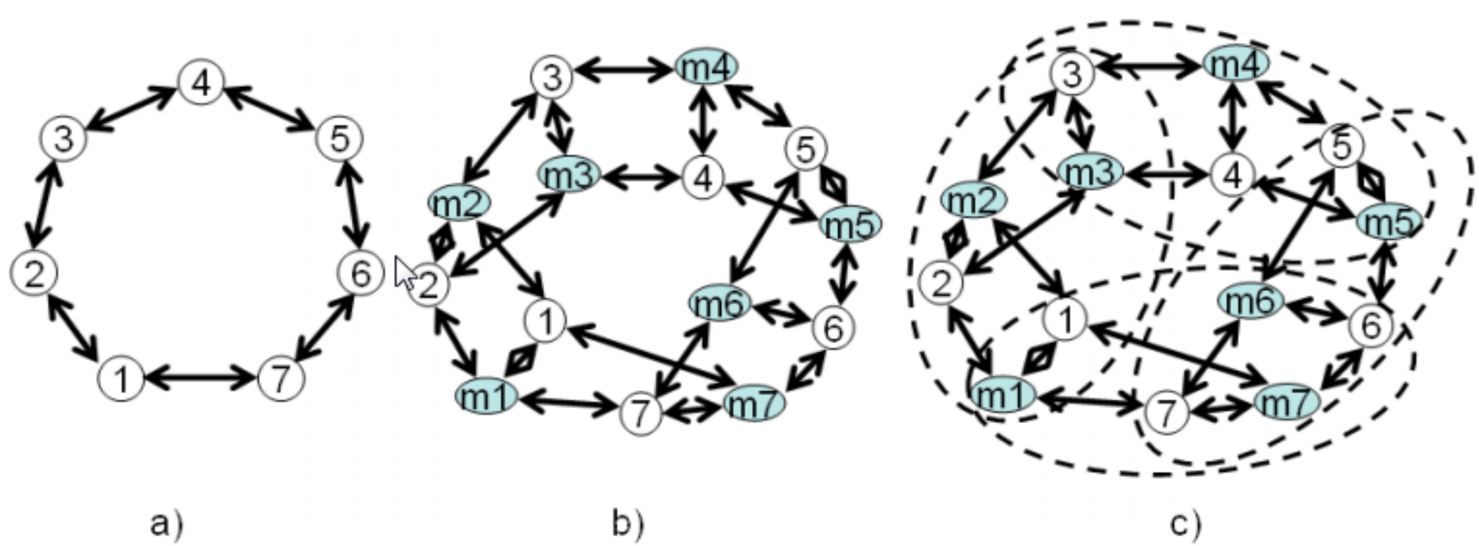
Step 5: **Assign each species seed** identified as the **gBest** to all individuals identified in the same species;

Step 6: **Adjusting particle positions** according to the PSO velocity and position update equation (1) and (2);

Step 7: Go back to step 2), unless termination condition is met.

Ring topology based niching PSO

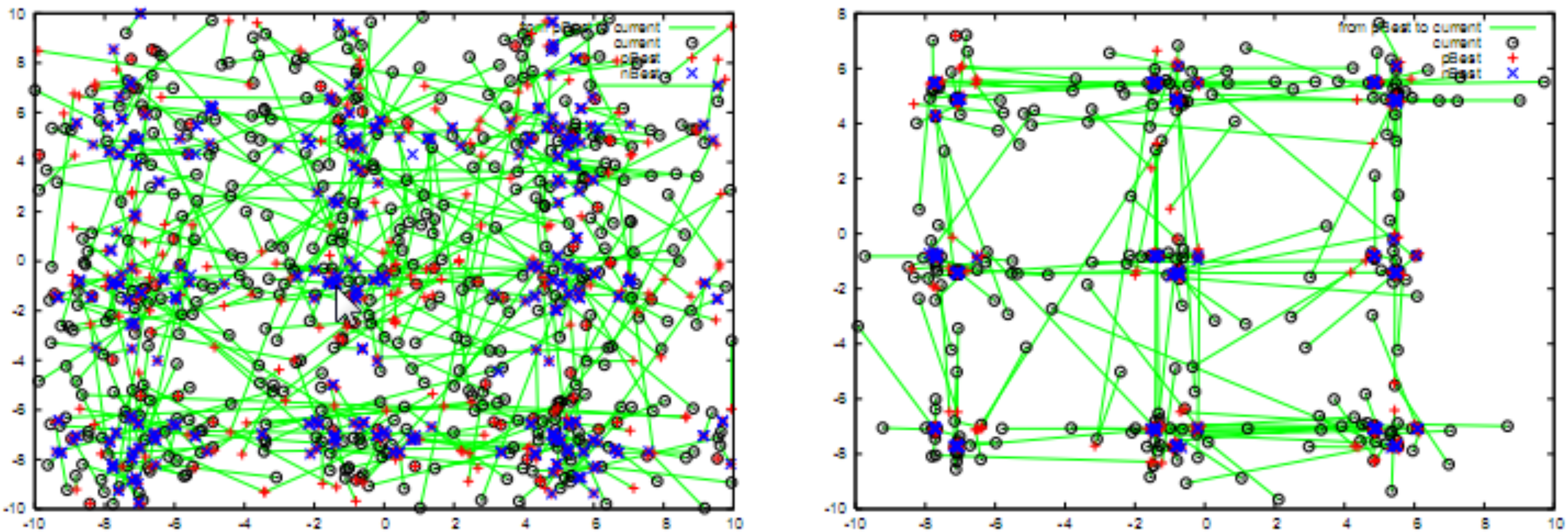
- Given a reasonably large population uniformly distributed in the search space, the ***ring topology based niching PSOs*** are able to ***form stable niches*** across different ***local neighbourhoods***, eventually locating multiple global/local optima
- This method can operate as a niching algorithm by using individual ***particles' local memories*** to form a stable network retaining the best positions found so far



X. Li, "Niching without niching parameters: Particle swarm optimization using a ring topology," *IEEE Trans. on Evol. Comput.*, vol. 14, no. 1, pp. 150 – 169, February 2010.

Ring topology based niching PSO

- Results on Shubert 2D function (two snapshots during a simulation run)



X. Li, "Niching without niching parameters: Particle swarm optimization using a ring topology," *IEEE Trans. on Evol. Comput.*, vol. 14, no. 1, pp. 150 – 169, February 2010.

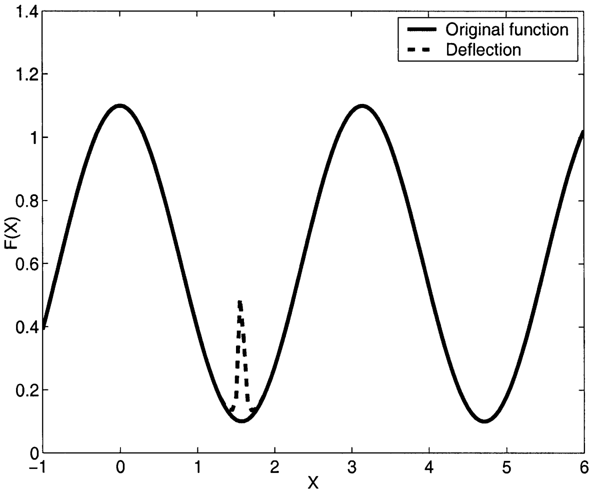
Stretching and Deflation in PSO

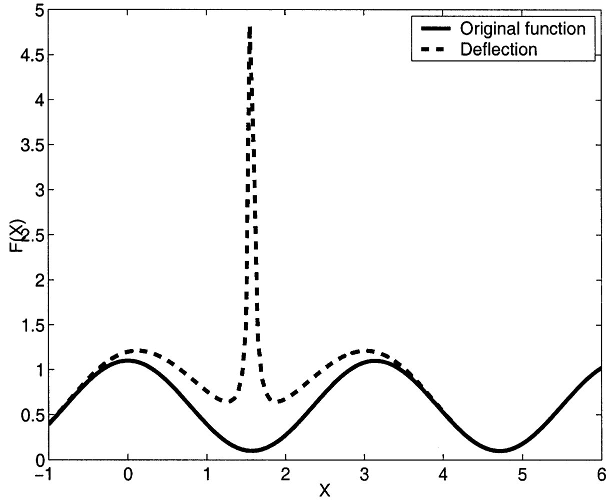
- Aims to compute **all global minimizers**, while avoiding local minimizers, through PSO
- Iteratively **modifies the objective function** by **deflection** and **stretching**
  - Knowledge of previously detected optima are incorporated in the new form
  - ***"Mexican hat" effect***: introduction of new local optima
  - Overcome such issues by using **repulsion** technique
  - Addition of **new control parameters**
  - Applications in non-linear **dynamic systems** (periodic orbits) & **game theory** (Nash equilibria)

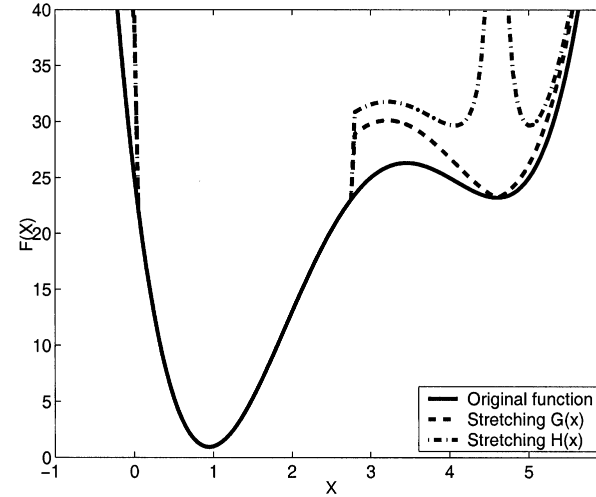
K. E. Parsopoulos, V. P. Plagianakos, G. D. Magoulas, and M. N. Vrahatis, "Objective function "stretching" to alleviate convergence to local minima," Nonlinear Analysis, vol. 47, no. 5, pp. 3419–3424, 2001.

K. E. Parsopoulos and M. N. Vrahatis, "On the computation of all global minimizers through particle swarm optimization," IEEE Trans. on Evol. Compu., vol. 8, no. 3, pp. 211–224, June 2004.

Stretching and Deflation in PSO (II)







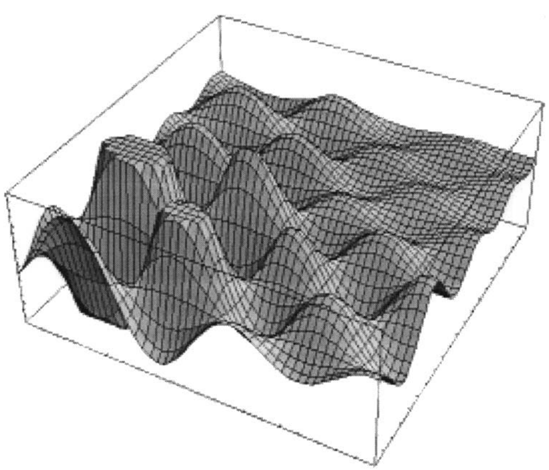


Fig. 14. Original plot of the *Levy no. 5* function in the range $[-2, 2]^2$.

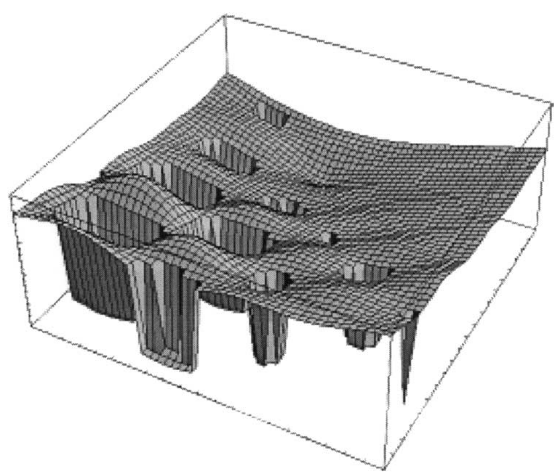


Fig. 15. First stage $G(x)$ of the stretching transformation (13) for the *Levy no. 5* function.

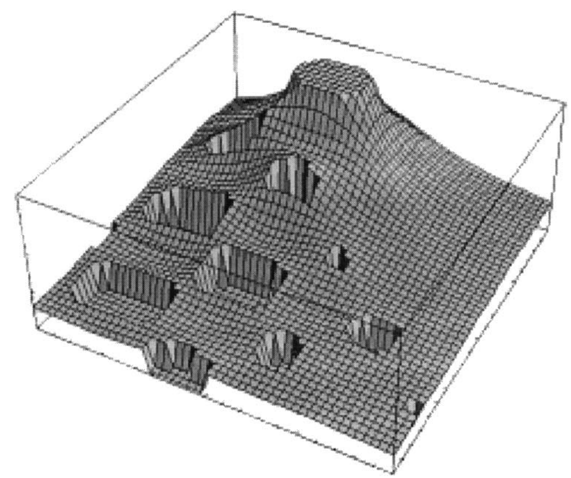


Fig. 16. *Levy no. 5* function after the stretching transformation (14).

Other recent niching PSO variants

- **LIPS**: Euclidean-distance-based niching PSO forms niches by using the *nearest neighbors* to each personal best in the Fully Informed PSO (FIPS)
  - B. Y. Qu, P. N. Suganthan, and S. Das, "A distance-based locally informed particle swarm model for multimodal optimization," *IEEE Transactions on Evolutionary Computation*, vol. 17, no. 3, pp. 387– 402, June 2013.
- ***NichePSO***, ***nbest*** **PSO**, and **Multi-swarms**
  - A. P. Engelbrecht. R. Brits and F. van den Bergh, "A niching particle swarm optimizer," *SEAL 2002*, pp. 692–696.
  - R. Brits, A. P. Engelbrecht, and F. van den Bergh, "Solving systems of unconstrained equations using particle swarm optimizers," *Proc. of the IEEE Conf. on Systems, Man, Cybernetics*, pp. 102–107, 2002.
  - T. Blackwell and J. Branke, "Multi-swarms, exclusion, and anti- convergence in dynamic environments," *Evolutionary Computation, IEEE Transactions on*, vol. 10, no. 4, pp. 459–472, 2006.
- **Adaptive Niching PSO** (ANPSO) adaptively determines the niche radius by calculating population statistics at each iteration
  - S. Bird and X. Li, "Adaptively choosing niching parameters in a PSO," in *GECCO 2006*, 2006, pp. 3–10.
- **Vector-based PSO** (VPSO) treats each particle as a vector and niche identification is done by carrying out vector operations of the particles. A niche is determined by the radius value based on the distance between the swarm best and the nearest particle with a negative dot product (i.e., moving in an opposite direction)
  - I. L. Schoeman and A. P. Engelbrecht, "Using vector operations to identify niches for particle swarm optimization," in *Proc. of the 2004 IEEE Conf. on Cybernetics and Intelligent Systems*, 2004, pp. 361 – 366.
- **Recent Developments in PSO**, please refer to:
  - J. Barrera and C. A. C. Coello, "A review of particle swarm optimization methods used for multimodal optimization," in *Innovations in Swarm Intelligence*, ser. Studies in Computational Intelligence, C. Lim, L. Jain, and S. Dehuri, Eds. Springer Berlin Heidelberg, 2009, vol. 248, pp. 9–37.
  - X.Li,"Developing niching algorithms in particle swarm optimization," in *Handbook of Swarm Intelligence*, ser. Adaptation, Learning, and Optimization, B. Panigrahi, Y. Shi, and M.-H. Lim, Eds. Springer Berlin Heidelberg, 2011, vol. 8, pp. 67–88.

Niching in Differential Evolution

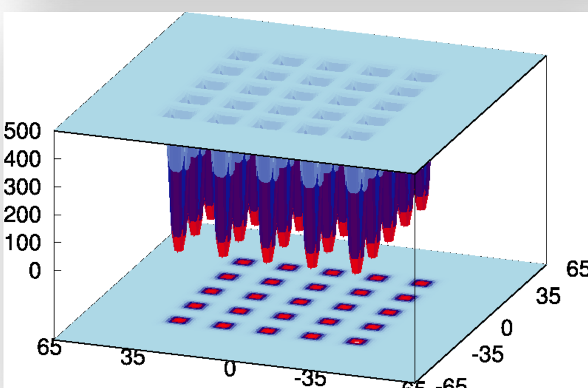
- Differential Evolution (Storn & Price 1995)
- DE belongs in the class of EAs
- Population-based, few control parameters
- Basic Operations:
  - Mutation, Crossover, Selection
- DE in MMO:
  - niching technics,
  - specialized search operators (mostly mutation str.)

Mining Differential Evolution's dynamics

Observation:

DE mutation strategies tend to distribute the individuals of the population in the vicinity of the minima of the objective function.

- Exploitative strategies: rapidly gather all the individuals to the basin of attraction of a single minimum,
- Explorative strategies: tend to spread the individuals around many minima.



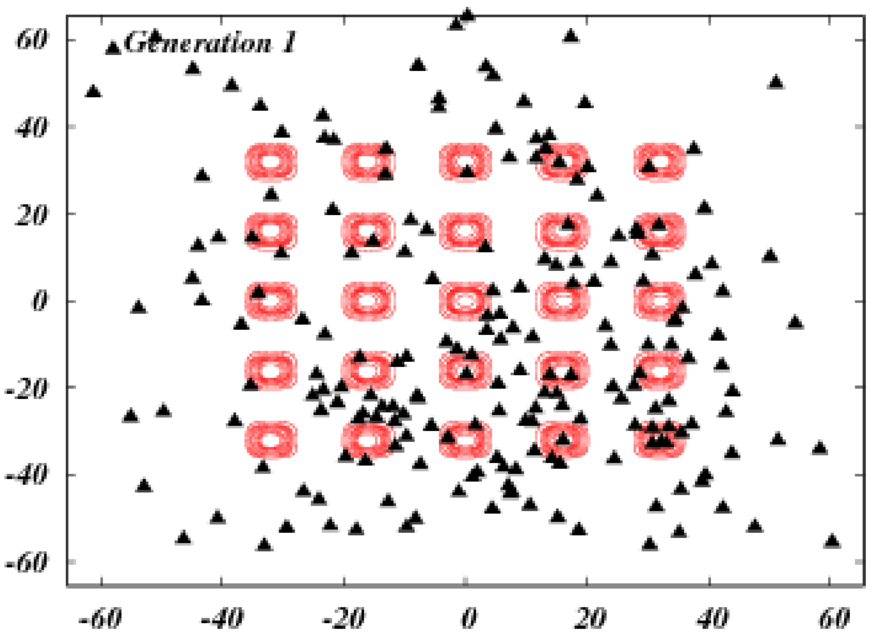
Case study: Shekel's Foxholes

- Twenty four separable local minima
- One global minimum
  - @ f(-32,32) = 0.998004

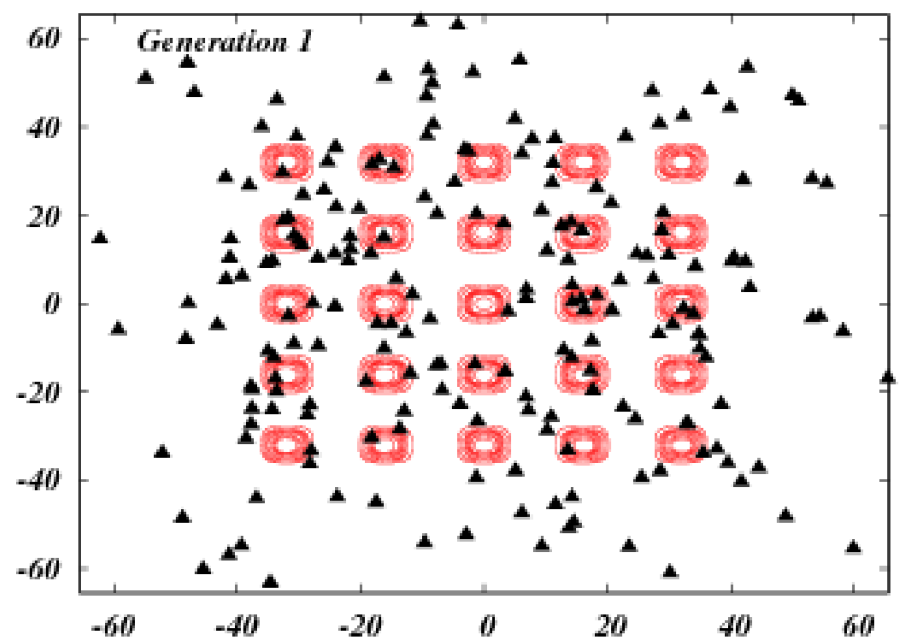
M.G. Epitropakis, D.K. Tasoulis, N.G. Pavlidis, V.P. Plagianakos, and M.N. Vrahatis, "Enhancing differential evolution utilizing proximity- based mutation operators,", IEEE Transactions on Evolutionary Computation, vol. 15, no. 1, pp. 99-119, 2011.

DE dynamics

DE/best/1 population



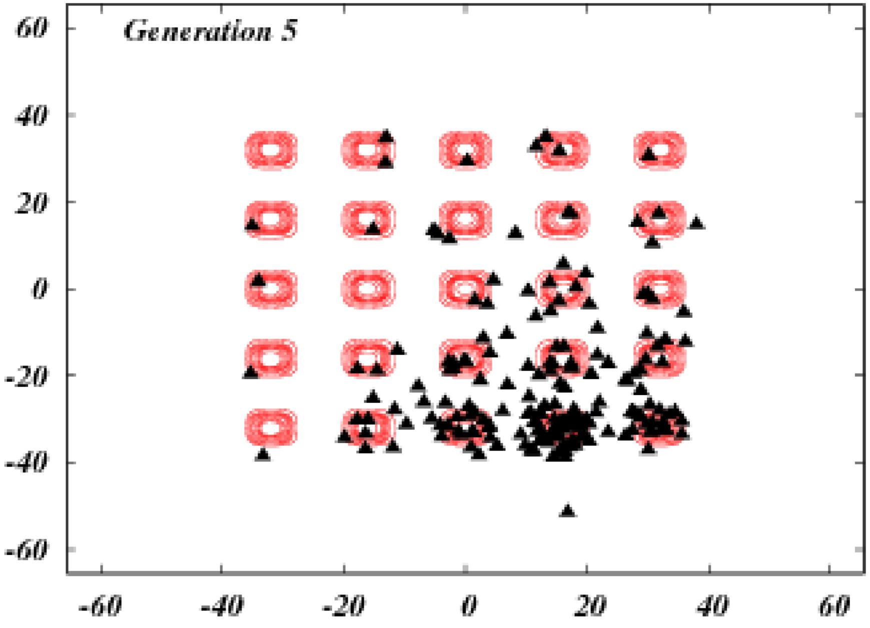
DE/rand/1 population



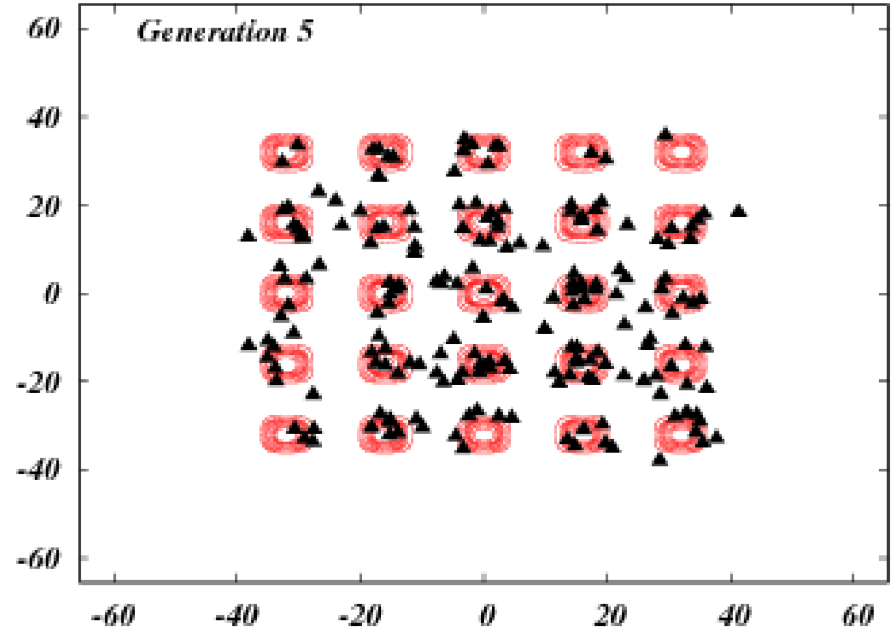
DE/best/1 and DE/rand/1 population's positions after **1** generation

DE dynamics

DE/best/1 population



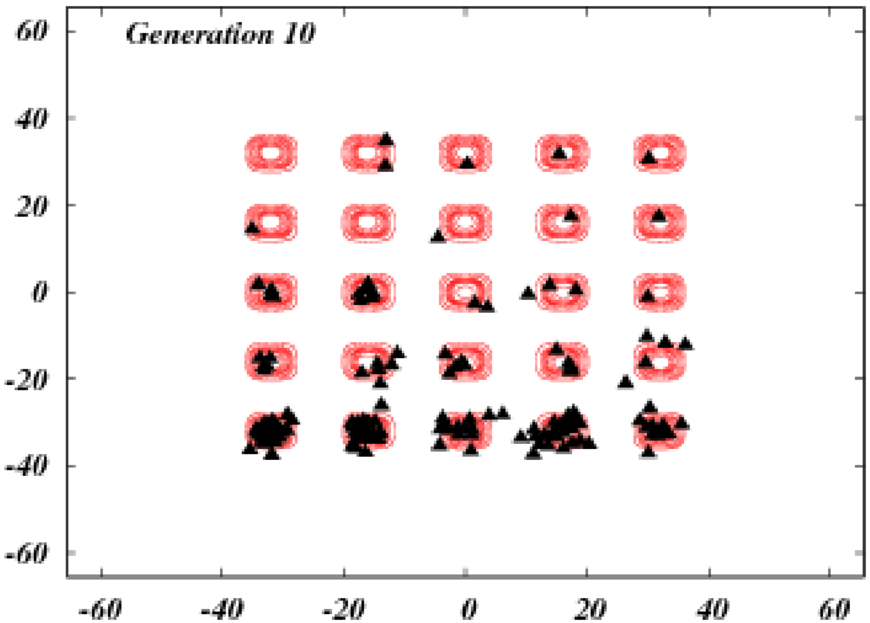
DE/rand/1 population



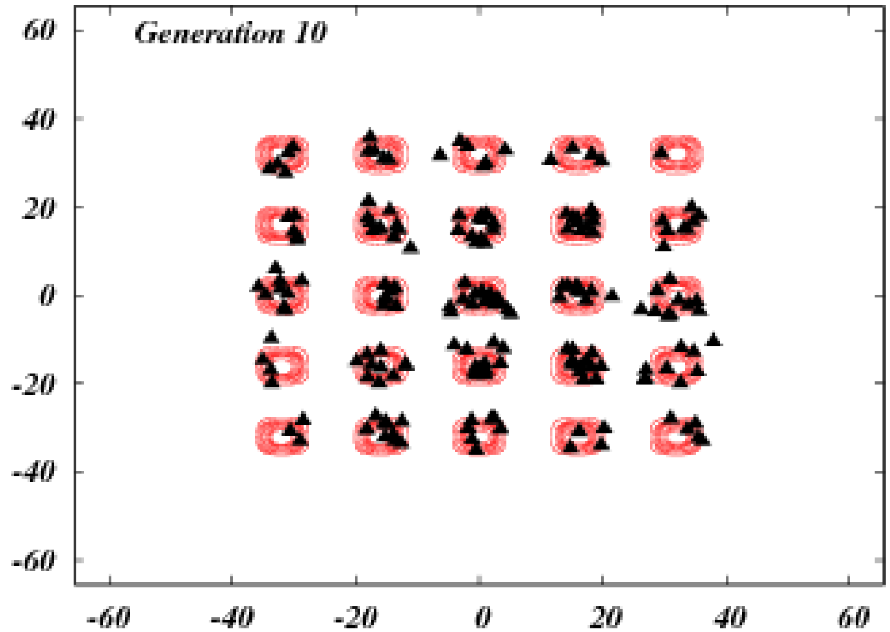
DE/best/1 and DE/rand/1 population's positions after **5** generation

DE dynamics

DE/best/1 population



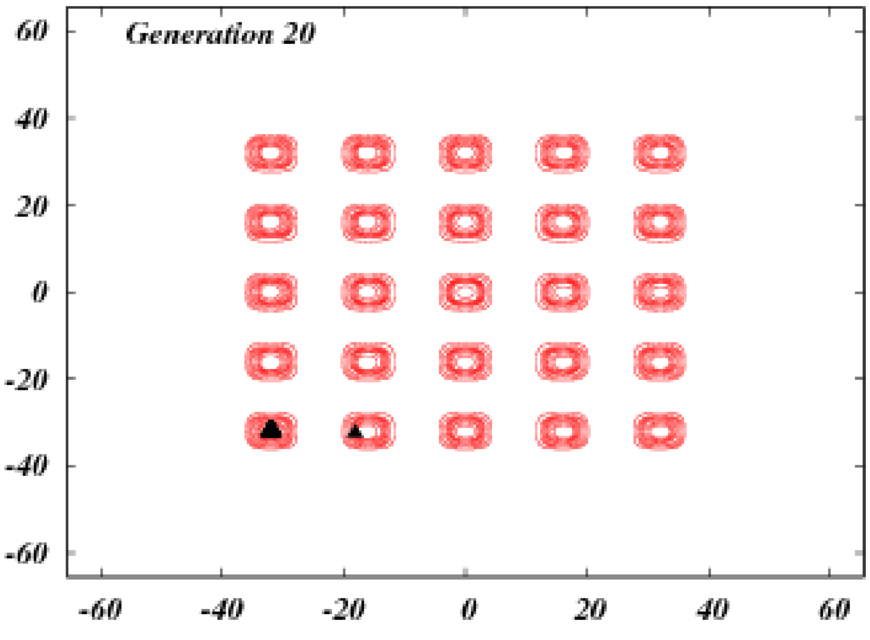
DE/rand/1 population



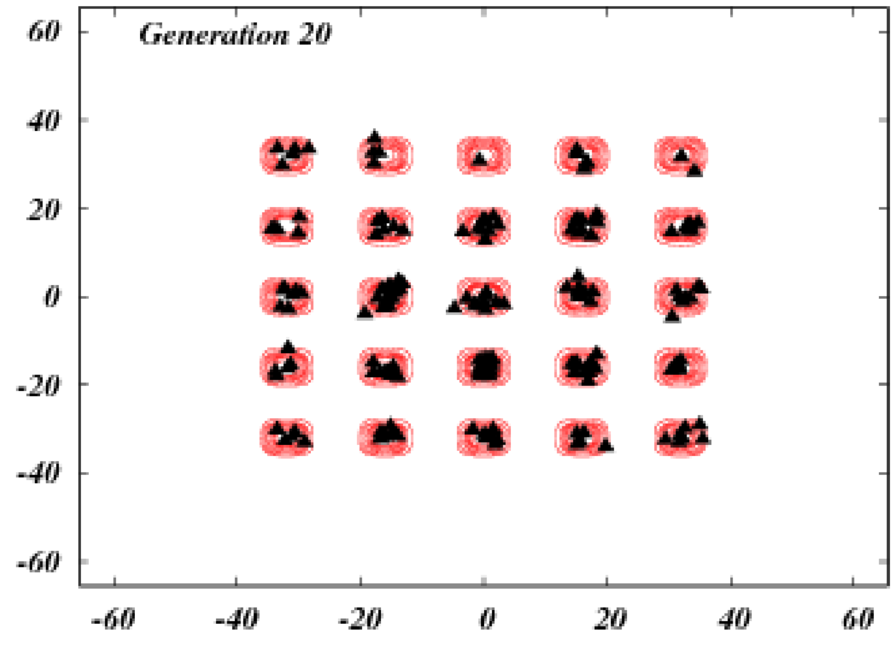
DE/best/1 and DE/rand/1 population's positions after **10** generation

DE dynamics

DE/best/1 population



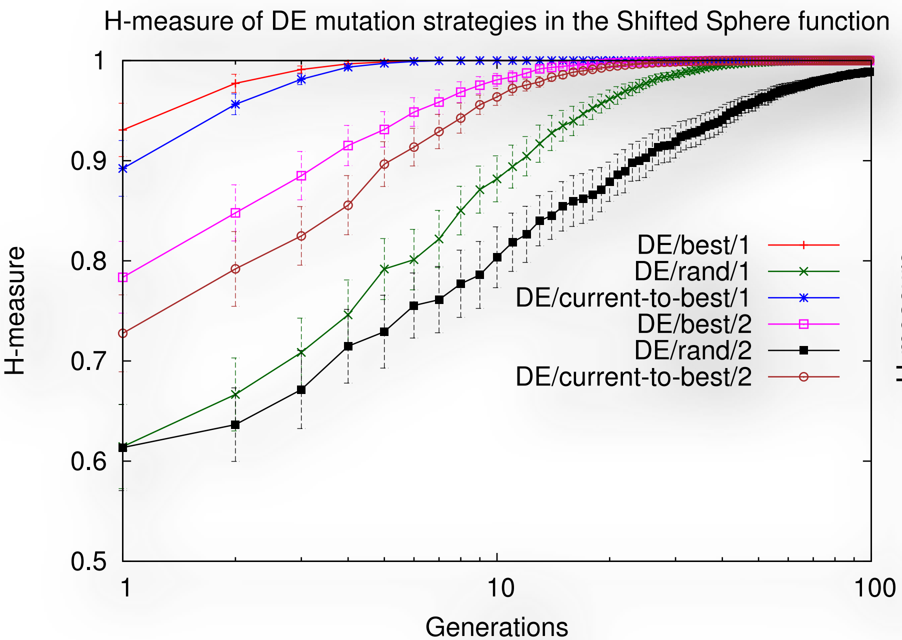
DE/rand/1 population

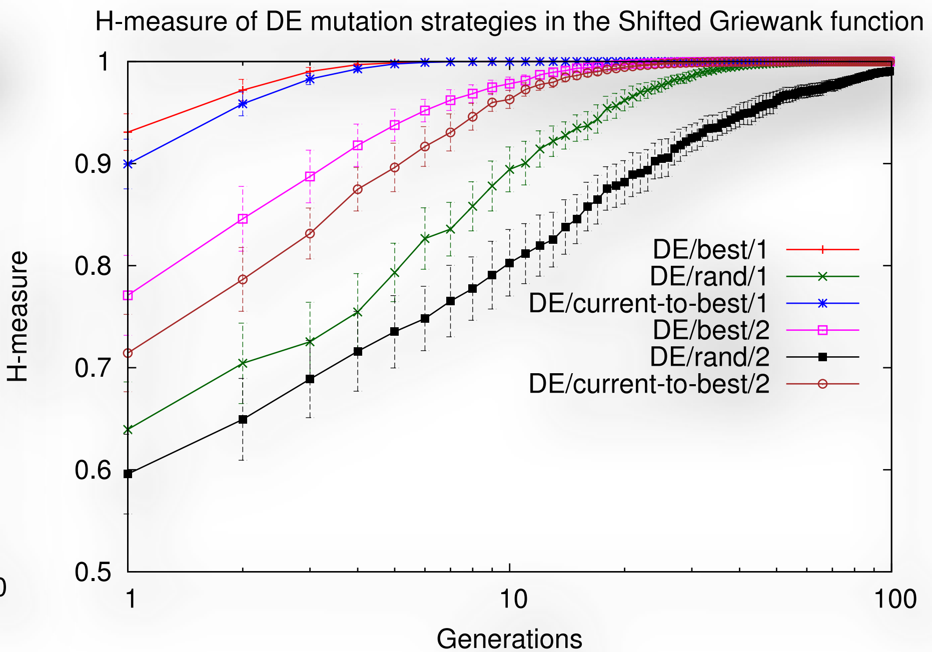


DE/best/1 and DE/rand/1 population's positions after **20** generation

DE's cluster tendency

- Cluster Tendency (H-measure, Hopkins test)
  - Determines the presence or absence of a clustering structure in a data set





B. Hopkins and J. G. Skellam, "A new method for determining the type of distribution of plant individuals," Ann. Botany, vol. 18, no. 2, pp. 213–227, 1954.
M.G.Epitropakis, D.K.Tasoulis, N.G.Pavlidis, V.P.Plagianakos, and M. N. Vrahatis, "Enhancing differential evolution utilizing proximity- based mutation operators," *Evolutionary Computation, IEEE Transactions on*, vol. 15, no. 1, pp. 99–119, Feb 2011.

Niching DE: DE/nrand family

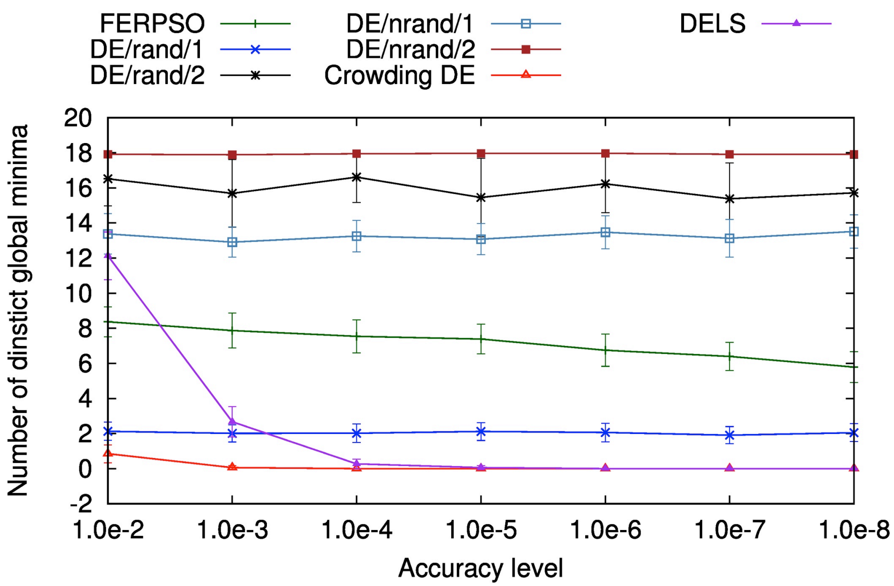
- Inspired by this observation, **classic DE mutation operators** were altered to ***incorporate spatial information*** about the ***nearest neighbour*** concept
- **Induce the niching effect**, **without** using any **additional parameter**
- Instead of using the **base vector** the usual way, its **nearest neighbour** is always chosen as the actual base vector

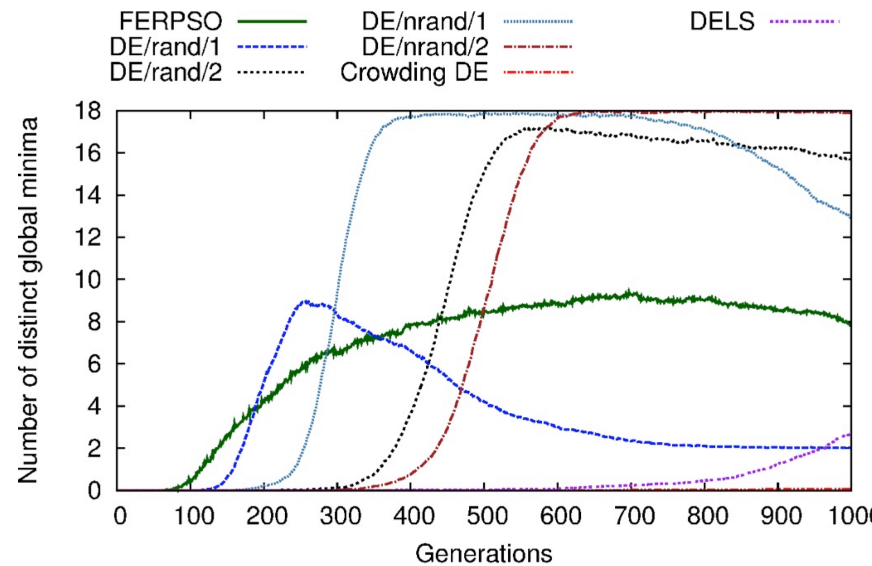
1. **DE/nrand/1:** $v_{g+1}^{i} = x_{g}^{NN_i} + F(x_{g}^{r_1} - x_{g}^{r_2}),$
2. **DE/nrand/2:** $v_{g+1}^{i} = x_{g}^{NN_i} + F(x_{g}^{r_1} - x_{g}^{r_2}) + F(x_{g}^{r_3} - x_{g}^{r_4}),$

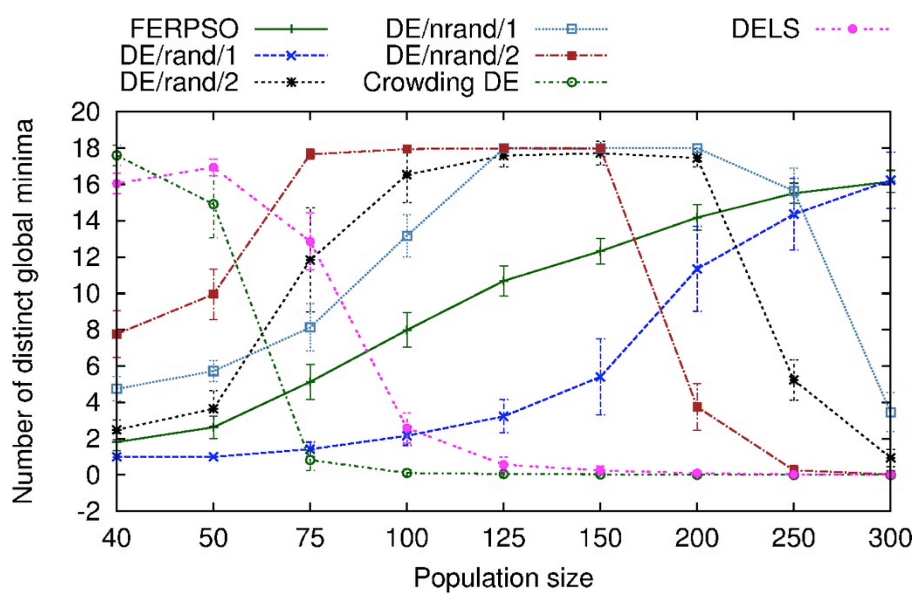
$x_{g}^{NN_i}$ is the nearest neighbor of the current individual x_{g}^{i}.

M. G. Epitropakis, V. P. Plagianakos, and M. N. Vrahatis, , “Finding multiple global optima exploiting differential evolution’s niching capability,” in *Differential Evolution (SDE), 2011 IEEE Symposium on*, April 2011, pp. 1–8.

DE/nrand family behavior







Multi-Modal Optimization:

https://mikeagn.github.io/DeMatDEnrand/

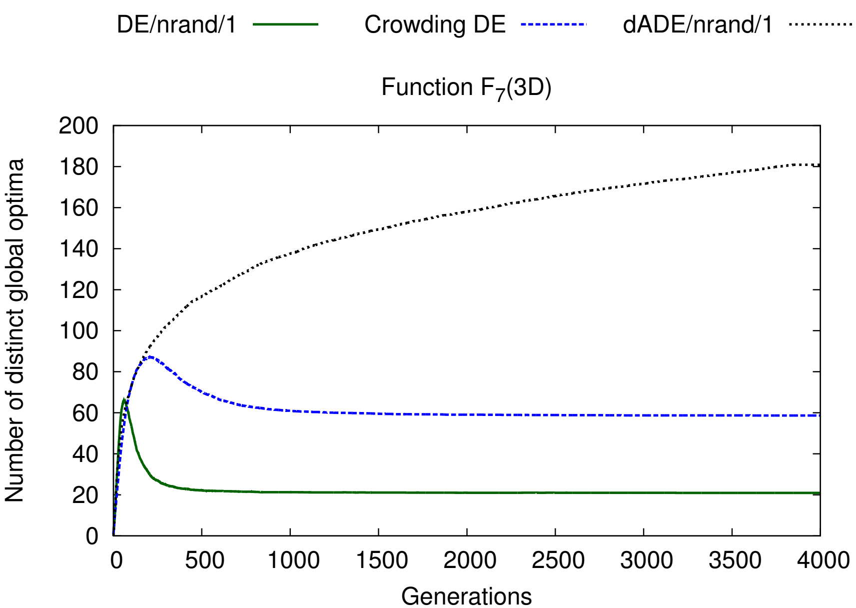
DE/NRAND/1 DEMO

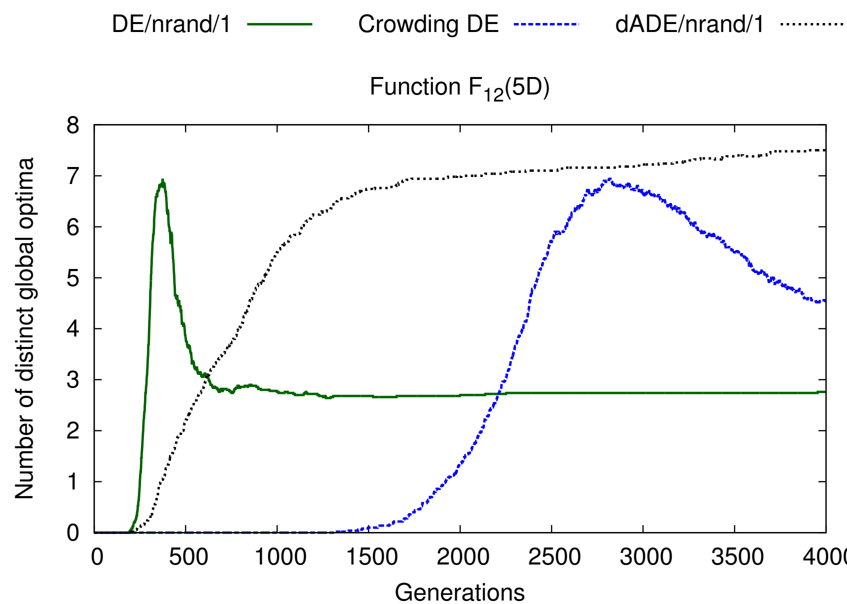
Niching DE: dADE/nrand family

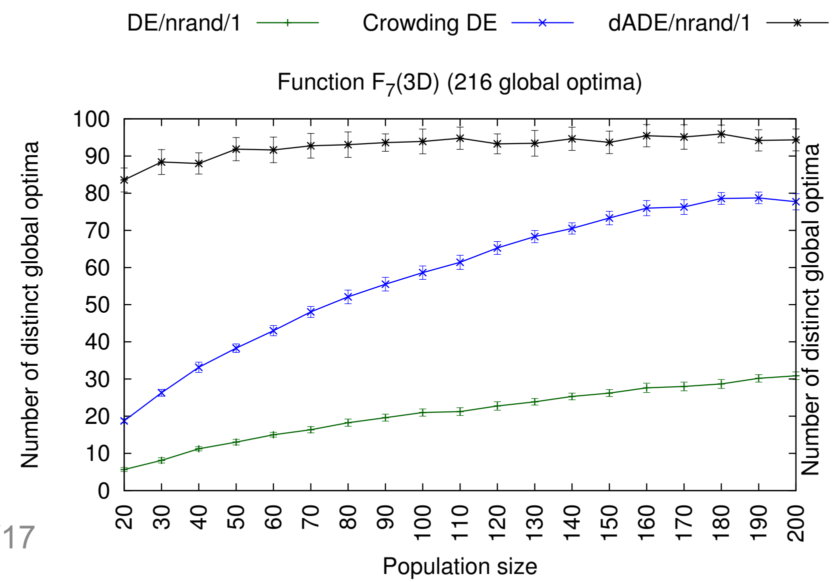
- Similar search as **DE/nrand** family
- **Self-adaptive control parameters** (JADE self-adaptation)
- Utilize **dynamic archive**:
  - put only better solutions in
  - if near better contained, re-initialize individual
  - identification radius R adapted during run
- Substantially **improved performance**
- **Less sensitive to the population size**
- Top participant in CEC 2013/2015 competitions

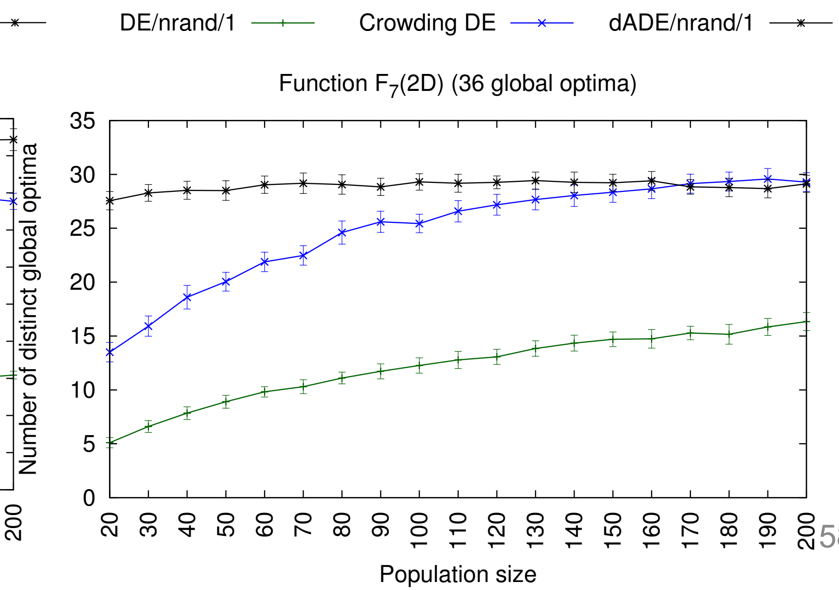
M. G. Epitropakis, X. Li, and E. K. Burke, “A dynamic archive niching differential evolution algorithm for multimodal optimization,” in Evolutionary Computation (CEC), 2013 IEEE Congress on, June 2013, pp. 79–86.

Niching DE: dADE/nrand family









Niching DE: Other recent variants

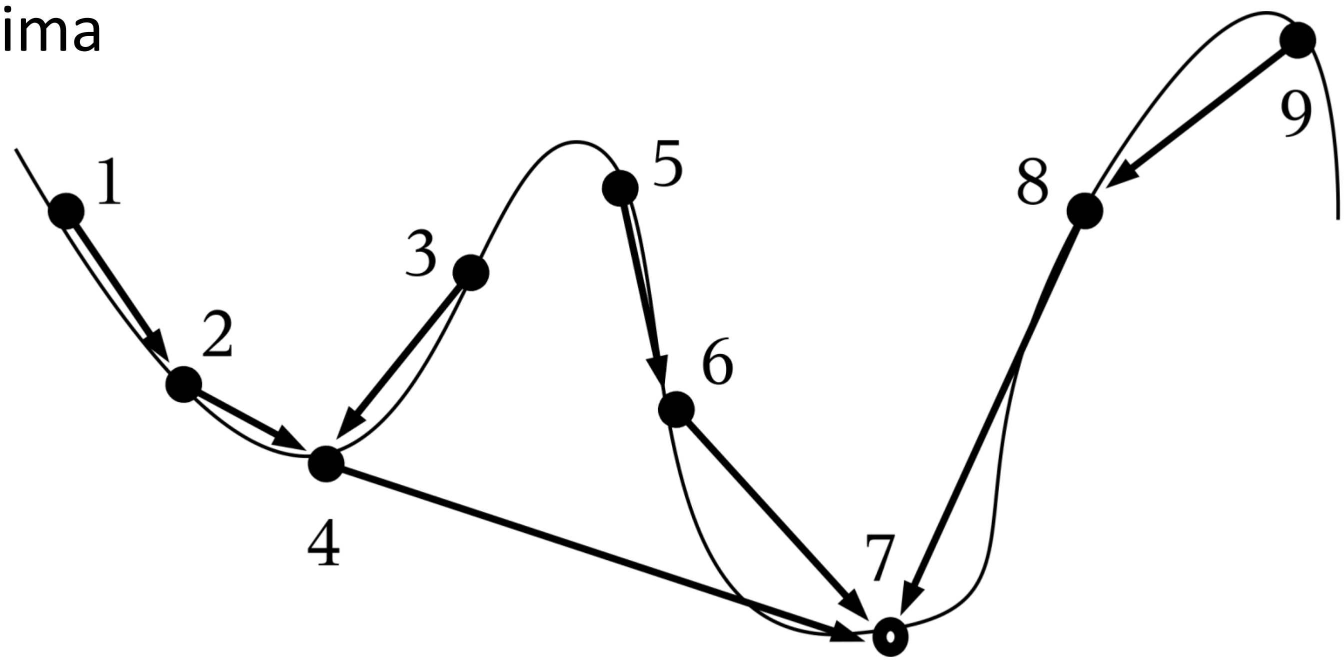
- **DE with restricted neighborhood mutations** (speciation, crowding, sharing): each individual is mutated by randomly selecting individuals within the m-th neighborhood niche of its base vector
  – B. Y. Qu, P. N. Suganthan, and J. J. Liang, "Differential evolution with neighborhood mutation for multimodal optimization," *Evolutionary Computation, IEEE Transactions on*, vol. 16, no. 5, pp. 601–614, Oct 2012.
- **DE with probabilistic parent selection** scheme based on ***fitness and proximity*** information
  – S. Biswas, S. Kundu, and S. Das, "Inducing niching behavior in differential evolution through local information sharing," *Evolutionary Computation, IEEE Transactions on*, vol. 19, no. 2, pp. 246–263, April 2015.
- **DE with parent centric mutation strategies** combined with crowding
  – S. Biswas, S. Kundu, and S. Das, "An improved parent-centric mutation with normalized neighborhoods for inducing niching behavior in differential evolution," *IEEE Transactions on Cybernetics*, vol. 44, no. 10, pp. 1726–1737, Oct 2014.
- **Ensemble of niching techniques**
  – S. Hui and P. N. Suganthan, "Ensemble and arithmetic recombination- based speciation differential evolution for multimodal optimization," *IEEE Transactions on Cybernetics*, vol. 46, no. 1, pp. 64–74, Jan 2016
- **DE with index-based neighborhoods** to induce the niching effect
  – M. G. Epitropakis, V. P. Plagianakos, and M. N. Vrahatis, "Multi- modal optimization using niching differential evolution with index- based neighborhoods," in *Proceedings of 2012 IEEE Congress on Evolutionary Computation (CEC'12)*, June 2012, pp. 1–8.

Multi-Modal Optimization

SOME OTHER STATE-OF-THE-ART NICHING METHODS

Nearest-better Clustering (I)

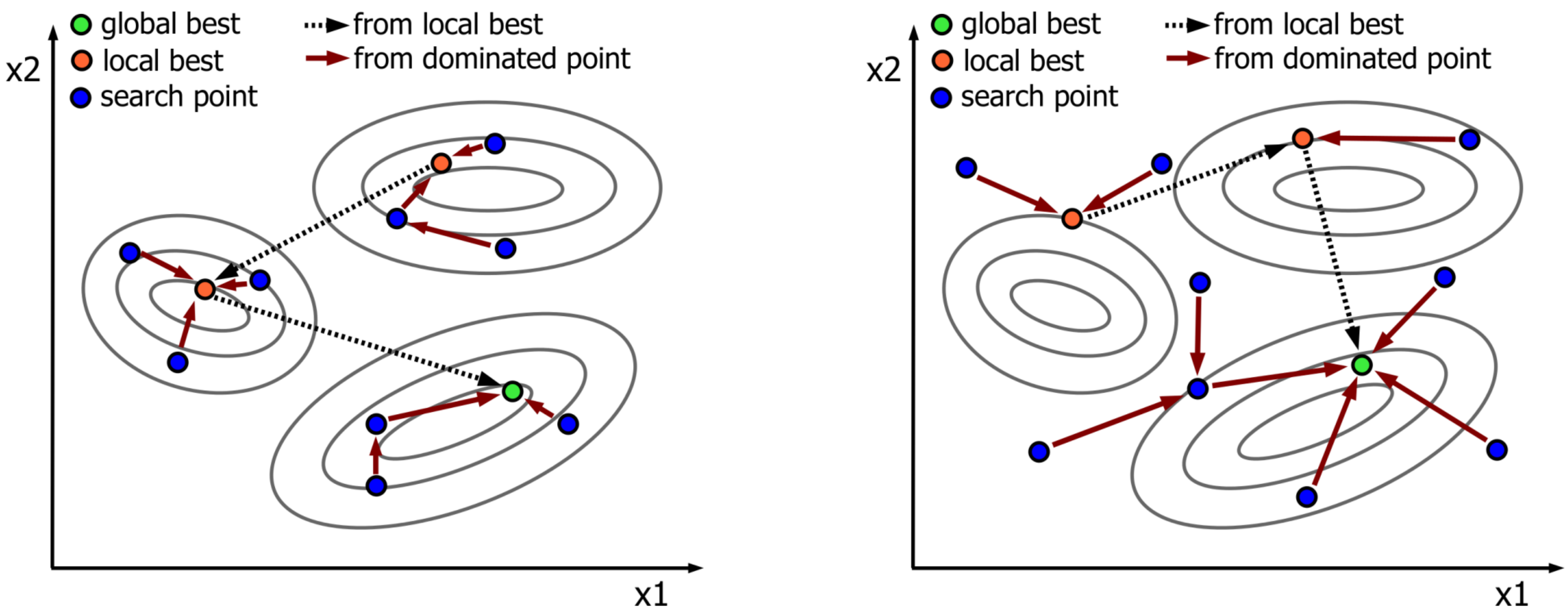
- **The basic idea:** Connect every solution to the nearest one that is better (in terms of fitness), clustering is done via cutting the longest lines
- ***Assumption***: Longest edges are connections between optima



M. Preuss. "Niching the CMA-ES via nearest-better clustering." In *Proceedings of the 12th annual conference companion on Genetic and evolutionary computation* (GECCO '10). ACM, New York, NY, USA, pp. 1711-1718, 2010

Nearest-better Clustering (II)

- NBC works with clustered (left) and randomized (right) samples
- It incorporates (&needs) heuristic rule to remove "the right" longest edges



M. Preuss. "Niching the CMA-ES via nearest-better clustering." In *Proceedings of the 12th annual conference companion on Genetic and evolutionary computation* (GECCO '10). ACM, New York, NY, USA, pp. 1711-1718, 2010

NEA2: Niching Evolutionary Algorithm 2

Algorithm 1: NEA2

1 distribute an evenly spread sample over the search space;
2 apply NBC: separate sample into populations according to clusters;
3 **forall the** *populations* **do**
4 	run local optimization (e.g. CMA-ES) until stop criterion is hit;
	`// start all over:`
5 **if** *!termination* **then**
6 	goto step 1

- **NEA2**: **clustering** + **local optimization**
- **NBC combined with CMA-ES** produces a niching algorithm that ***won*** the top place in the CEC'2013 niching competition
- However, it still needs to set a few niching parameters

M. Preuss. "Niching the CMA-ES via nearest-better clustering." In *Proceedings of the 12th annual conference companion on Genetic and evolutionary computation* (GECCO '10). ACM, New York, NY, USA, pp. 1711-1718, 2010

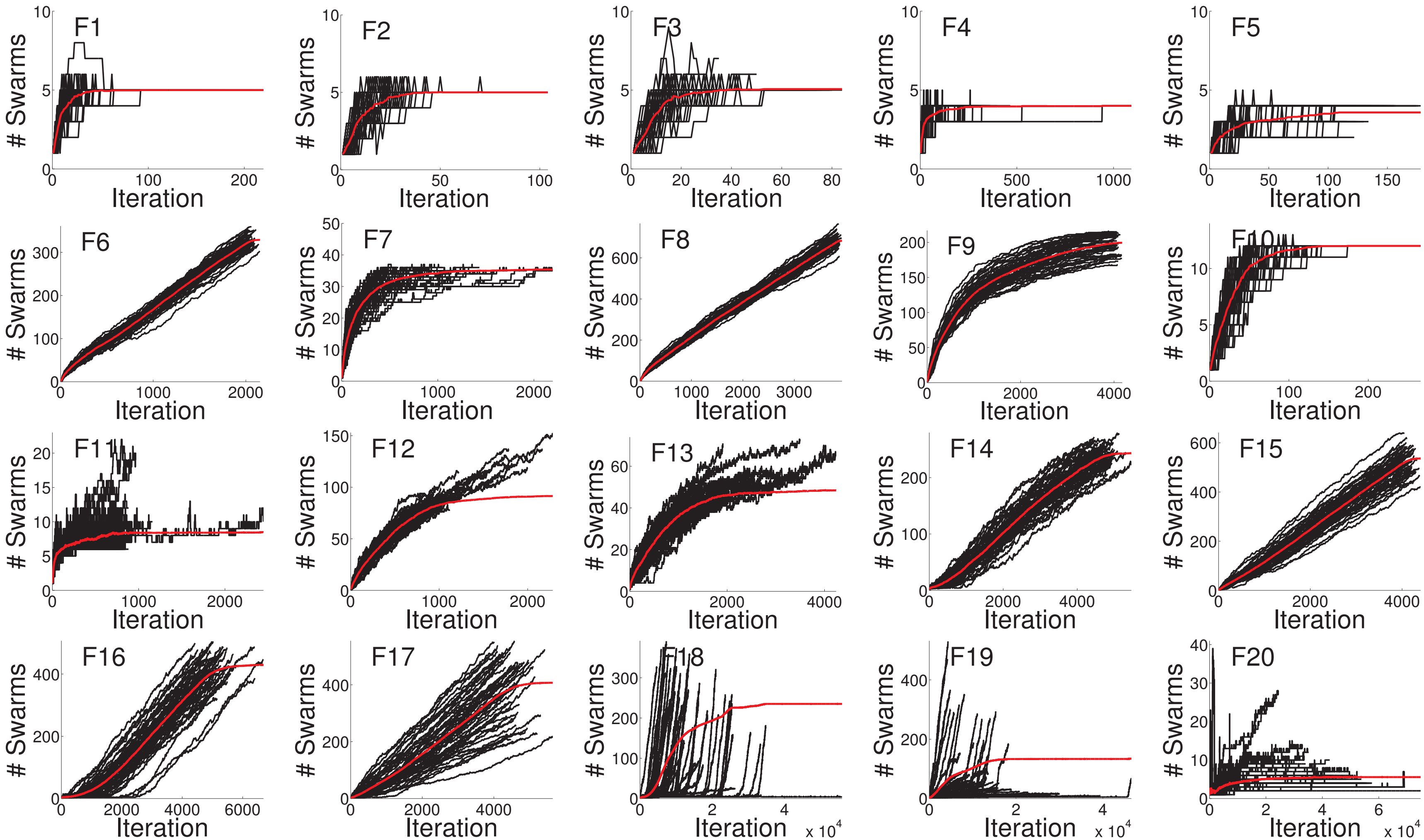
Niching Migratory Multi-swarm Optimiser (NMMSO)

- Built on an ***analysis of top ranked algorithms*** in CEC'2013 niching competition to exploit similar characteristics of the winners
  - ✓ ***self-adaption*** of search parameters
  - ✓ ***dynamic mode maintenance***
  - ✓ exploitative ***local search***

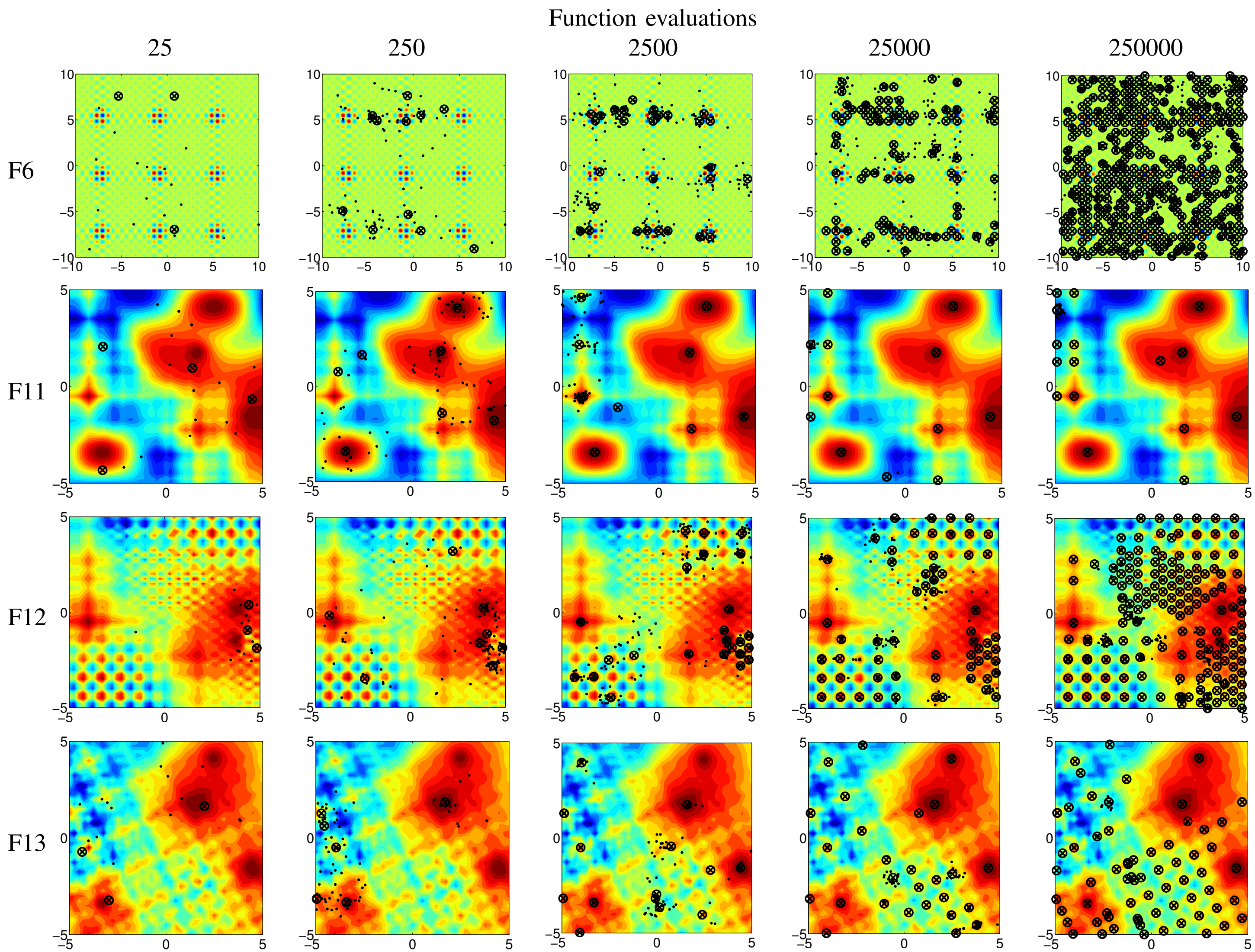
- **The basic idea**:
  - uses ***concurrent swarms*** each having ***strong local search***
  - each ***swarm fine-tunes its local mode*** estimates
  - ***swarms*** which have improved their mode/niche estimate are paired with their closest adjacent swarm for ***potential merging*** (preventing duplication of labour)
  - New regions in which to search for modes are identified by ***splitting*** away particles from ***existing*** (large) ***swarms***

J. E. Fieldsend, "Running Up Those Hills: Multi-Modal Search with the Niching Migratory Multi-Swarm Optimiser," in IEEE Congress on Evolutionary Computation, 2014, pp. 2593 - 2600.

NMMSO #swarms maintained



NMMSO distribution of swarms



RS-CMSA-ES: Covariance Matrix Self Adaption ES with Repelling Subpopulations

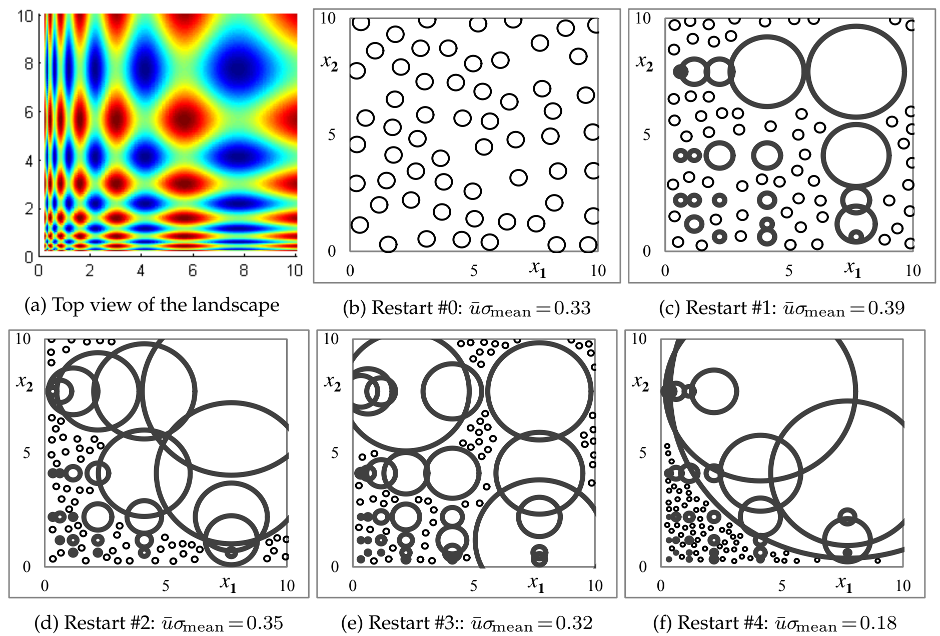
- RS-CMSA **adapts** and **reformulates** multiple **existing concepts** from different methods
- RS-CMSA can **learn the relative size** and possibly the **shape of the basins**, and properly handle the challenge of non-circular basins
- **No assumption** on the **distribution of global minima**, the **shape** or the **size of the basins** is required.
- **No niche radius** parameters and for all other parameters, the default values are shown to be robust (given a rough estimate for the expected number of minima)

Multimodal Optimization by Covariance Matrix Self-Adaptation Evolution Strategy with Repelling Subpopulations, Ali Ahrari, Kalyanmoy Deb and Mike Preuss, doi: 10.1162/EVCO_a_00182

RS-CMSA-ES: Main Components (I)

- **Sub-populations** to capture different niches
- **Archive with adaptive Taboo regions**

- **Taboo regions** are adaptively re-sized and reshaped, differ for different subpopulations.
- The size of the taboo region is adapted to the basin size.
- Taboo regions are the previously identified basins and the center of fitter subpopulations.
- Only critical taboo regions are checked.



Multimodal Optimization by Covariance Matrix Self-Adaptation Evolution Strategy with Repelling Subpopulations, Ali Ahrari, Kalyanmoy Deb and Mike Preuss, doi: 10.1162/EVCO_a_00182

RS-CMSA-ES: Main Components (II)

- Core algorithm: **Adapted CMSA-ES with elitism** (Beyer and Sendhoff, 2008)
- A **restart mechanism** with dependent restarts
- **Hill-valley function** (Ursem, 1999): checks whether two solutions share the same basin.
- **Semi-random Initialization:** maximize the initial distance between subpopulations and the archive
- **Very promising performance: Winner** of the **GECCO 2017 competition on niching methods.**

Multimodal Optimization by Covariance Matrix Self-Adaptation Evolution Strategy with Repelling Subpopulations, Ali Ahrari, Kalyanmoy Deb and Mike Preuss, doi: 10.1162/EVCO_a_00182

RLSIS: Restarted Local Search with Improved Selection of Starting Points

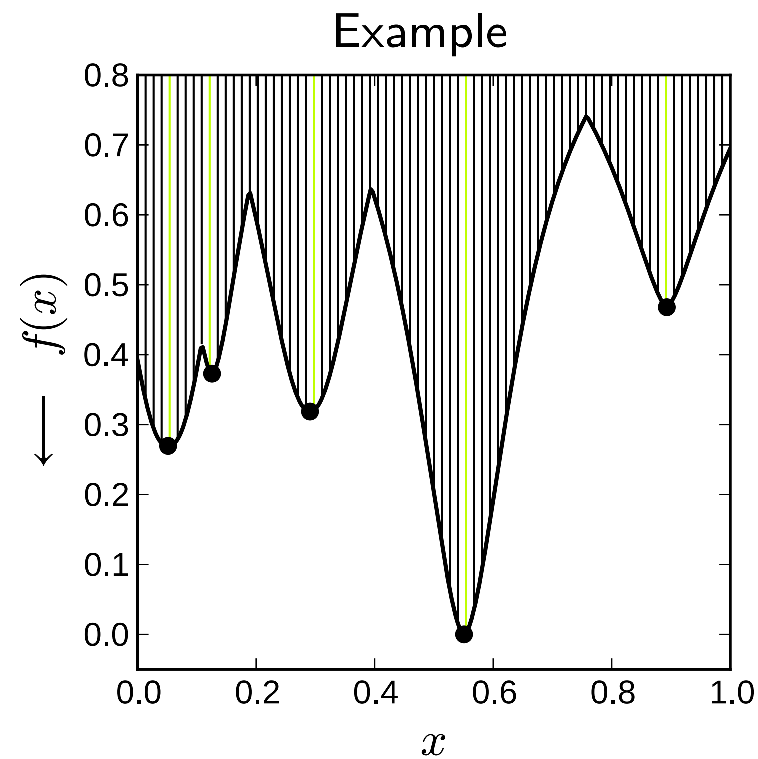
- **Stratified sampling** procedure for efficient initialization
- **Nearest-better clustering** to approximate modes
- **Local search** procedure to accurately locate mode (Nelder-Mead or CMAES)
- **Restart** procedure

S. Wessing, M. Preuss, and G. Rudolph. Assessing basin identification methods for locating multiple optima. In: Advances in Stochastic and Deterministic Global Optimization. Springer, 2016.
Simon Wessing. Two-stage methods for multimodal optimization. PhD thesis, Technische Universität Dortmund, 2015.
Many thanks to Simon Wessing for providing us slides for RLSIS

RLSIS: One-dimensional Problems

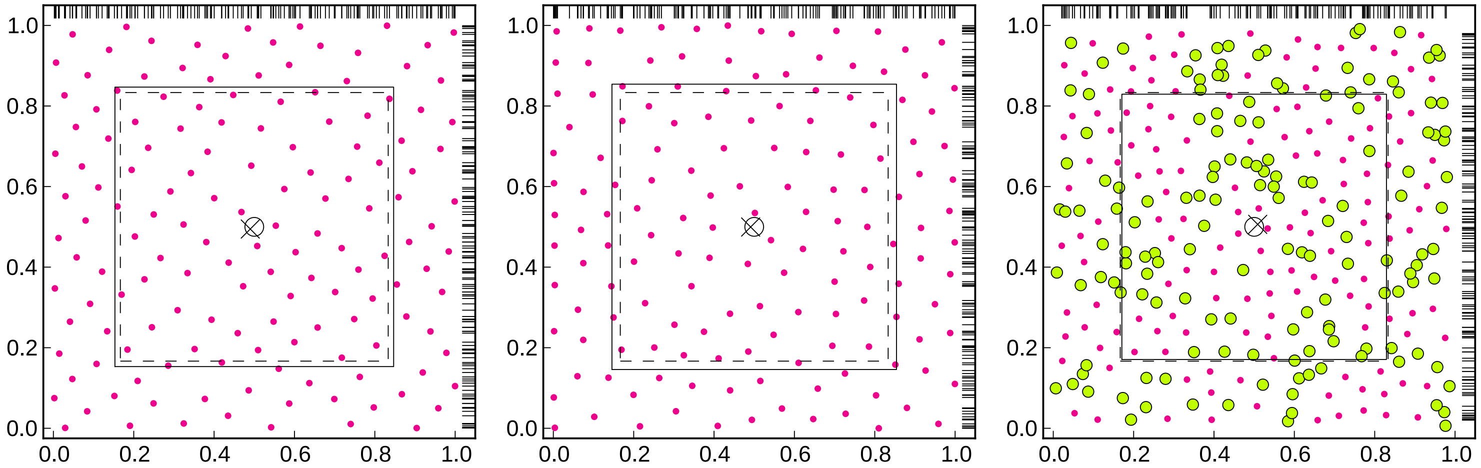
1. **Stratified sample** of 300 points
2. **Nearest-better clustering** of the sample, to obtain approx. local optima
3. **Nelder-Mead** simplex search for each selected point, to increase precision (initial step size 10−4)

- Consumed budget ≈ 1000 evaluations



RLSIS: Maximin Reconstruction Algorithm (MmR)

- **Basic principle**: *maximization of minimal distance*
  Complement with correction methods for edge effects
  - **Torus** → *periodic edge correction (PEC)*
  - **Mirroring** → *reflection edge correction (REC)* (not used here)
- Optional: consider a set of existing points (green points in (c))



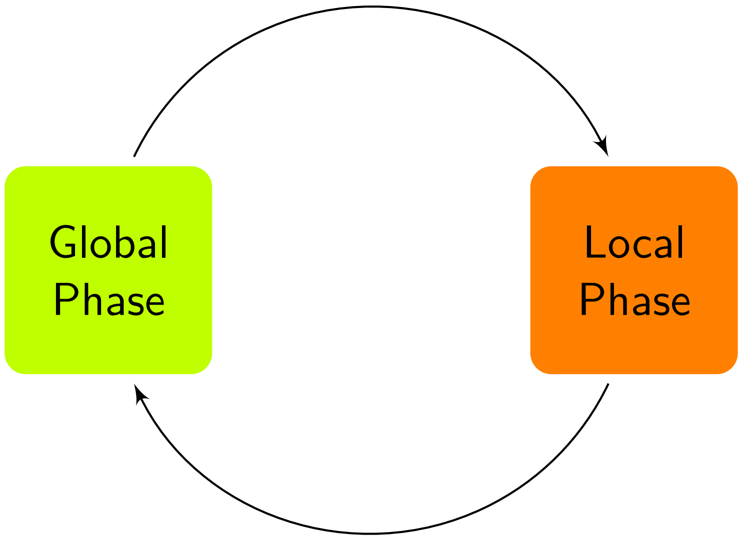
(a) $p = 0.5$ (b) $p = 2$ (c) +PEC, +REC, $p = 2$

S. Wessing, M. Preuss, and G. Rudolph. Assessing basin identification methods for locating multiple optima. In: Advances in Stochastic and Deterministic Global Optimization. Springer, 2016.

RLSIS: Multi-dimensional Problems

- **Restarted Local Search (RLS):**

1. Determine a **starting point**
2. Execute **local search** (CMA-ES) with this starting point
3. Go to 1. (**restart**)



New: starting points and found optima are saved in an archive and considered by MmR in following iterations

Multi-Modal Optimization

CHALLENGES IN NICHING METHODS

Challenges in Niching Methods

- **Searching Efficiency** (local/global search)
- **Maintaining found solutions** (archives,…)
- **Specifying niching parameters** (No parameters!)
  - Attempting to find a single uniform niche radius
  - Dynamic niche radius, instead of fixed one
  - Avoid to specify/use niche parameters
- **Scalability** (dimension, #optima)
- **Measuring performance**
  - Benchmarks or Real-World applications

Multi-Modal Optimization

BENCHMARK TEST FUNCTIONS FOR MULTI-MODAL OPTIMIZATION

CEC'2013 niching benchmark

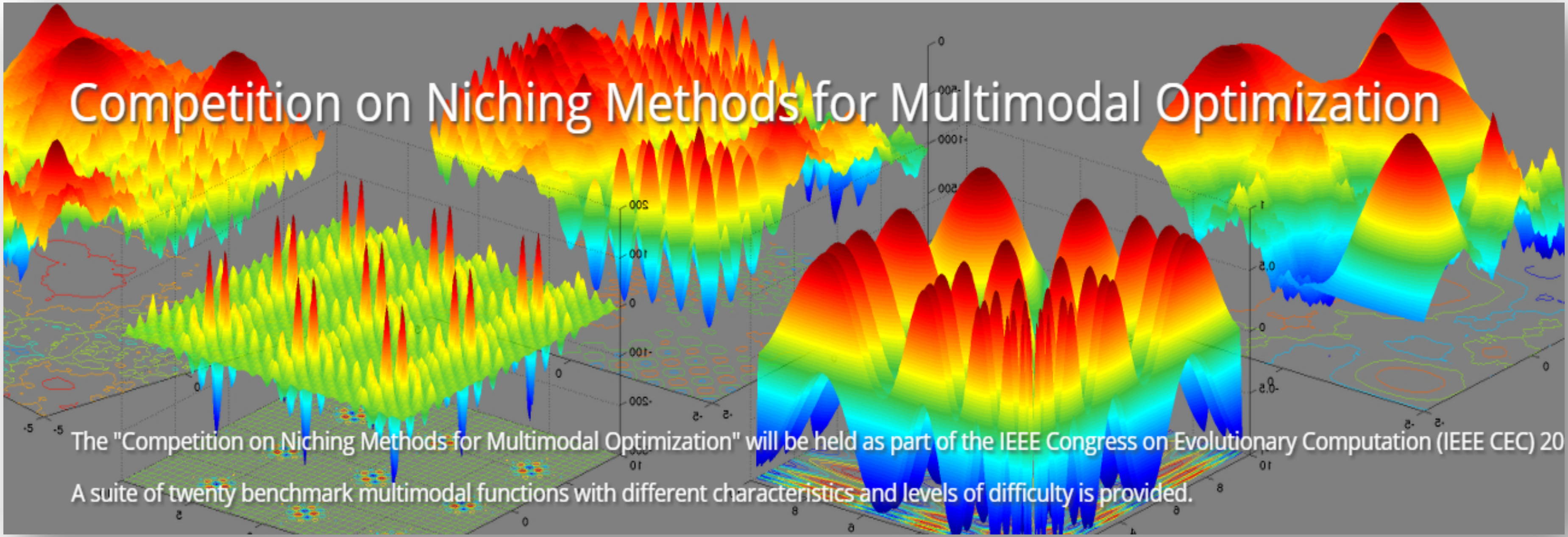
- A **common platform** for fair and easy **evaluation** and **comparisons** of different niching algorithms
- **20 benchmark** multimodal functions with different characteristics
- **5 accuracy levels**: $\varepsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$
- The benchmark suite and the performance measures have been implemented in: **C/C++, JAVA, MATLAB & Python (R to come soon)**

X. Li, A. Engelbrecht, and M.G. Epitropakis, "Benchmark Functions for CEC'2013 Special Session and Competition on Niching Methods for Multimodal Function Optimization", Technical Report, Evolutionary Computation and Machine Learning Group, RMIT University, Australia, 2013.

CEC 2013/2015/2016 competitions

- IEEE CEC niching competitions at 2013,2015 and 2016, with the latest results available at the following URL:

  http://goanna.cs.rmit.edu.au/~xiaodong/cec13-niching/competition/
  http://goanna.cs.rmit.edu.au/~xiaodong/cec15-niching/competition/
  http://www.epitropakis.co.uk/cec16-niching/competition/
  https://github.com/mikeagn/CEC2013

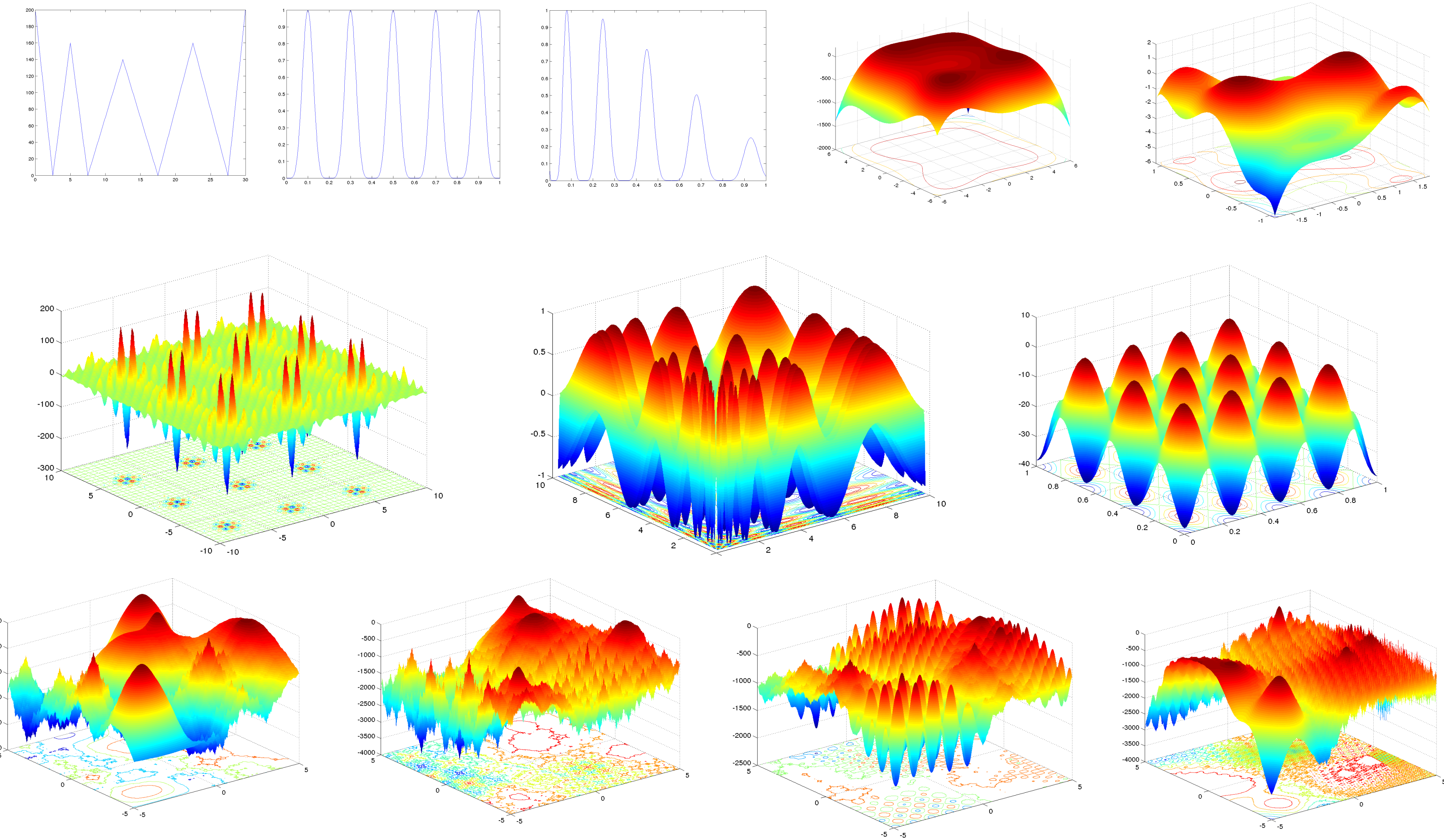


20 test functions

Id	Dim.	# GO	Name	Characteristics
F_1	1	2	Five-Uneven-Peak Trap	Simple, deceptive
F_2	1	5	Equal Maxima	Simple
F_3	1	1	Uneven Decreasing Maxima	Simple
F_4	2	4	Himmelblau	Simple, non-scalable, non-symmetric
F_5	2	2	Six-Hump Camel Back	Simple, not-scalable, non-symmetric
F_6	2,3	18,81	Shubert	Scalable, #optima increase with D, unevenly distributed grouped optima
F_7	2,3	36,216	Vincent	Scalable, #optima increase with D, unevenly distributed optima
F_8	2	12	Modified Rastrigin	Scalable, #optima independent from D, symmetric
F_9	2	6	Composition Function 1	Scalable, separable, non-symmetric
F_{10}	2	8	Composition Function 2	Scalable, separable, non-symmetric
F_{11}	2,3,5,10	6	Composition Function 3	Scalable, non-separable, non-symmetric
F_{12}	2,3,5,10	8	Composition Function 4	Scalable, non-separable, non-symmetric

X. Li, A. Engelbrecht, and M.G. Epitropakis, “Benchmark Functions for CEC’2013 Special Session and Competition on Niching Methods for Multimodal Function Optimization”, Technical Report, Evolutionary Computation and Machine Learning Group, RMIT University, Australia, 2013.

20 test functions (graphs)



Determining found global optima

- First we need to ***specify a level of accuracy*** (typically $0 < \varepsilon \leq 1$)
  - a threshold value under which we would consider a global optimum is found
- Second, we ***assume*** that for each test function, the following ***information is available***:
  - The number of global optima
  - The fitness of the global optima (or peak height), which is known or can be estimated
  - A niche radius value that can *sufficiently distinguish* two closest global optima

X. Li, A. Engelbrecht, and M. Epitropakis, "Benchmark functions for cec'2013 special session and competition on niching methods for multimodal function optimization," Technical Report, Evolutionary Computation and Machine Learning Group, RMIT University, 2013.

Determining found global optima

input : L_{sorted} - a list of individuals (candidate solutions) sorted in decreasing fitness values;
ϵ - accuracy level; r - niche radius;
ph - the fitness of global optima (or peak height)
output: S - a list of best-fit individuals identified as solutions

```
begin
  S = ∅;
  while not reaching the end of L_sorted do
    Get best unprocessed p ∈ L_sorted;
    found ← FALSE;
    if d(ph, fit(p)) ≤ ϵ) then
      for all s ∈ S do
        if d(s, p) ≤ r then
          found ← TRUE;
          break;
        end
      end
      if not found then
        let S ← S ∪ {p};
      end
    end
  end
end
```

PARAMETERS USED FOR PERFORMANCE MEASUREMENT.

Function	r	Peak height	No. global optima
F_1 (1D)	0.01	200.0	2
F_2 (1D)	0.01	1.0	5
F_3 (1D)	0.01	1.0	1
F_4 (2D)	0.01	200.0	4
F_5 (2D)	0.5	1.03163	2
F_6 (2D)	0.5	186.731	18
F_7 (2D)	0.2	1.0	36
F_6 (3D)	0.5	2709.0935	81
F_7 (3D)	0.2	1.0	216
F_8 (2D)	0.01	-2.0	12
F_9 (2D)	0.01	0	6
F_{10} (2D)	0.01	0	8
F_{11} (2D)	0.01	0	6
F_{11} (3D)	0.01	0	6
F_{12} (3D)	0.01	0	8
F_{11} (5D)	0.01	0	6
F_{12} (5D)	0.01	0	8
F_{11} (10D)	0.01	0	6
F_{12} (10D)	0.01	0	8
F_{12} (20D)	0.01	0	8

Performance measures

Measures:

Peak Ratio (PR) measures the average percentage of all known global optima found over multiple runs:

$$\mathrm{PR} = \frac{\sum_{run=1}^{NR} \text{\# of Global Optima}_i}{(\text{\# of known Global Optima}) * (\text{\# of runs})}$$

Who is the winner:

- The participant with the highest average Peak Ratio performance on all benchmarks wins.
- In all functions the following holds: the higher the PR value, the better

CEC 2013/2015 niching competition top 4 entries

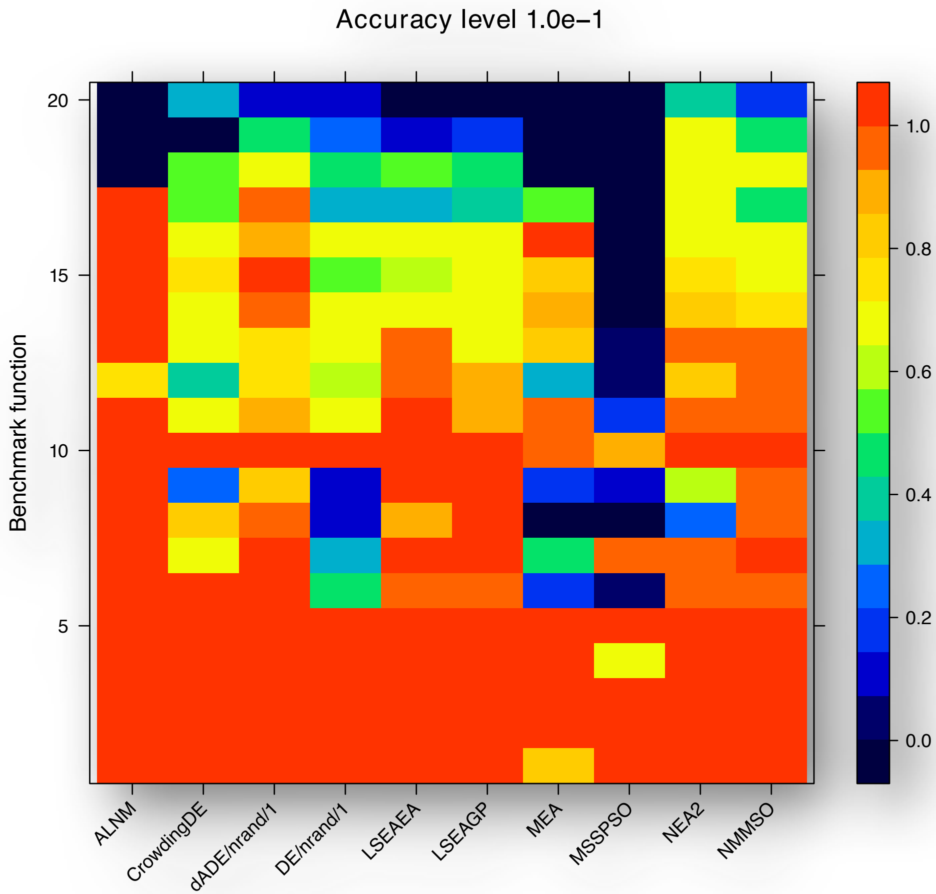
- **(NMMO) Niching Migratory Multi-Swarm Optimiser:**
  – J. E. Fieldsend, "Running Up Those Hills: Multi-Modal Search with the Niching Migratory Multi-Swarm Optimiser," in IEEE Congress on Evolutionary Computation, 2014, pp. 2593 – 2600.
- **(NEA2) Niching the CMA-ES via Nearest-Better Clustering**:
  – M. Preuss. "Niching the CMA-ES via nearest-better clustering." In Proceedings of the 12th annual conference companion on Genetic and evolutionary computation (GECCO ’10). ACM, New York, NY, USA, pp. 1711-1718, 2010.
- **(LSEAGP) Localised Search Evolutionary Algorithm using a Gaussian Process**:
  – J. E. Fieldsend, "Multi-Modal Optimisation using a Localised Surrogates Assisted Evolutionary Algorithm," in UK Workshop on Computational Intelligence (UKCI 2013), 2013, pp. 88-95.
- **(dADE/nrand/1) A Dynamic Archive Niching Differential Evolution algorithm for Multimodal Optimization:**
  – M. G. Epitropakis, Li, X., and Burke, E. K., "A Dynamic Archive Niching Differential Evolution Algorithm for Multimodal Optimization", IEEE Congress on Evolutionary Computation, 2013. CEC 2013. Cancun, Mexico, pp. 79-86, 2013.

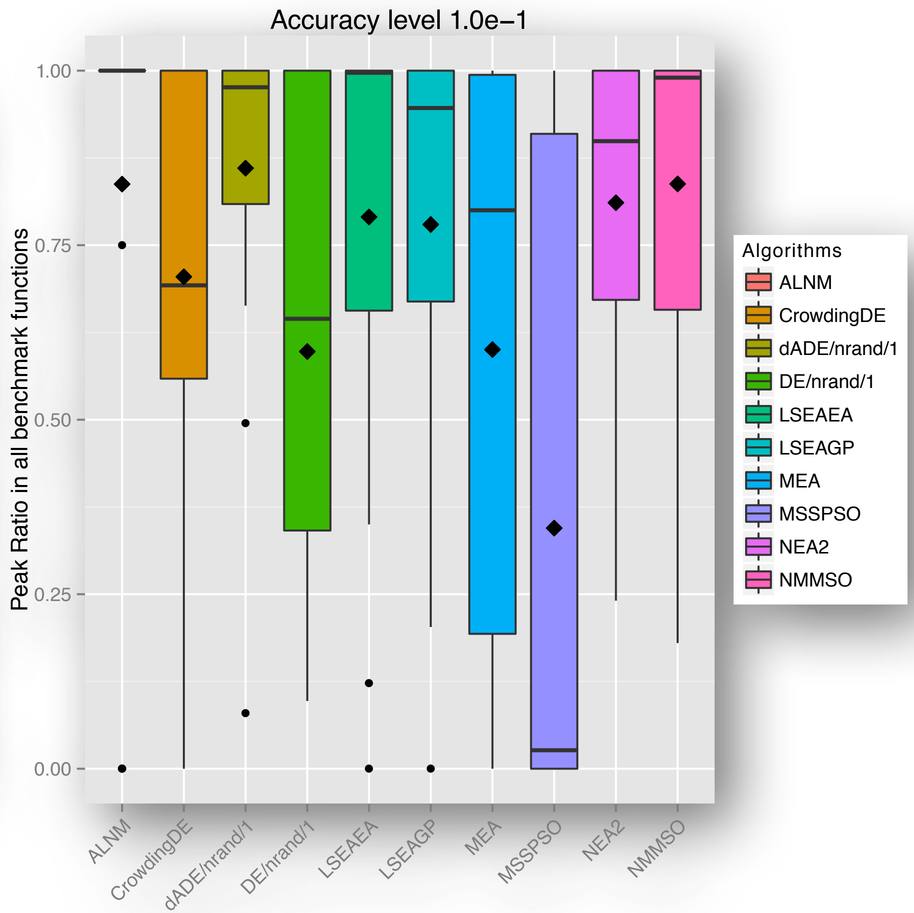
CEC 2013/2015 niching competitions

Summary:

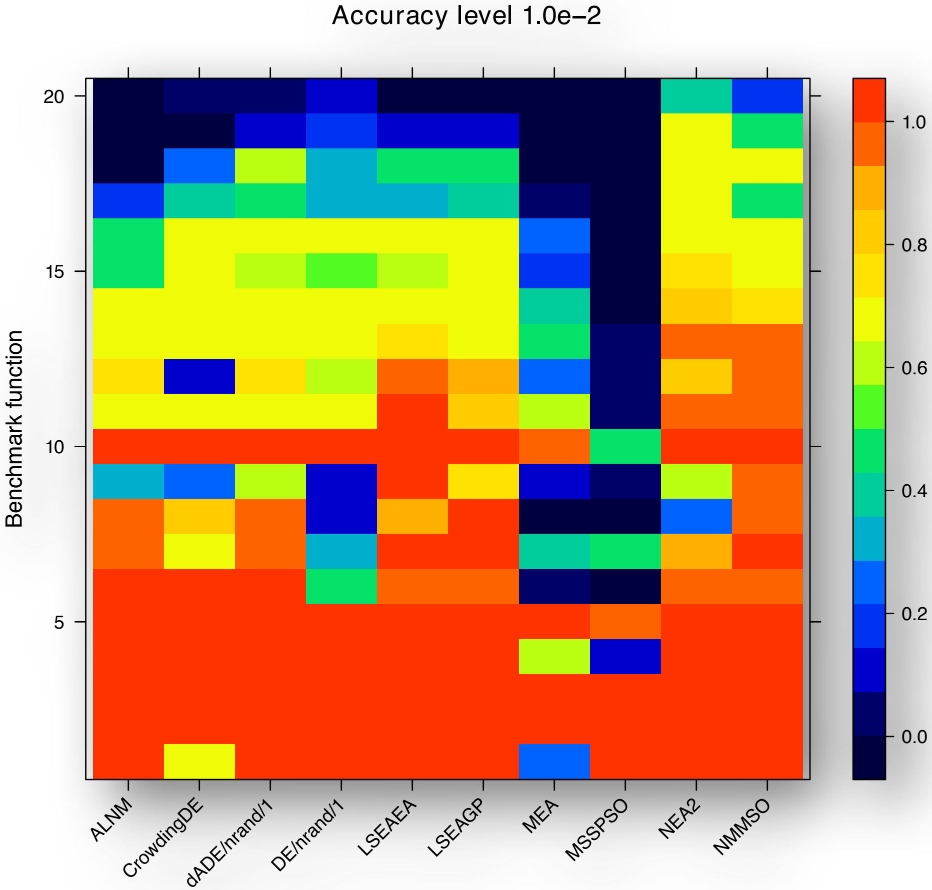
- **6** new search algorithms
- **4** comparators based on the competition @ CEC2013
- **20** multi-modal benchmark functions
- **5** accuracy levels $\varepsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$
- Results: **per accuracy level & over all accuracy levels**
- In total (CEC2013 & CEC2015) 21 algorithms in the repository: `https://github.com/mikeagn/CEC2013`

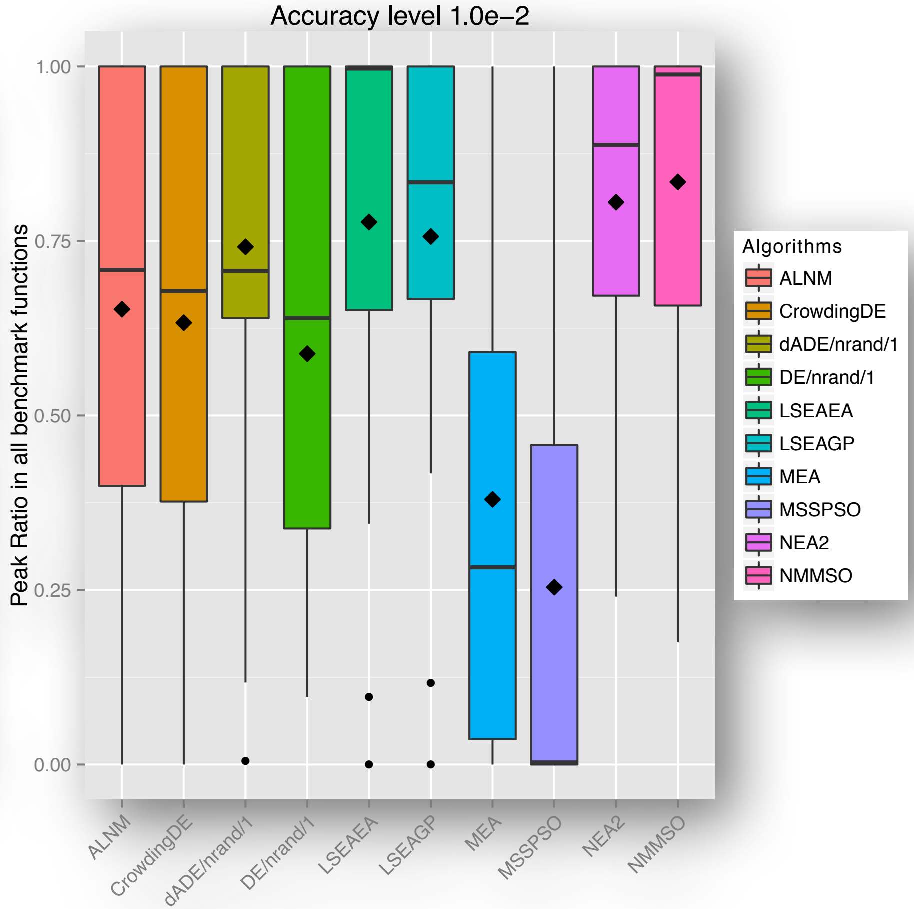
Accuracy level 10^{-1}



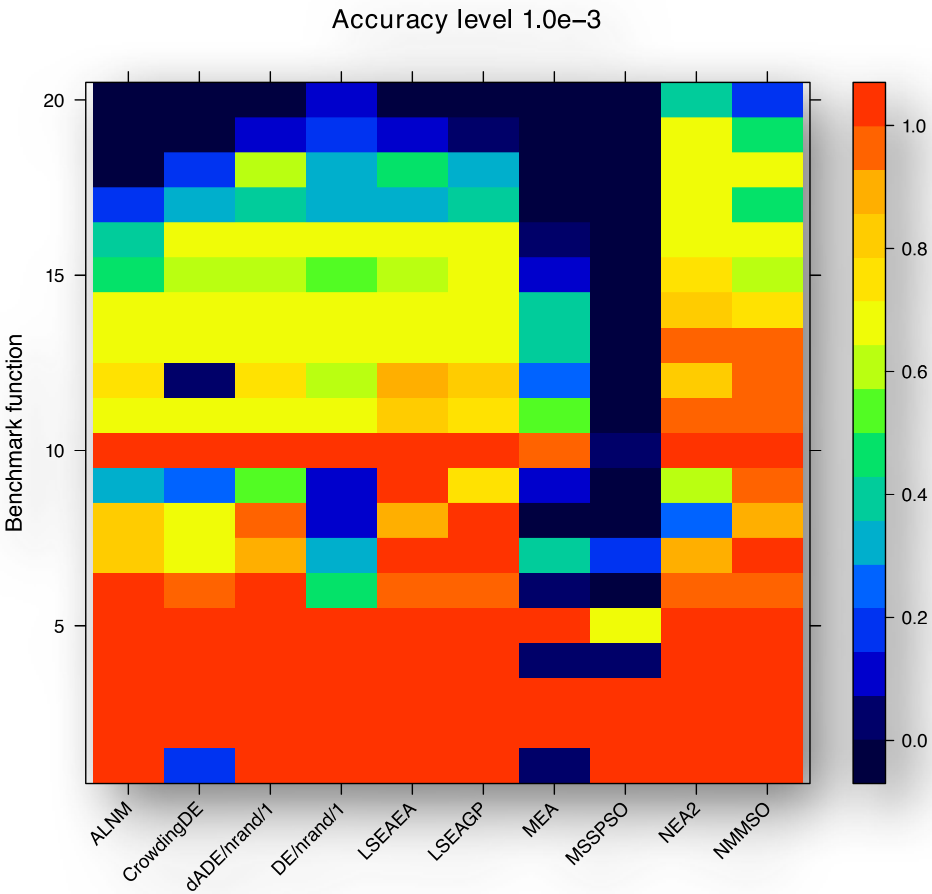


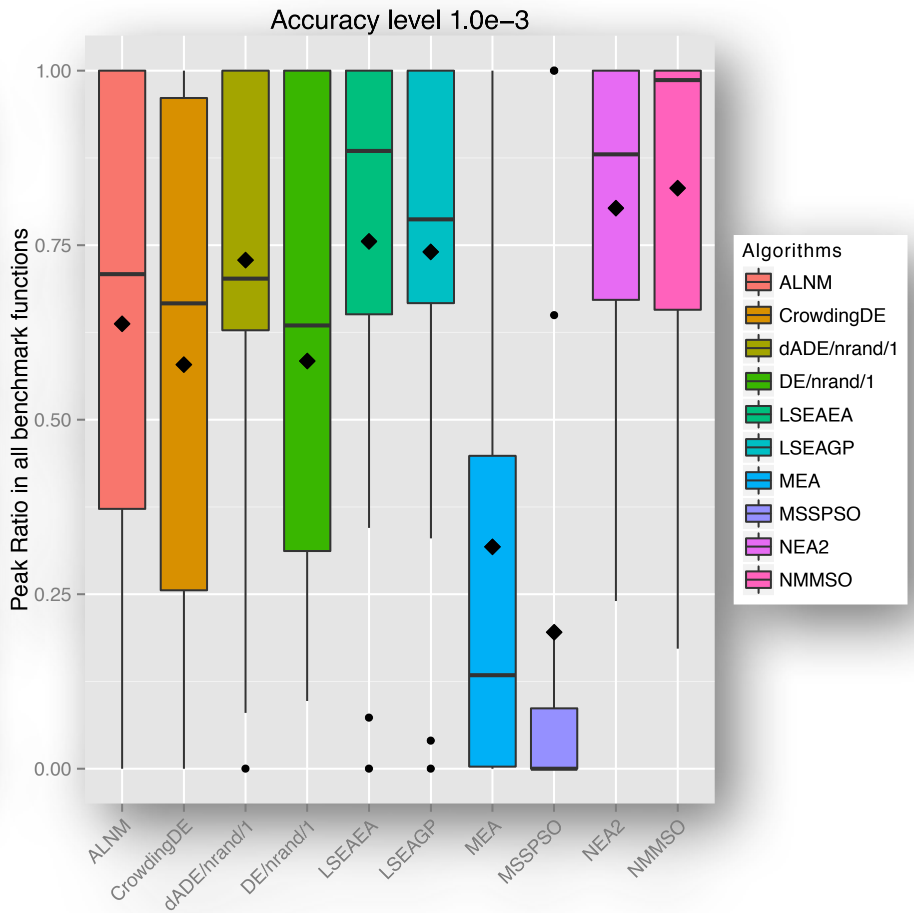
Accuracy level 10^{-2}



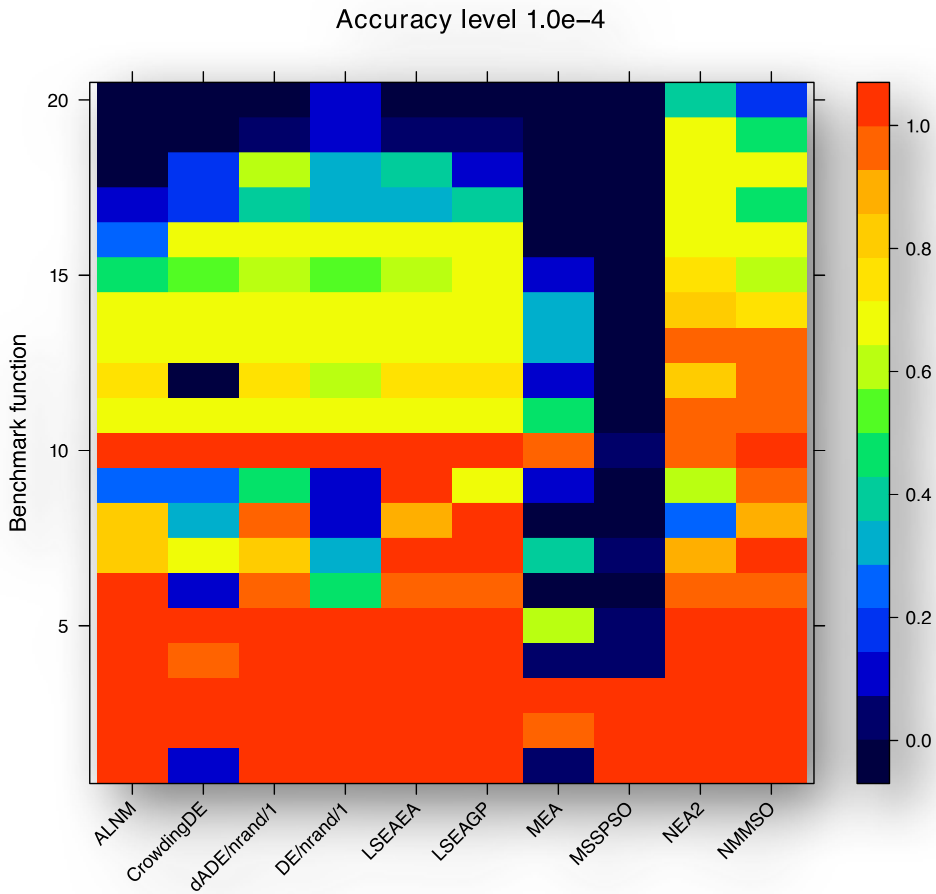


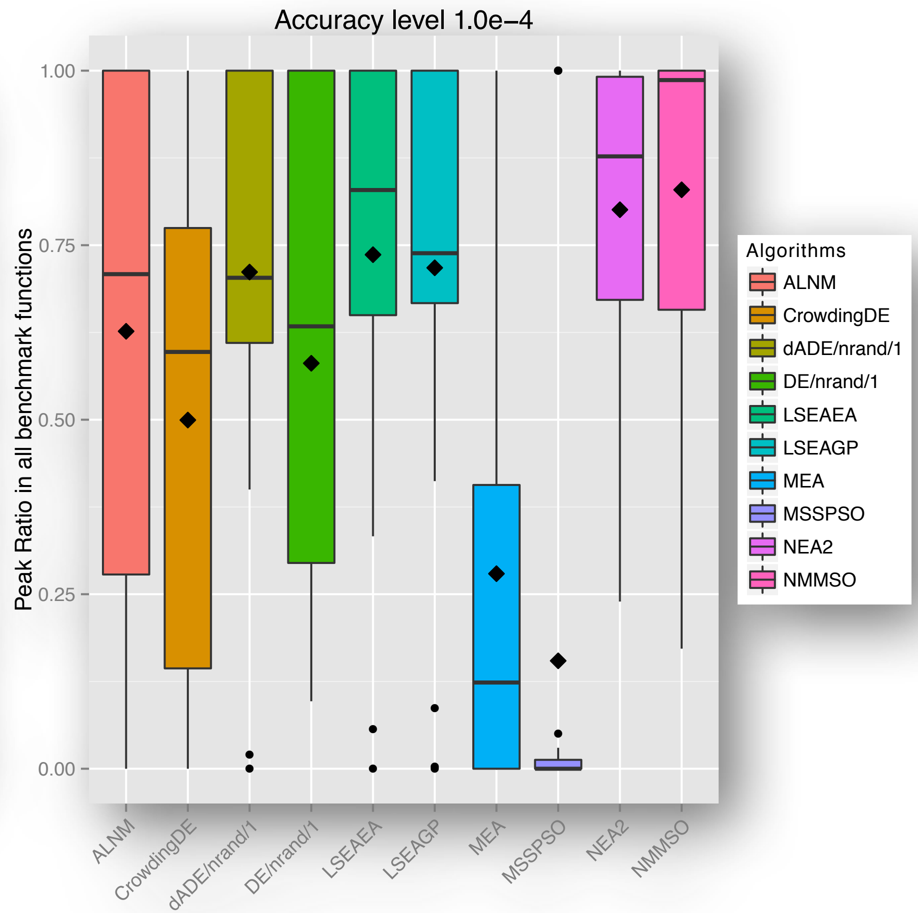
Accuracy level 10^{-3}



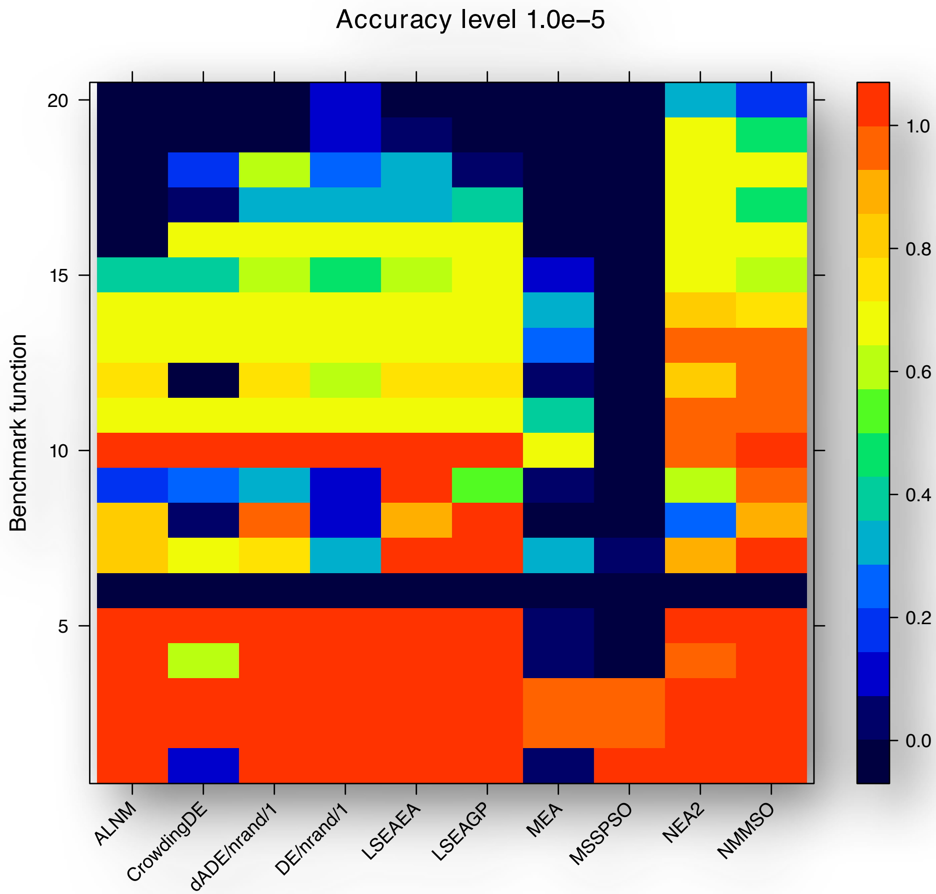


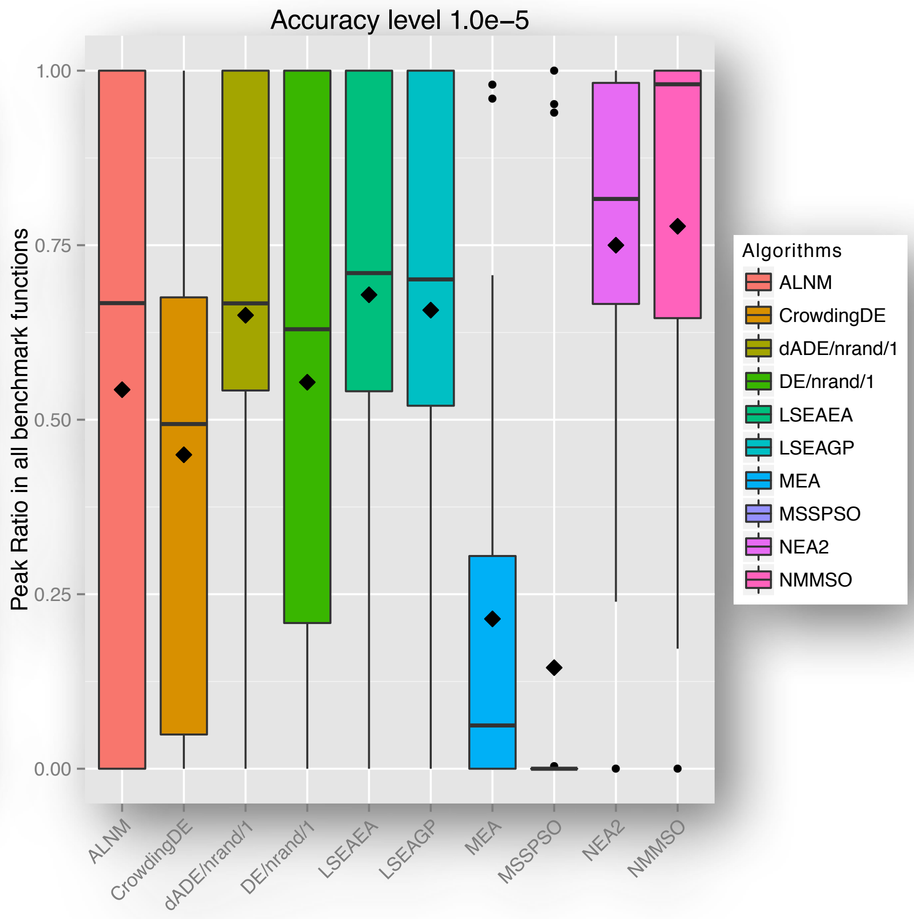
Accuracy level 10^{-4}



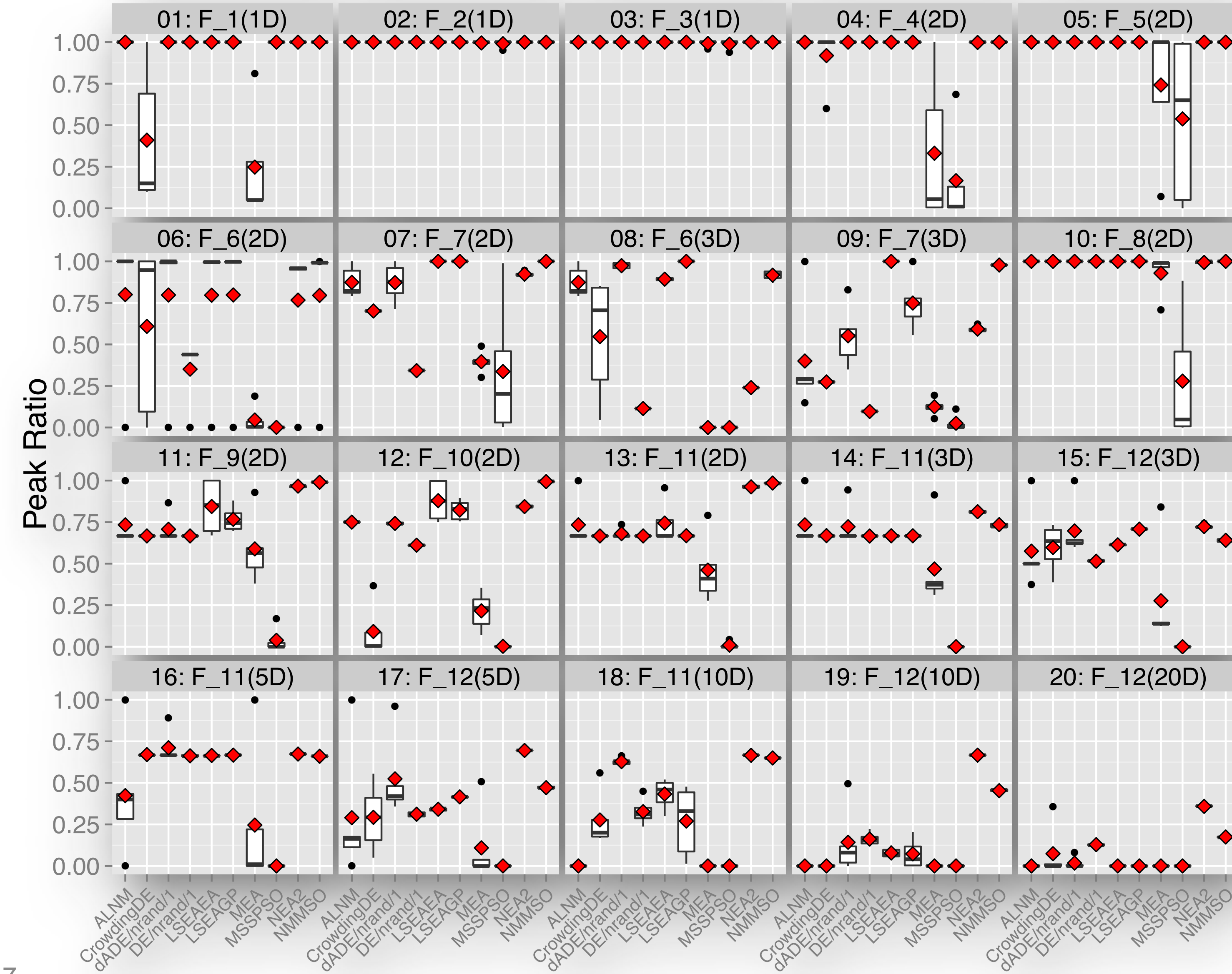


Accuracy level 10^{-5}

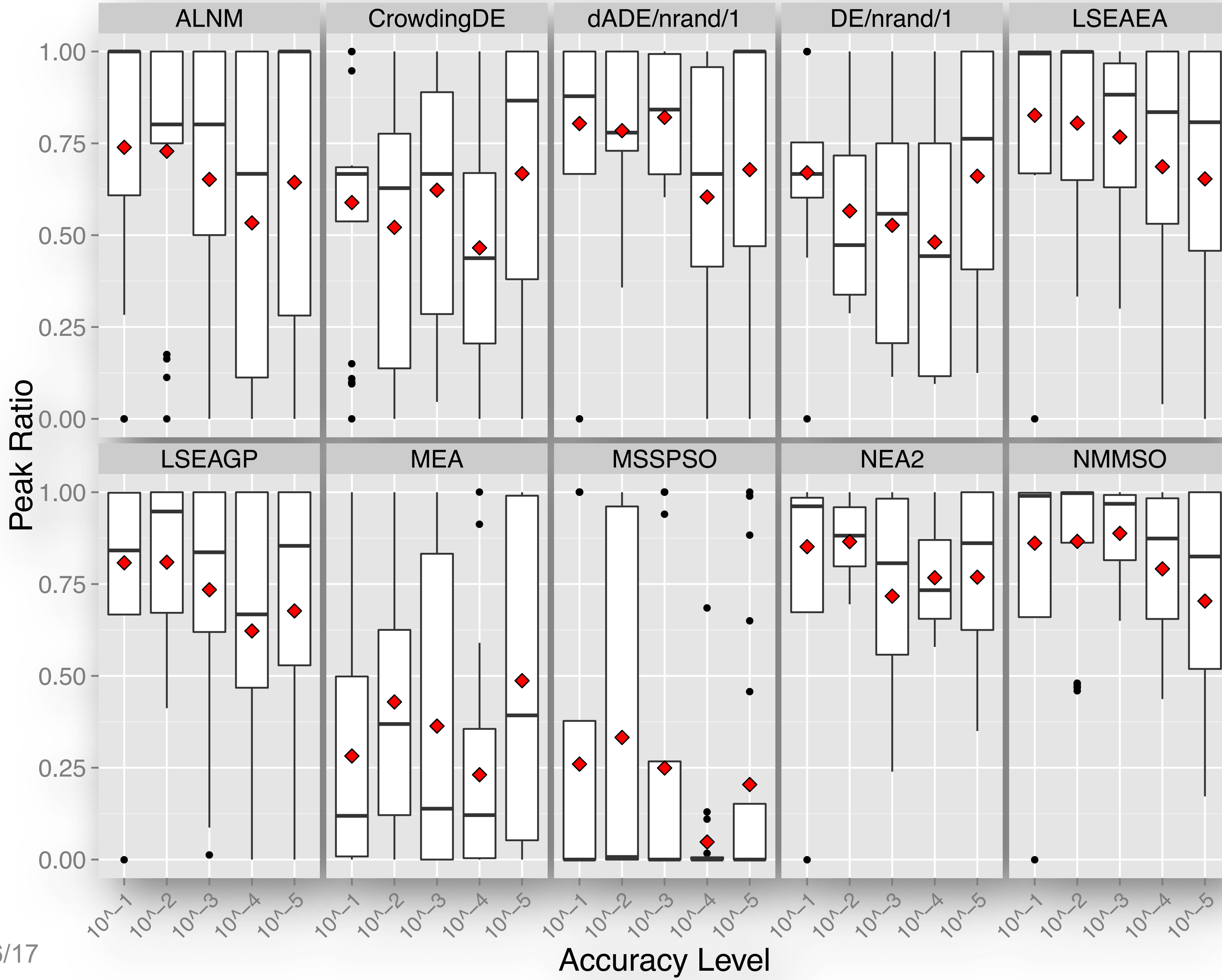




Performance per benchmark



Performance per algorithm

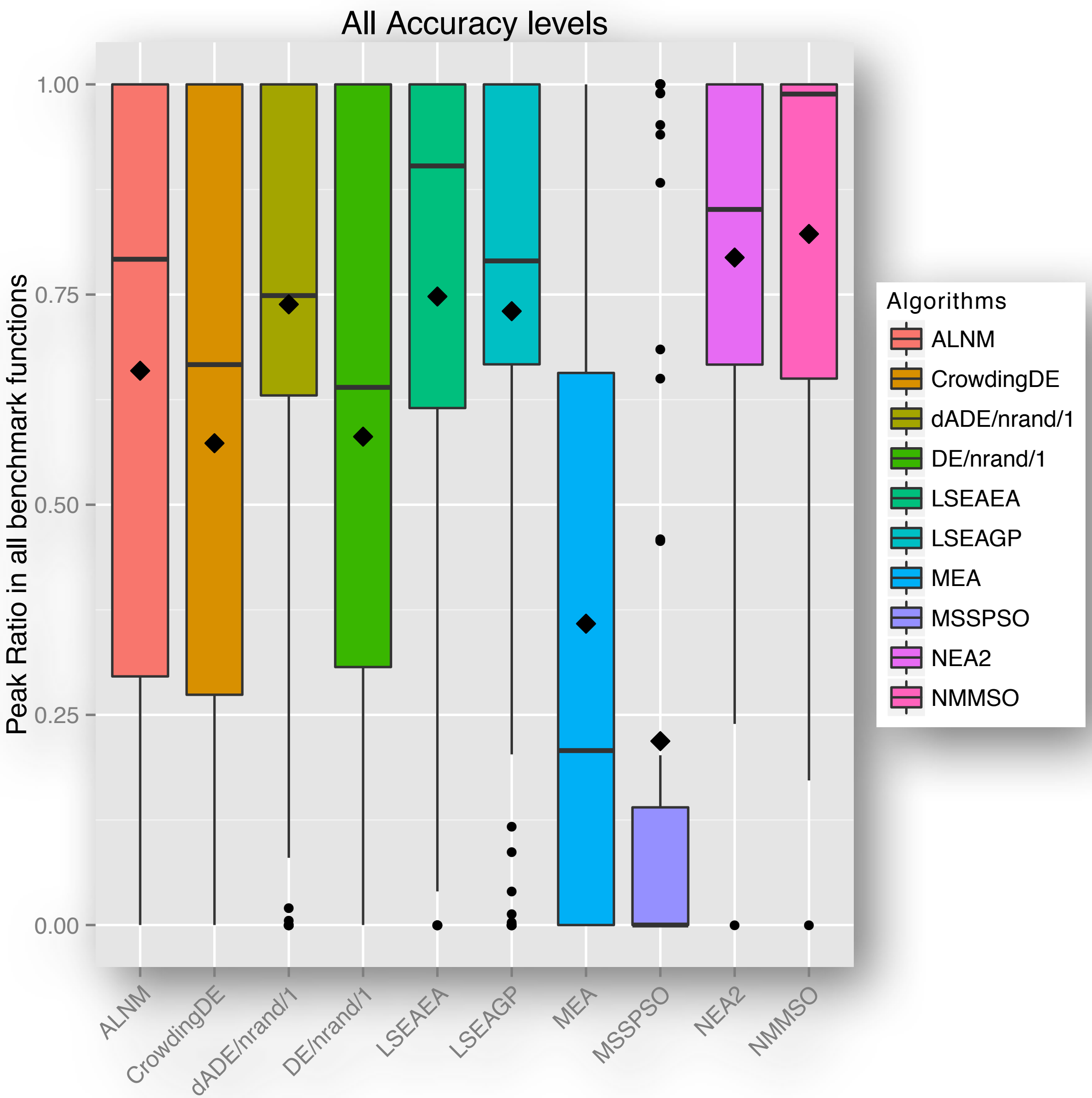


Statistical analysis

	ALNM	CrowdingDE	dADE/nrand/1	DE/nrand/1	LSEAEA	LSEAGP	MEA	MSSPSO	NEA2
	p/p_b	p/p_b	p/p_b	p/p_b	p/p_b	p/p_b	p/p_b	p/p_b	p/p_b
CrowdingDE	–/=	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A
dADE/nrand/1	–/–	+/+	N/A	N/A	N/A	N/A	N/A	N/A	N/A
DE/nrand/1	–/=	=/=	–/–	N/A	N/A	N/A	N/A	N/A	N/A
LSEAEA	+/+	+/+	=/=	+/+	N/A	N/A	N/A	N/A	N/A
LSEAGP	–/–	+/+	=/=	+/+	=/=	N/A	N/A	N/A	N/A
MEA	–/–	–/–	–/–	–/–	–/–	–/–	N/A	N/A	N/A
MSSPSO	–/–	–/–	–/–	–/–	–/–	–/–	–/–	N/A	N/A
NEA2	+/+	+/+	+/+	+/+	–/=	+/=	+/+	+/+	N/A
NMMSO	+/+	+/+	+/+	+/+	+/+	+/+	+/+	+/+	=/=

- p: Wilcoxon rank-sum test
- p_b: Bonferroni correction
- **+** row wins column,
- **–** row loses from column,
- **=** non-significant differences
- N/A: Not Applicable

Overall performance



Participants' performance

Algorithm	Statistics			Friedman's Test	
	Median	Mean	St.D.	Rank	Score
NMMSO	**0.9885**	**0.8221**	0.2538	**1**	16.1900
NEA2	**0.8513**	**0.7940**	0.2332	**2**	16.1150
LSEAEA	**0.9030**	**0.7477**	0.3236	**4**	14.5050
dADE/nrand/1	0.7488	**0.7383**	0.3010	5	14.2450
LSEAGP	**0.7900**	0.7302	0.3268	**3**	14.7550
CMA-ES	0.7550	0.7137	0.2807	6	14.0800
N-VMO	0.7140	0.6983	0.3307	7	13.7600
ALNM	0.7920	0.6594	0.3897	9	12.4900
PNA-NSGAII	0.6660	0.6141	0.3421	11	11.2700
NEA1	0.6496	0.6117	0.3280	14	10.5250
DE/nrand/2	0.6667	0.6082	0.3130	10	11.2950
dADE/nrand/2	0.7150	0.6931	0.3174	8	12.8100
DE/nrand/1	0.6396	0.5809	0.3338	13	10.6150
DELS-aj	0.6667	0.5760	0.3857	15	9.6950
CrowdingDE	0.6667	0.5731	0.3612	12	10.6200
DELG	0.6667	0.5706	0.3925	16	9.4400
DECG	0.6567	0.5516	0.3992	17	8.9900
IPOP-CMA-ES	0.2600	0.3625	0.3117	18	5.8700
MEA	0.2075	0.3585	0.3852	19	5.2750
A-NSGAII	0.0740	0.3275	0.4044	20	4.7200
MSSPSO	0.0000	0.2188	0.3913	21	3.7350

Discussion

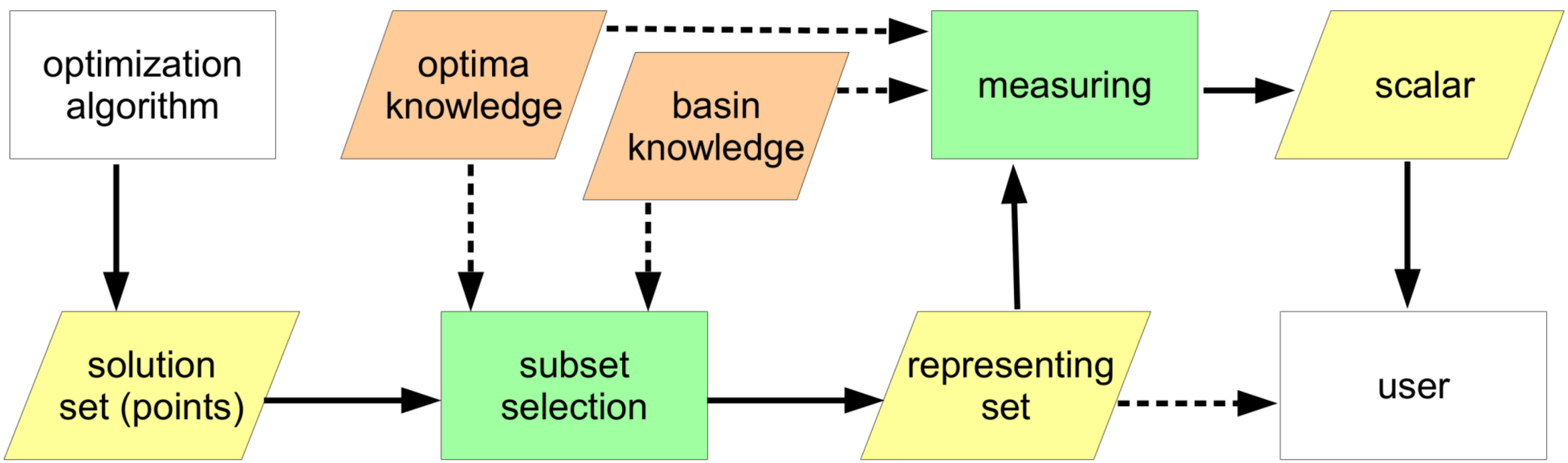
- The competitions gave ***a boost*** to the MMO community
- New ***competitive*** and ***very promising*** approaches
- Key characteristics of the algorithms:
  - **New methodologies**: active learning, surrogates, Gaussian Processes, probabilistic classifier for prediction, archives, hill-valley approaches
  - Usage of **local models** to **maintain diversity** and **exploit locally** the neighborhoods
  - **Algorithms**: EAs, DE, CMA-ES, Multi-swarms, and Bootstrap-LV sampling.

Other niching benchmark sets

- **Earliest work on designing niching benchmark** functions was carried out ***by Deb in his master thesis***!
  – K. Deb, "Genetic algorithms in multimodal function optimization (master thesis and tcga report no. 89002)," Tuscaloosa: University of Alabama, The Clearinghouse for Genetic Algorithms, 1989.
- Tunable **cosine and quadratic** function families
  – J. Ronkkonen, "Continuous Multimodal Global Optimization with Differential Evolution Based Methods. Acta Universitatis Lappeenrantaensis 363, 2009.
- Preuss/Lasarczyk generator: **mixture of polynomials** (MPM)
  – Preuss, Lasarczyk. On the importance of information speed in structured populations. In Proc. PPSN VIII, pp. 91–100, 2004
  – improved version (MPM2) in the dissertation of Simon Wessing: Two-stage Methods for Multimodal Optimization. TU Dortmund, 2015
- Gallagher/Yuan tunable generator: **mixture of Gaussian distributions**
  – Gallagher and B. Yuan. A general-purpose tunable landscape generator. IEEE Trans. Evolutionary Computation, 10(5):590–603, 2006
- **Simple and composition** multimodal functions
  – Qu et al. B. Y. Qu, J. J. Liang, Z. Y. Wang, Q. Chen, and P. N. Suganthan, "Novel benchmark functions for continuous multimodal optimization with comparative results," Swarm and Evolutionary Computation, vol. 26, pp. 23–34, 2016.

Performance Measuring

- Two main components:
  - Subset solution selection
  - Performance measuring



M. Preuss, Multimodal Optimization by Means of Evolutionary Algorithms, ser. Natural Computing Series. Springer International Publishing, 2016.

A note on Performance Measures

indicator	short	requires $f(\vec{x})$	subset sel.	optima known	basins known	param.
sum of distances	SD					
SD to nearest neighbor	SDNN					
Solow-Polasky diversity	SPD					✓
average objective value	AOV	✓				
peak ratio	PR		✓	✓		✓
quantity-adjusted PR	QAPR			✓		✓
peak distance	PD		✓	✓		
augmented PD	APD	✓	✓	✓		
peak accuracy	PA	✓	✓	✓		
averaged Hausdorff distance	AHD			✓		✓
augmented AHD	AAHD	✓		✓		✓
basin ratio	BR		✓	✓	✓	
quantity-adjusted BR	QABR			✓	✓	
basin accuracy	BA	✓	✓	✓	✓	
representative 5 selection	R5S	✓				

- Different advantages/disadvantages
- Many connections with multi-objective metrics.
- Mostly used currently in literature:
  - Peak Ratio (PR), (problematic)

Preuss, Wessing, Measuring Multimodal Optimization Solution Sets with a View to Multiobjective Techniques. In *EVOLVE IV*, pp. 123–137, 2013
M. Preuss, Multimodal Optimization by Means of Evolutionary Algorithms, ser. Natural Computing Series. Springer International Publishing, 2016.

www.epitropakis.co.uk/gecco2016/

Niching @ GECCO 2016

GECCO 2016 Competition on Niching Methods for Multimodal Optimization

In this instance of the competition, we continue the successful 2013 and 2015 competitions and rely on an unchanged problem set and similar measures. However, we try out two new aspects regarding the evaluation criteria of the participants:

1. we are including information on the resources (time, function evaluations) needed to find the global optima, not only the fraction of successes within a given time period (number of evaluations), and
2. we take into account the size of the final solution set, and reward small sets that mostly consist of the sought optima only.

Overview

GECCO 2016 Competition (I)

- Largely follows the procedures of the 2013/2015 CEC niching competitions, adopt new performance criteria:

Improved Scenarios

- Include information on the resources (time, function evaluations) needed to find the global optima, not only the fraction of successes within a given time period (number of evaluations), and
- Take into account the size of the final solution set, and reward small sets that mostly consist of the sought optima only.

GECCO 2016 Competition (II)

Three different Scenarios (performance evaluation):

- **Scenario I**: Adopt the CEC2013/2015 competition ranking procedure (based on average Peak Ratio), to facilitate straight forward comparisons with all previous competition entries.
- **Scenario II**: Adopt the (static) F1 measure to take into account the recall and precision of the final solution sets
- **Scenario III**: Adopt the (dynamic) F1 measure integral over the whole runtime to take into account the computational efficiency of the submitted algorithm

Ranking based on average values across all problems/accuracy levels of the aforementioned measures are used to decide the winner.

Participants

Submissions to the competition:

- (**rlsis**): Restarted Local Search with Improved Selection of Starting Points, Simon Wessing
- (**rs-cmsa-es**): Benchmarking Covariance Matrix Self Adaption Evolution Strategy with Repelling Subpopulations for GECCO 2016 Competition on Multimodal Optimization, Ali Ahrari, Kalyanmoy Deb and Mike Preuss
- (**ascga**): Adaptive species conserving genetic algorithm, Jian-Ping Li, Felician Campean
- (**nea2+**): Niching the CMA-ES via Nearest-Better Clustering: First Steps Towards an Improved Algorithm, Mike Preuss
- (**nmmso**) Niching Migratory Multi-Swarm Optimiser, J. Fieldsend

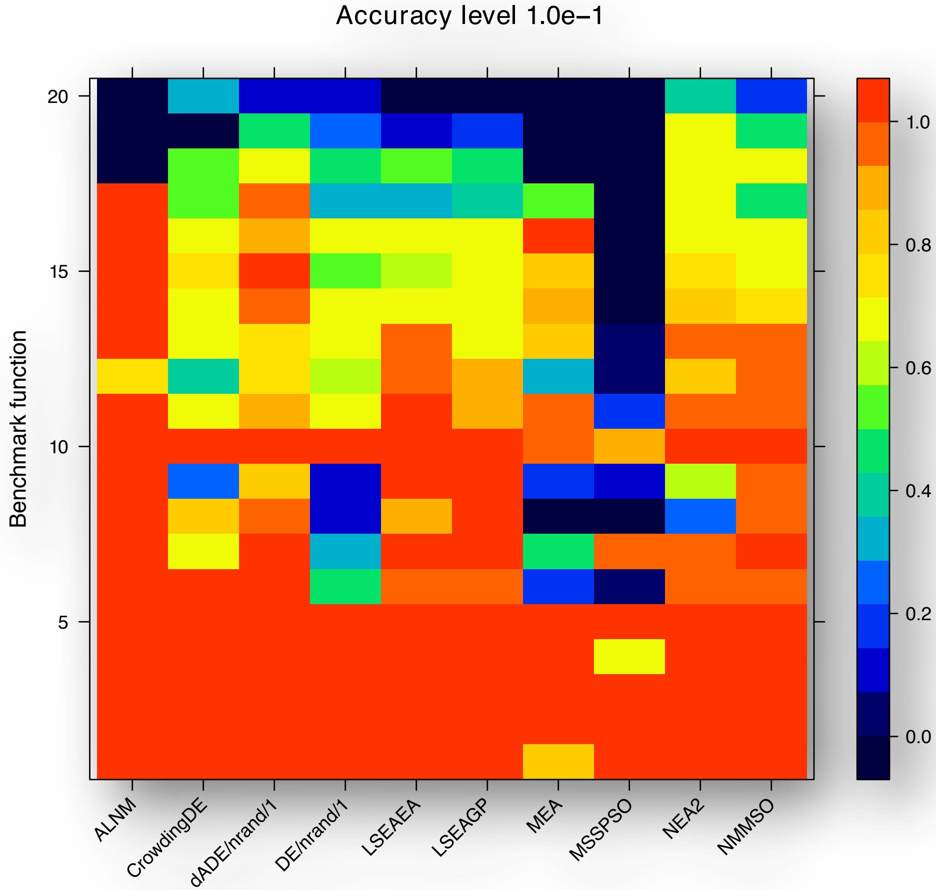
Baseline algorithms: CMA-ES, IPOP-CMA-ES, NEA1, NEA2

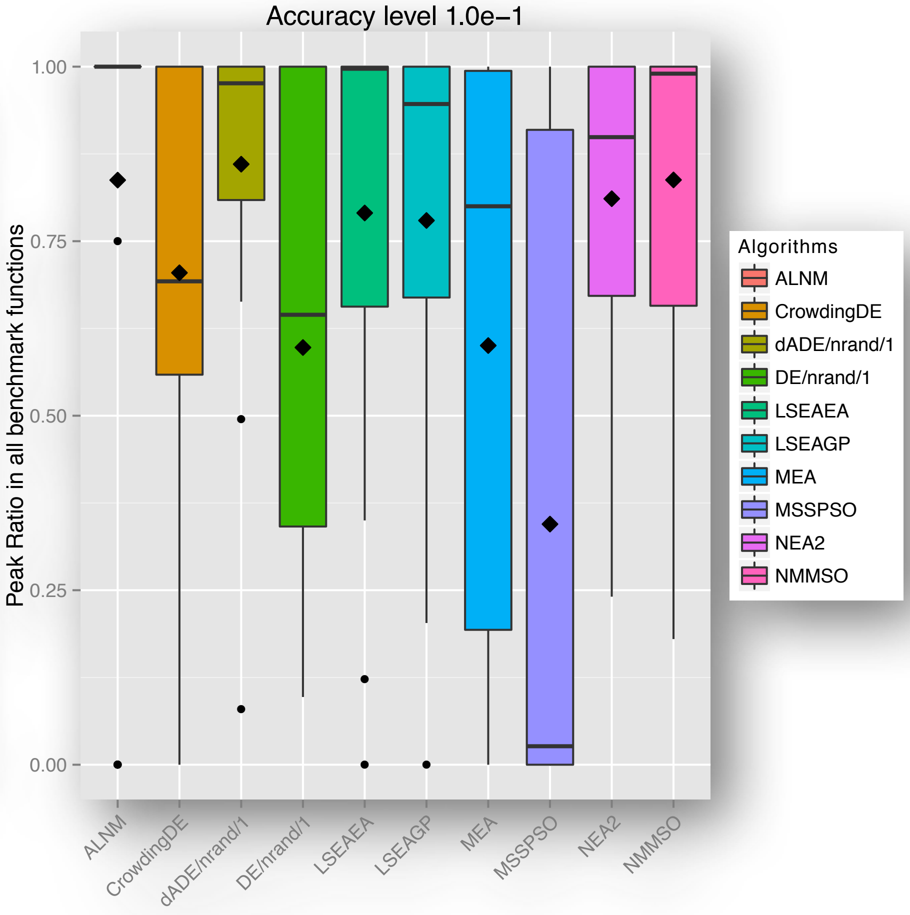
GECCO 2016 Competition Setup

Summary:

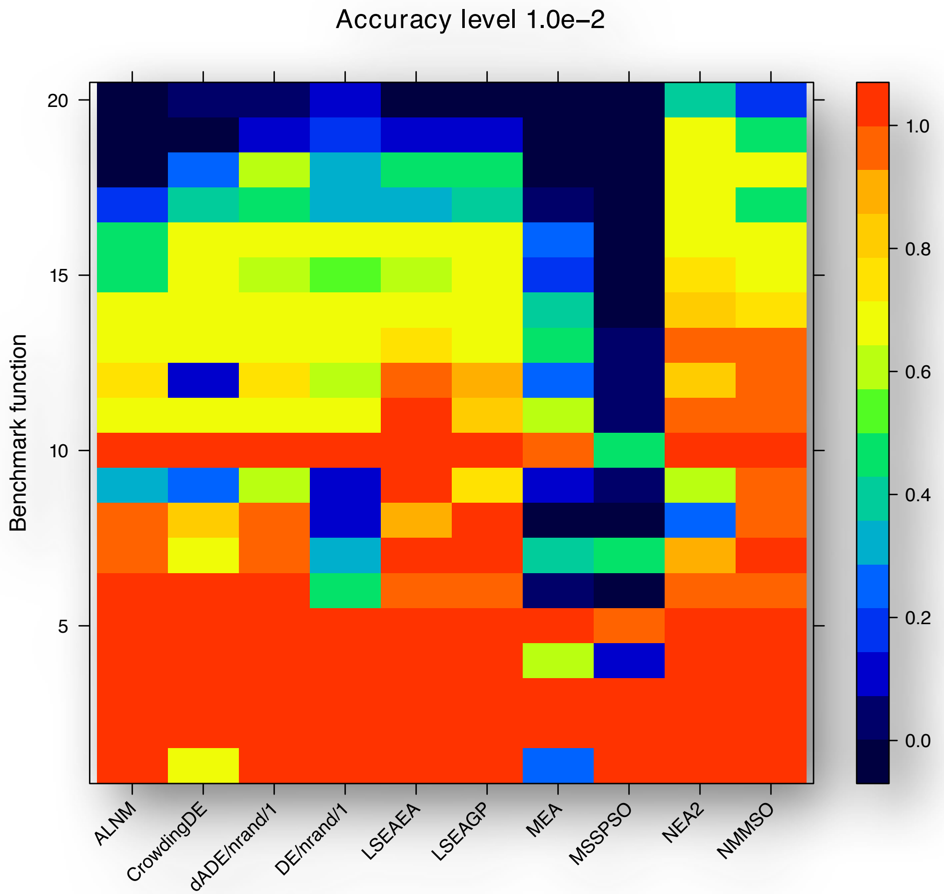
- **5** new search algorithms
- **4** classic algorithm comparators
- **20** multi-modal benchmark functions
- **5** accuracy levels $\varepsilon \in \{10^{-1}, 10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$
- Results: **per accuracy level & over all accuracy levels**
- Latest version always in the repository:
  `https://github.com/mikeagn/CEC2013`

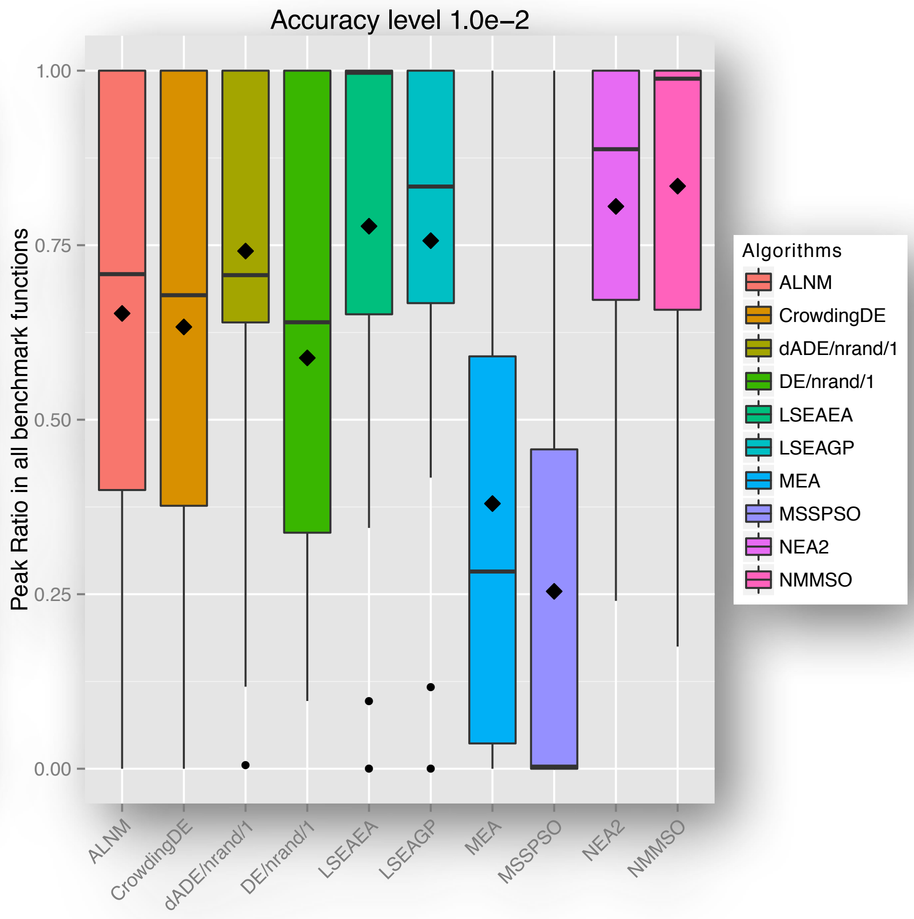
Scenario I: Accuracy level 10^{-1}



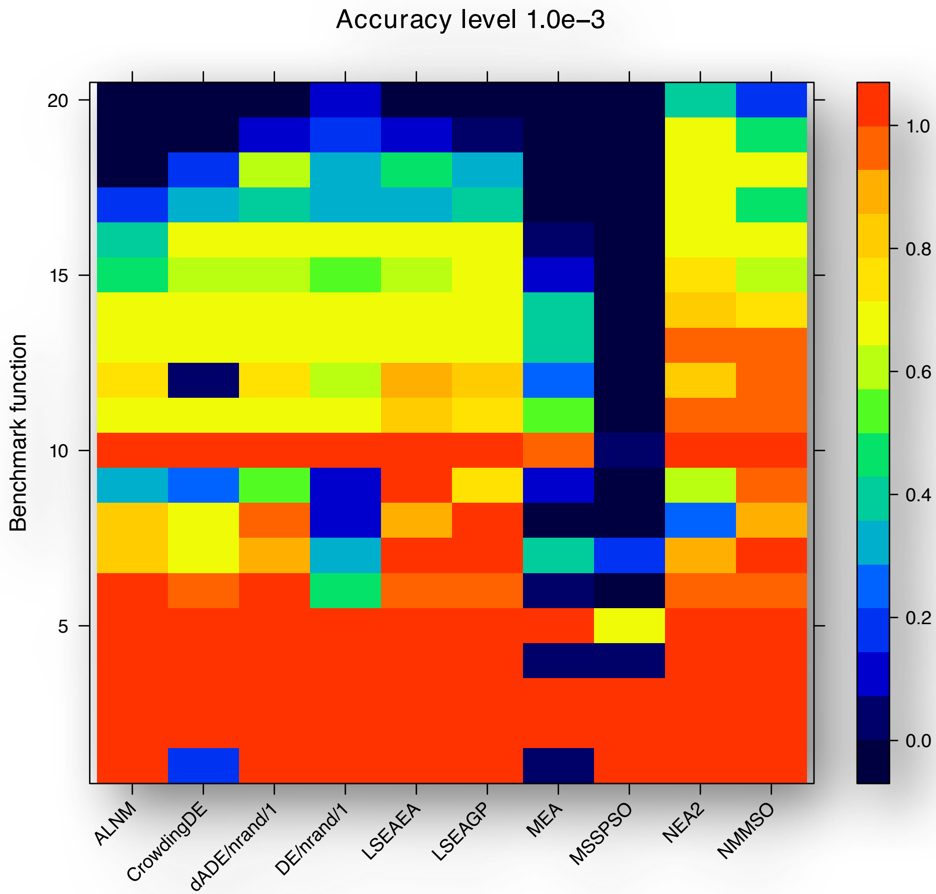


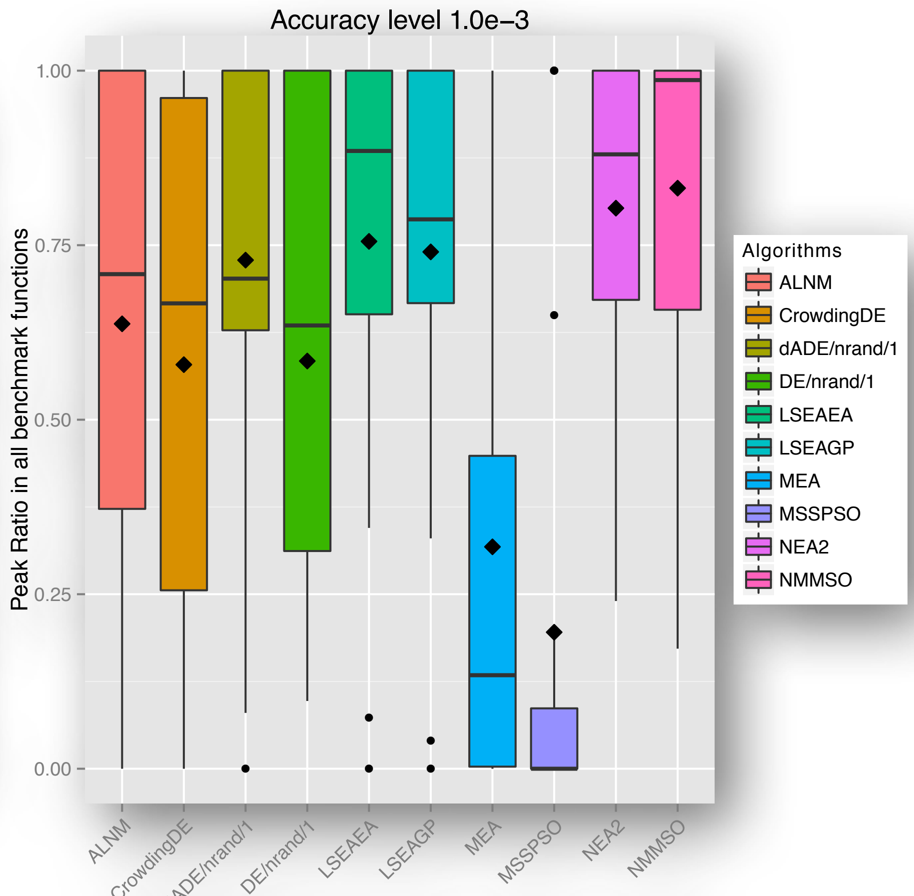
Scenario I: Accuracy level 10^{-2}



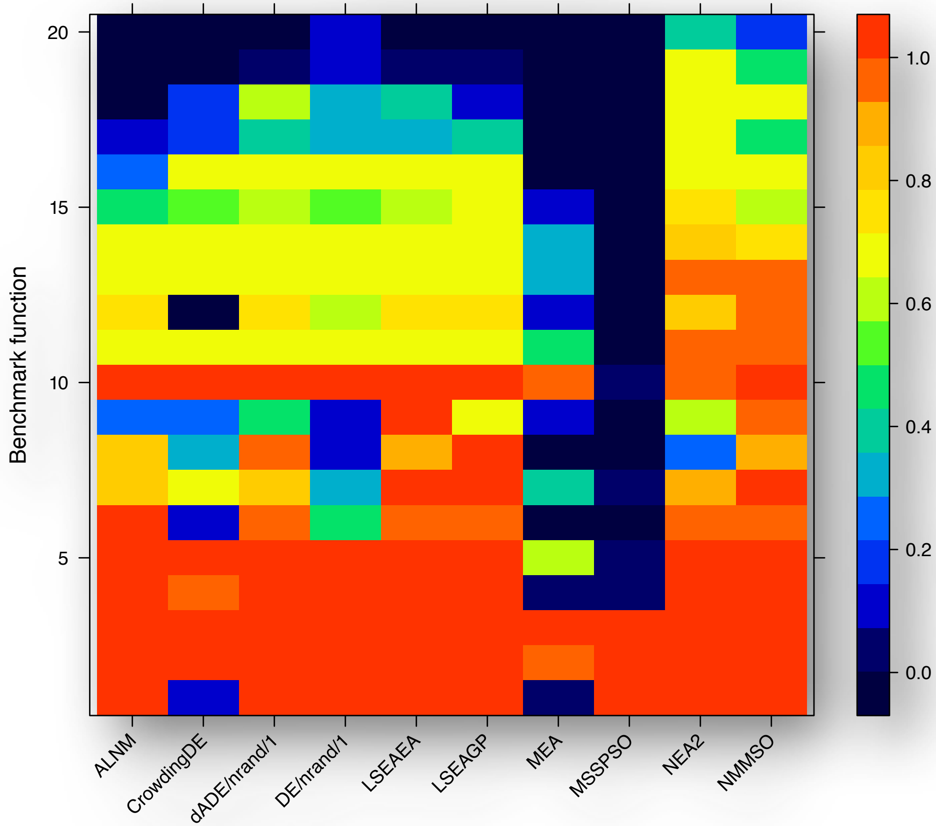


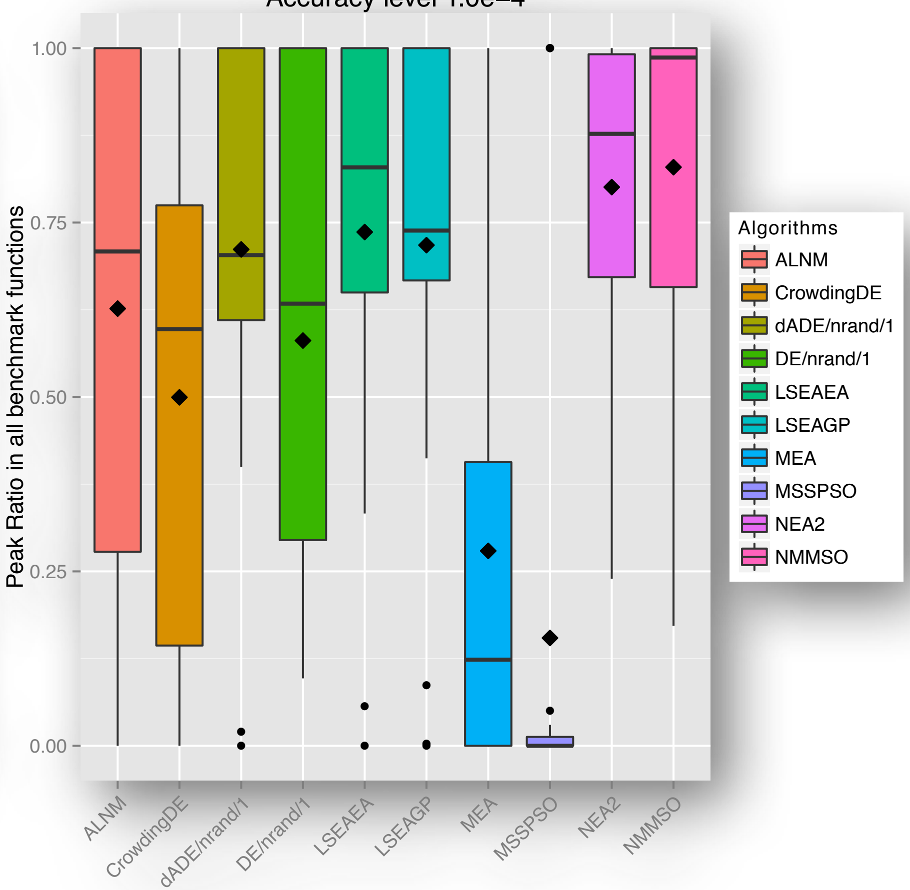
Scenario I: Accuracy level 10^{-3}



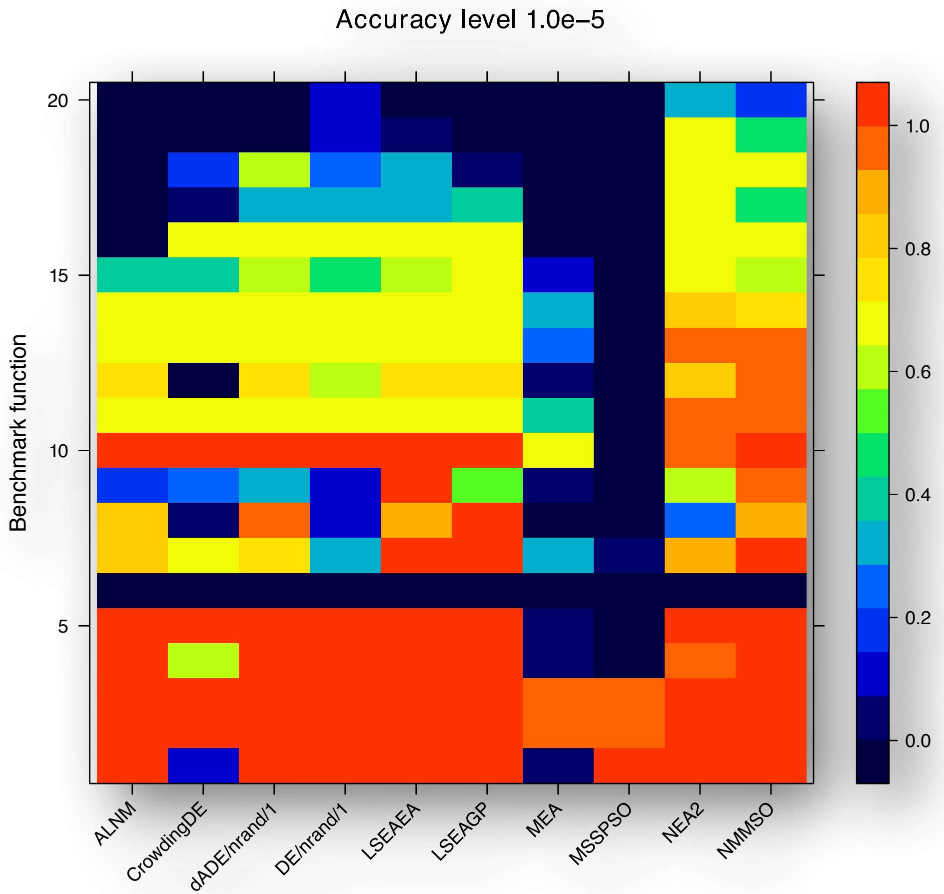


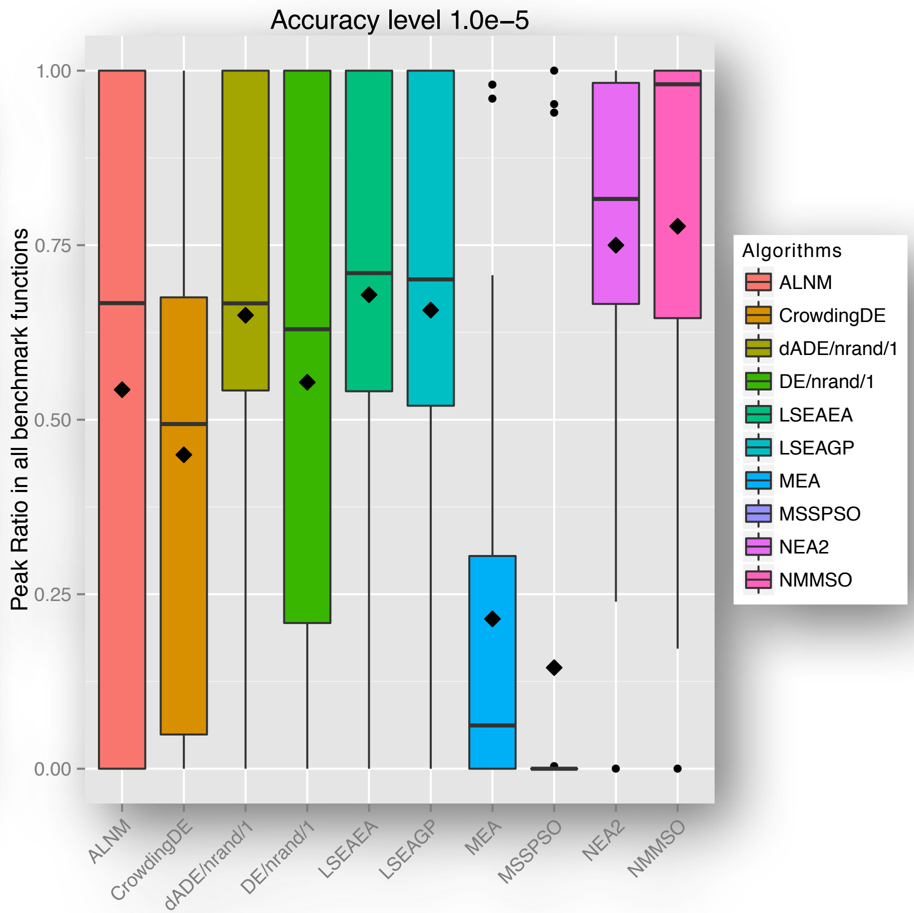
Scenario I: Accuracy level 10^{-4}



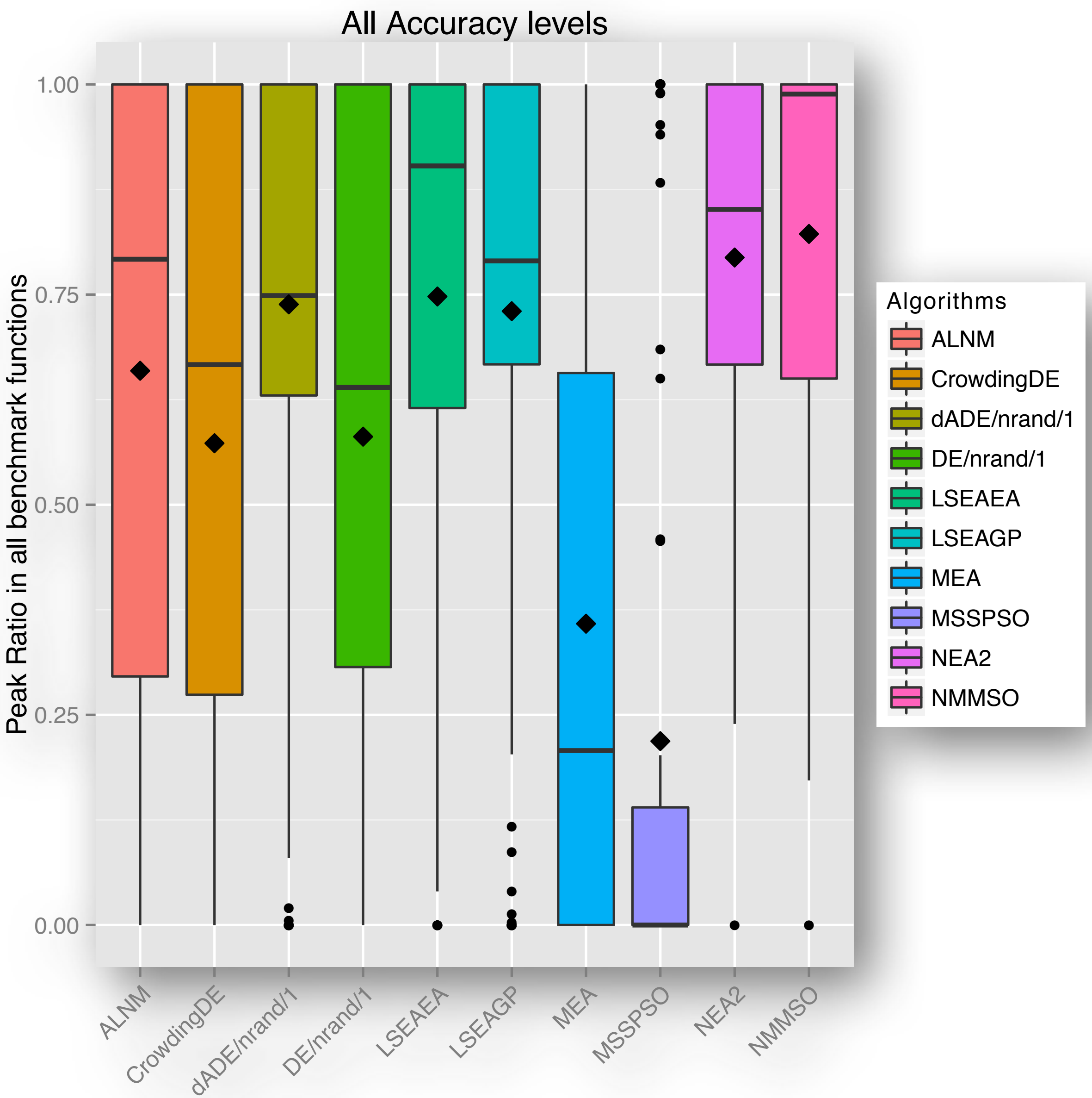


Scenario I: Accuracy level 10^{-5}

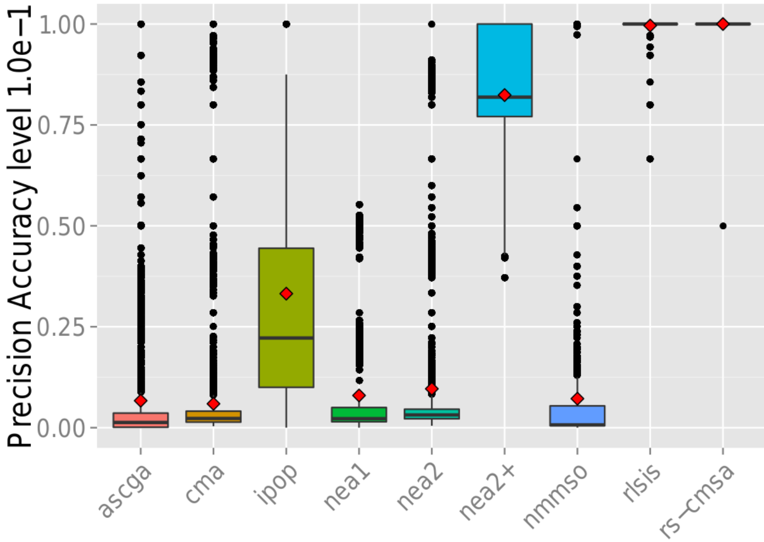


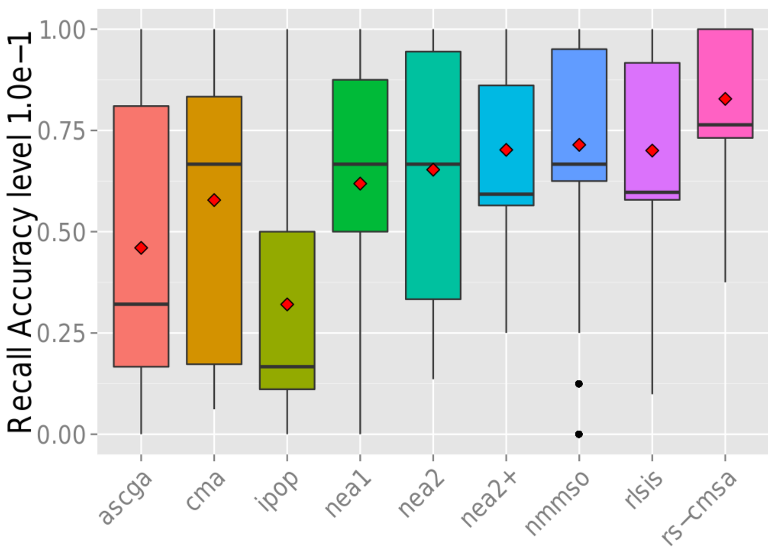


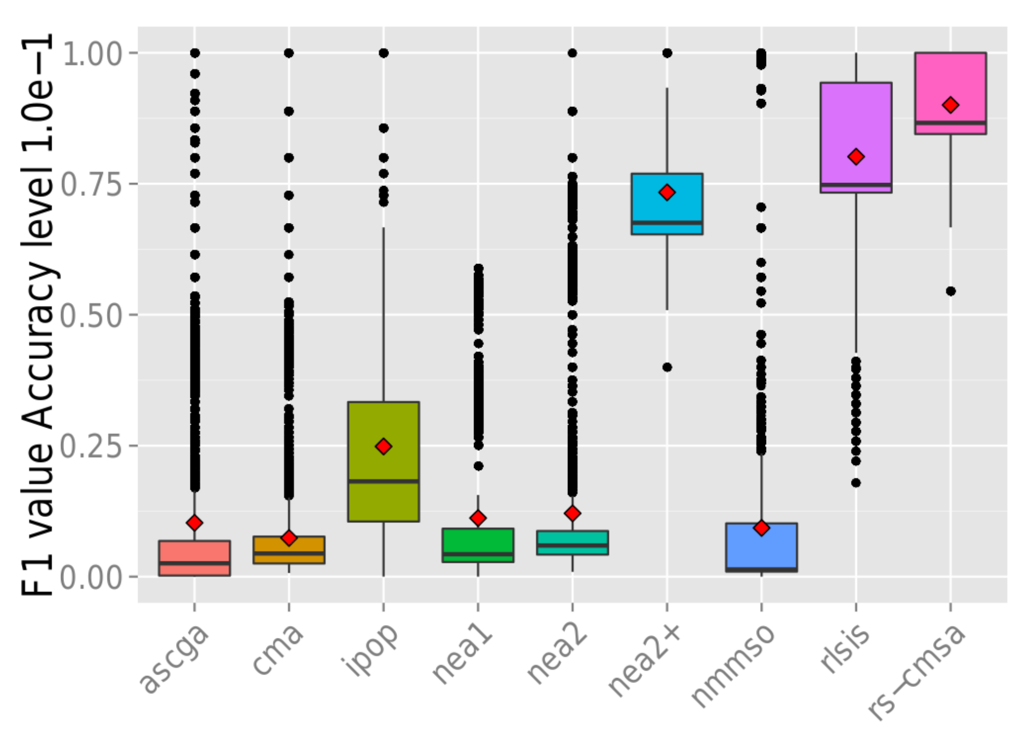
Scenario I: Overall performance



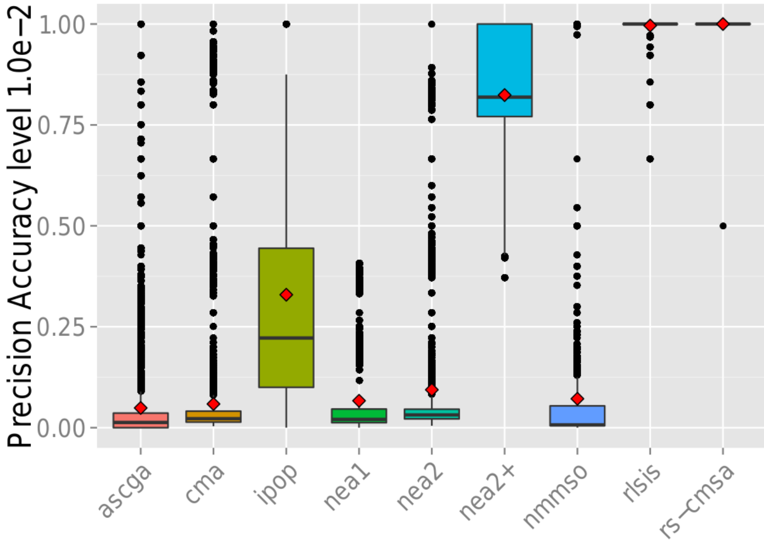
Scenario II: Accuracy level 10^{-1}

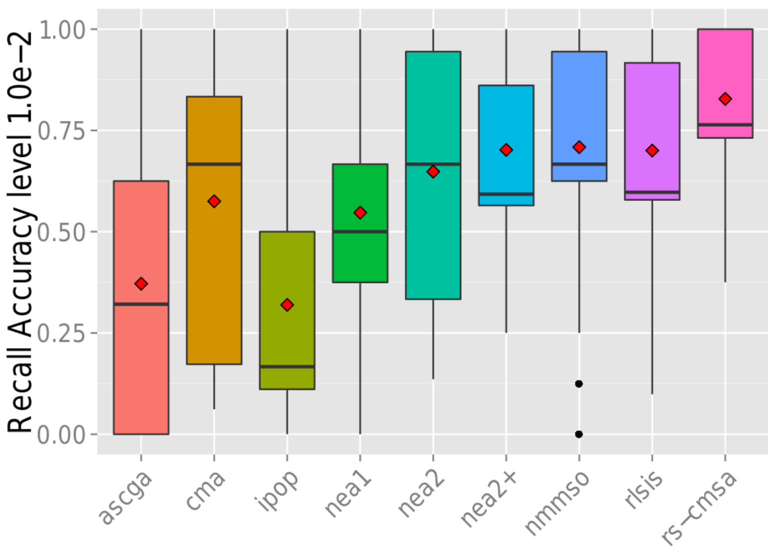


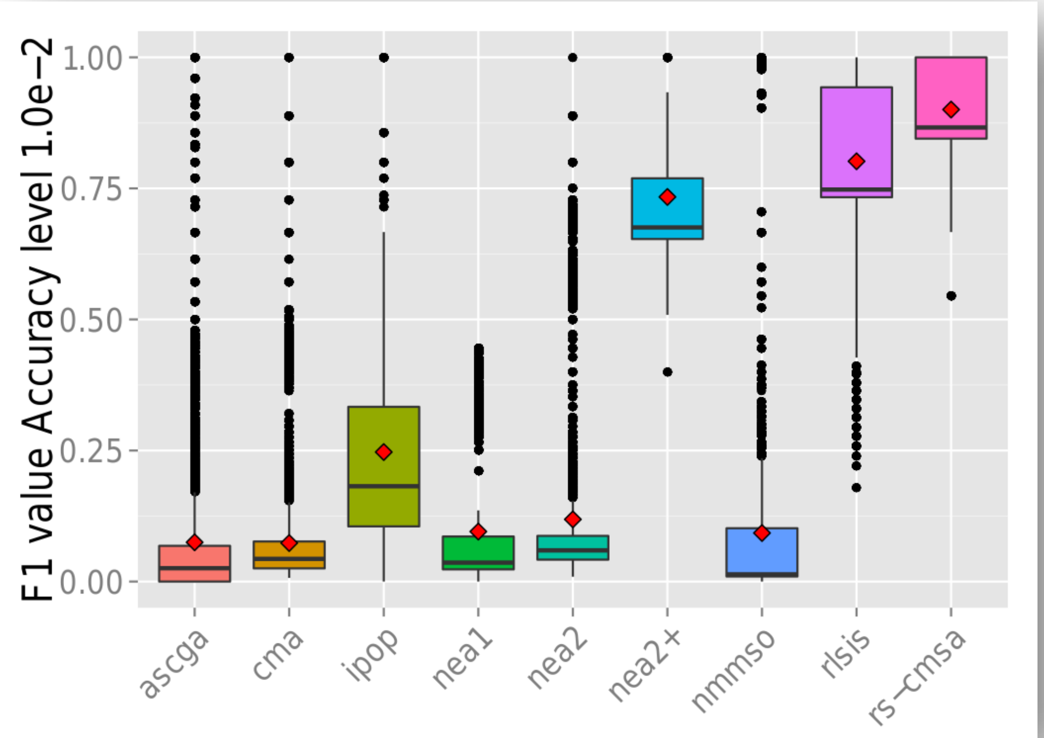




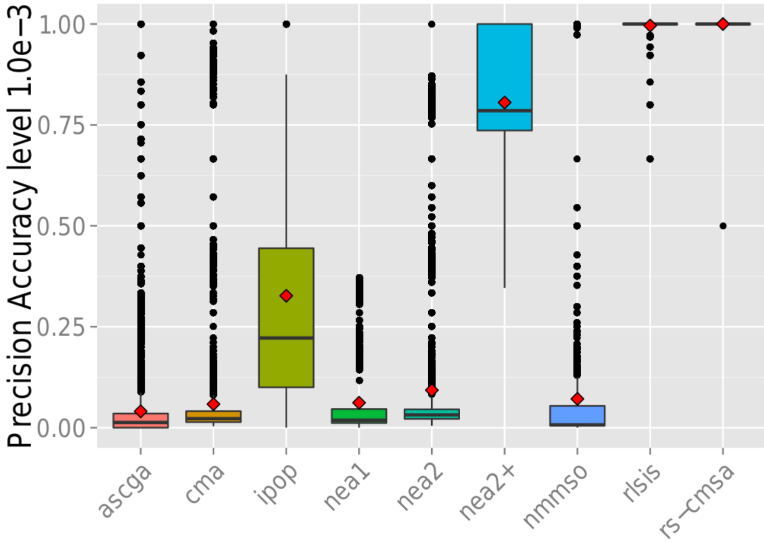
Scenario II: Accuracy level 10^{-2}

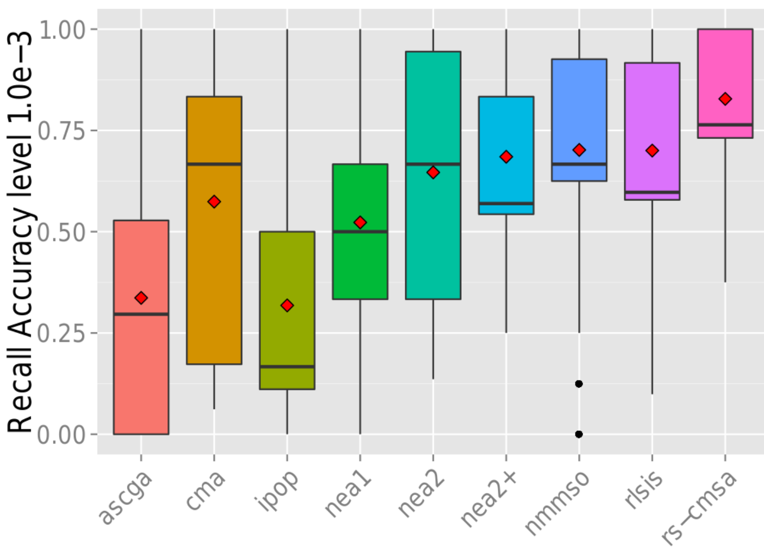


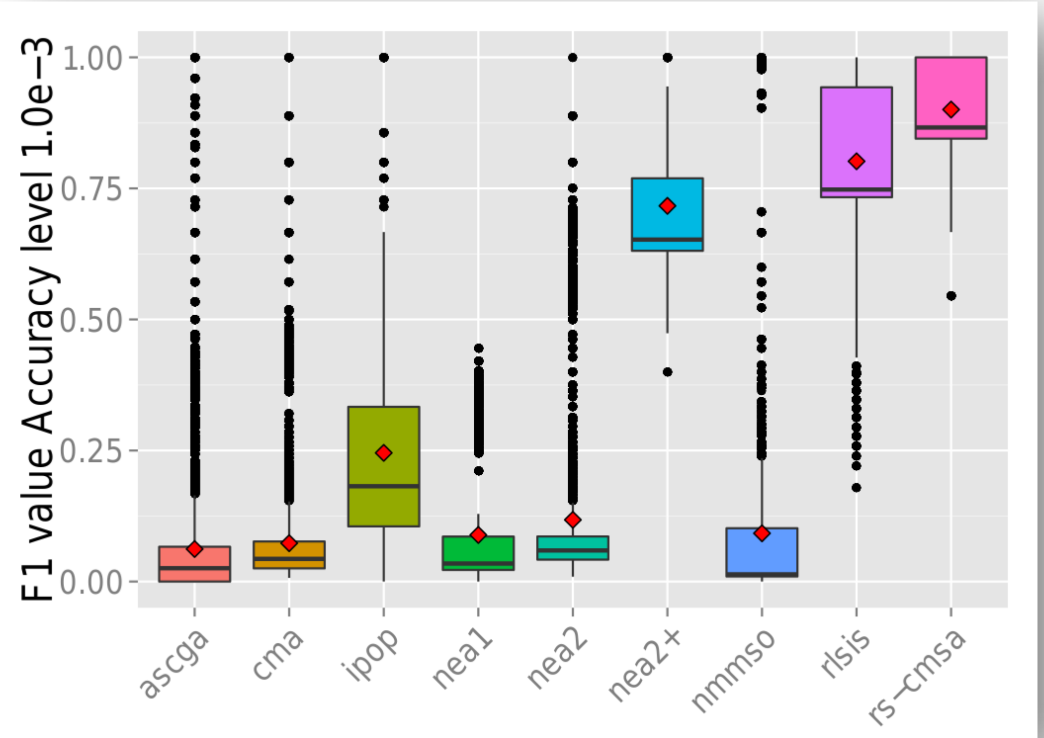




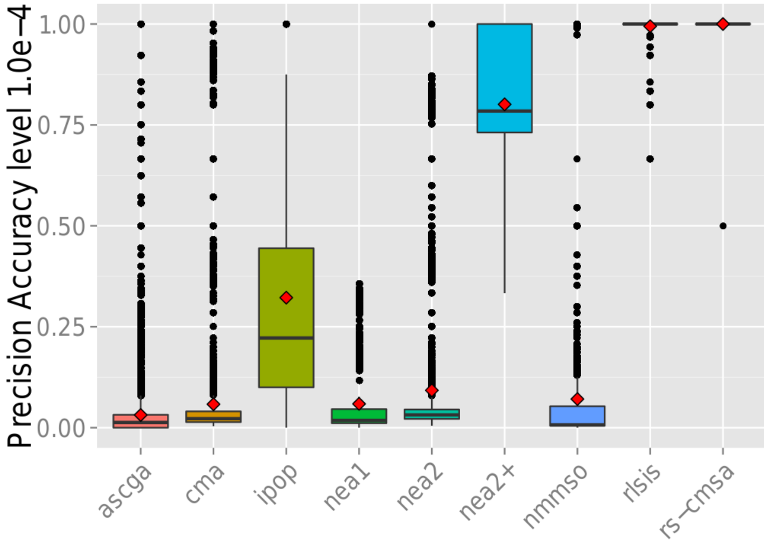
Scenario II: Accuracy level 10^{-3}

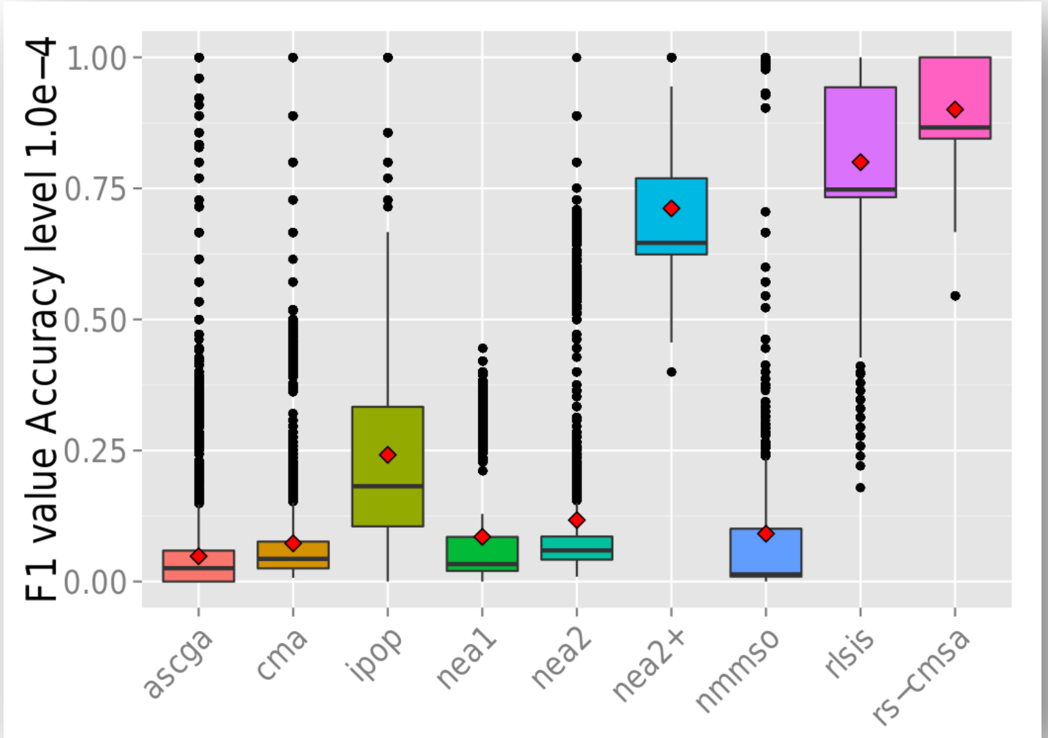


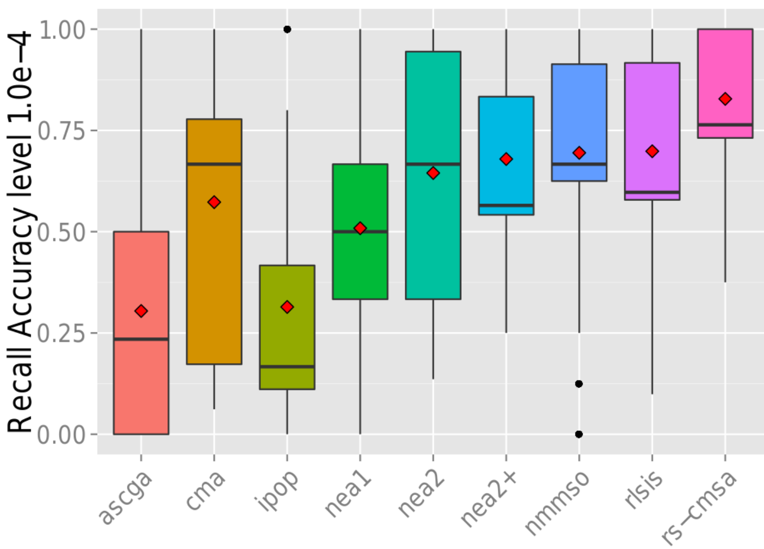




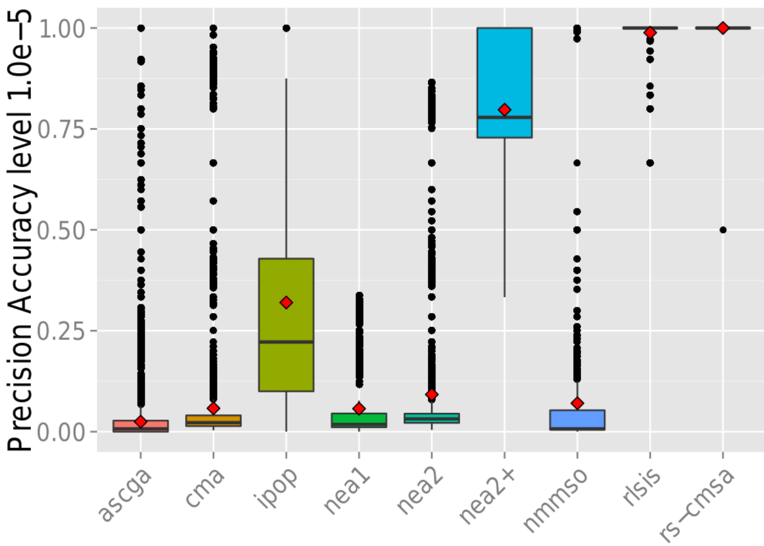
Scenario II: Accuracy level 10^{-4}

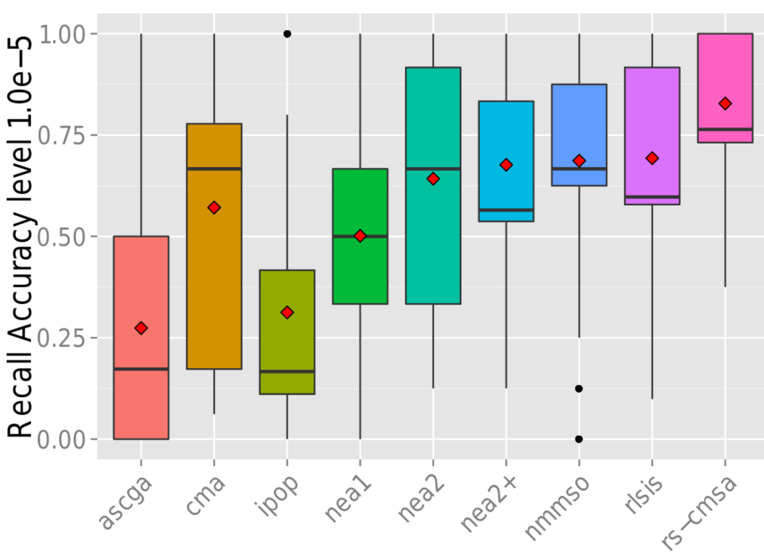


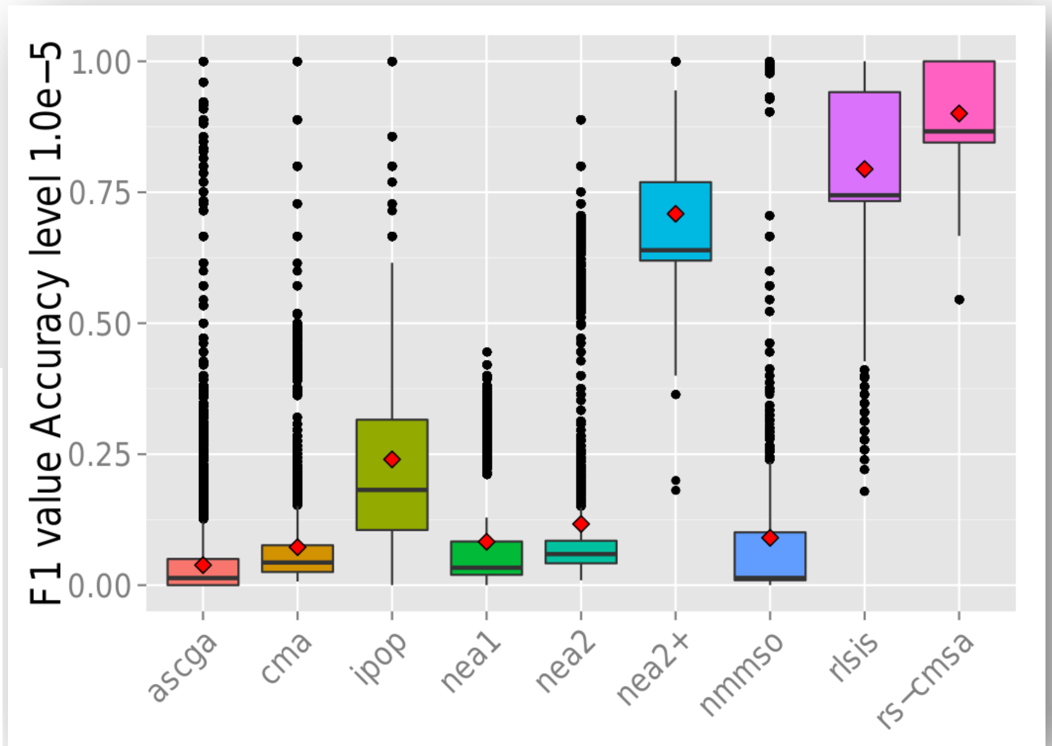




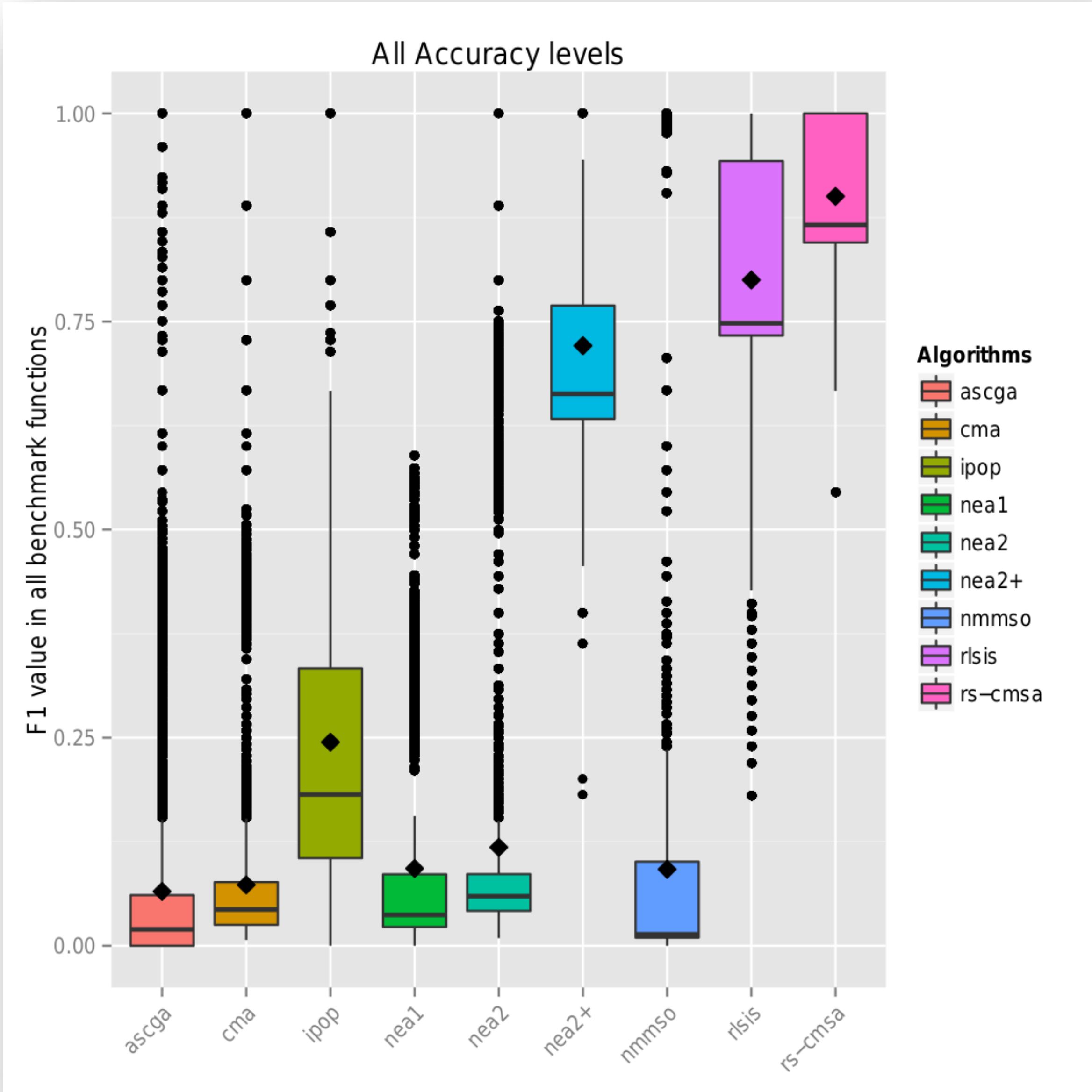
Scenario II: Accuracy level 10^{-5}



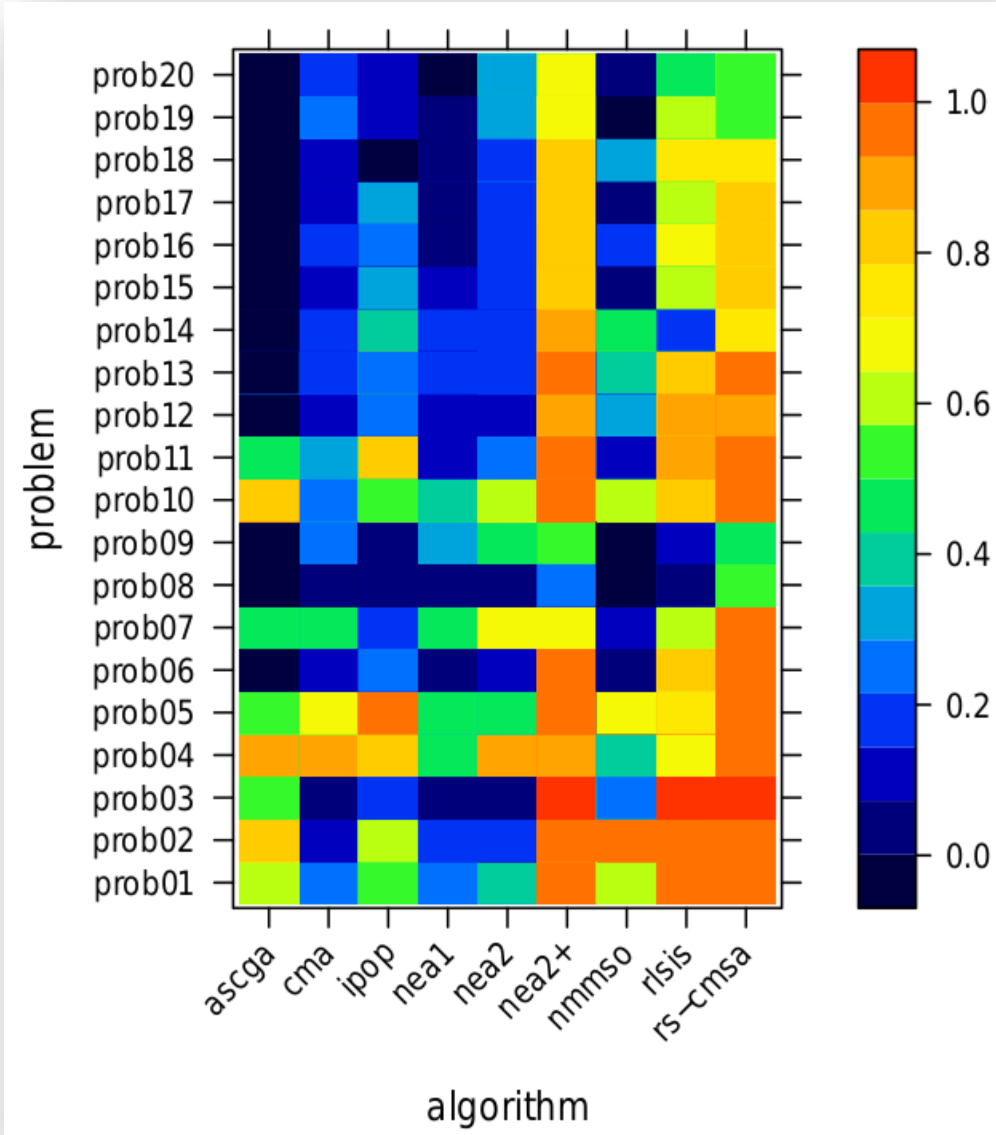




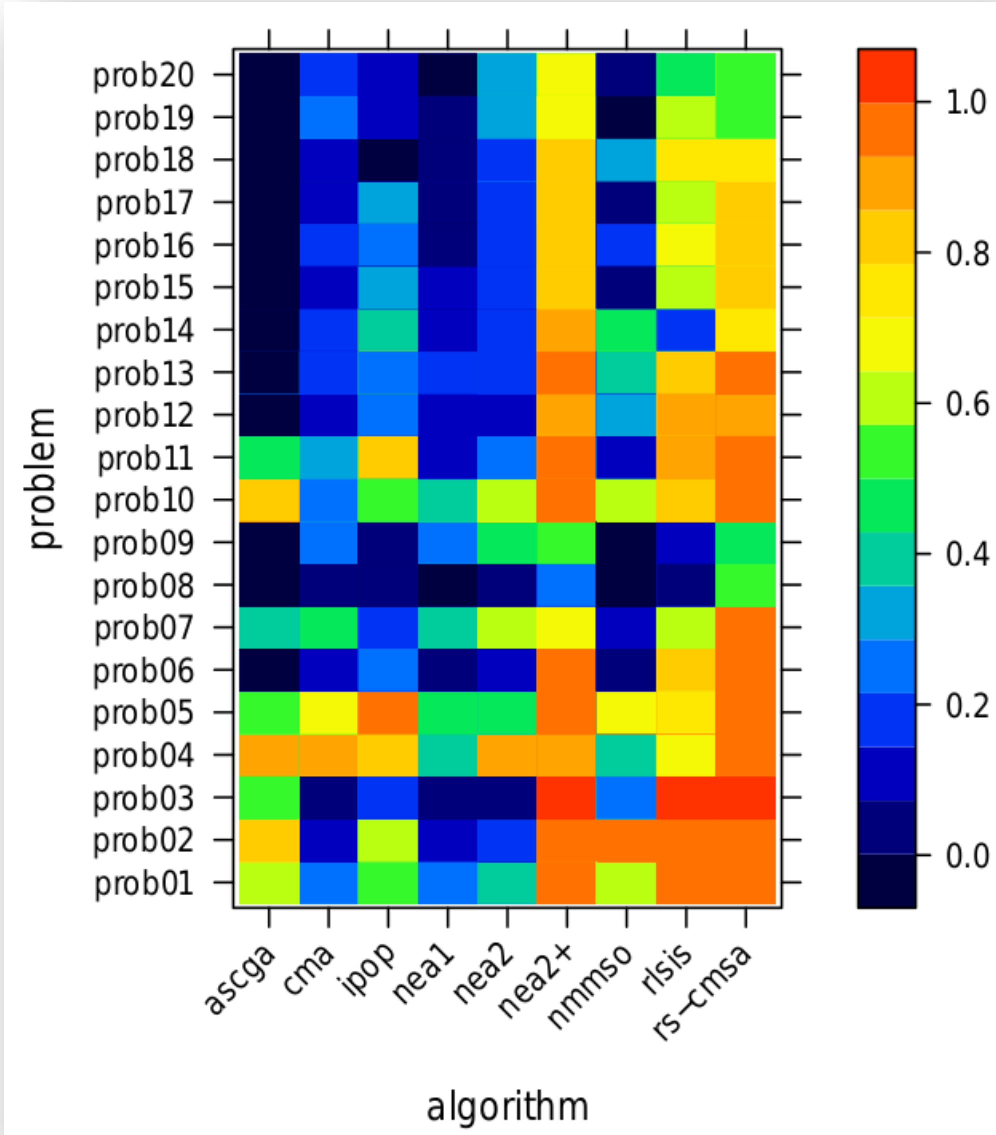
Scenario II: Overall performance



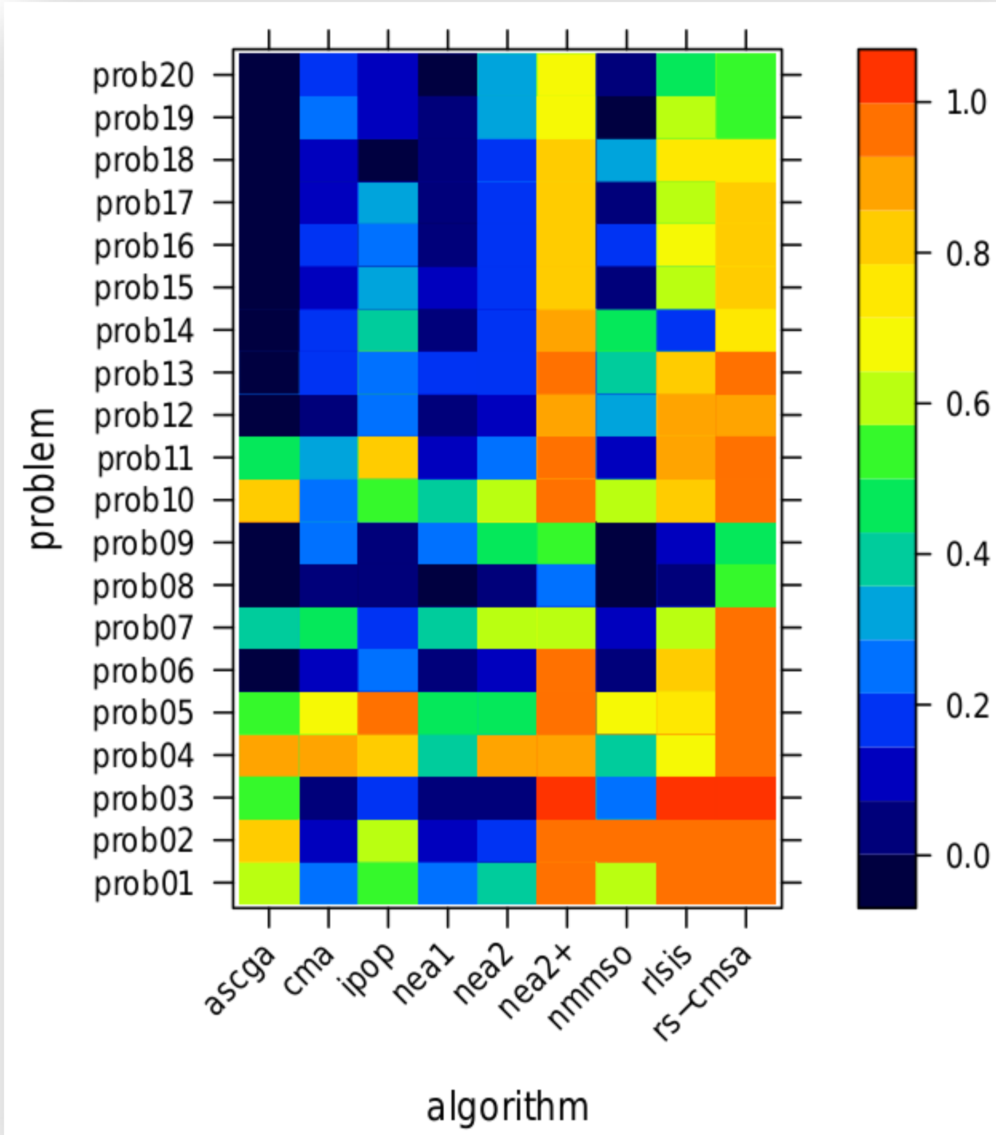
Scenario III: Accuracy level 10^{-1}



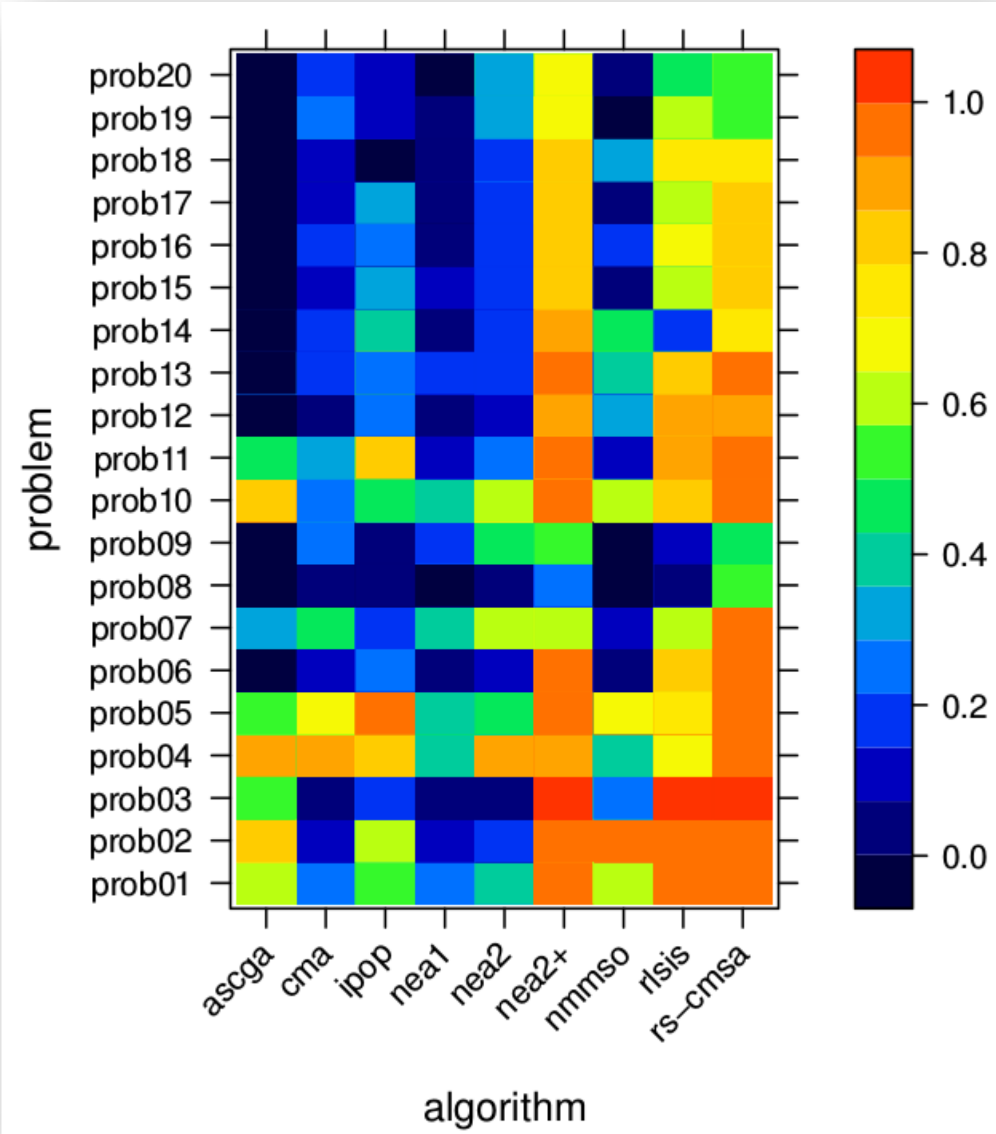
Scenario III: Accuracy level 10^{-2}



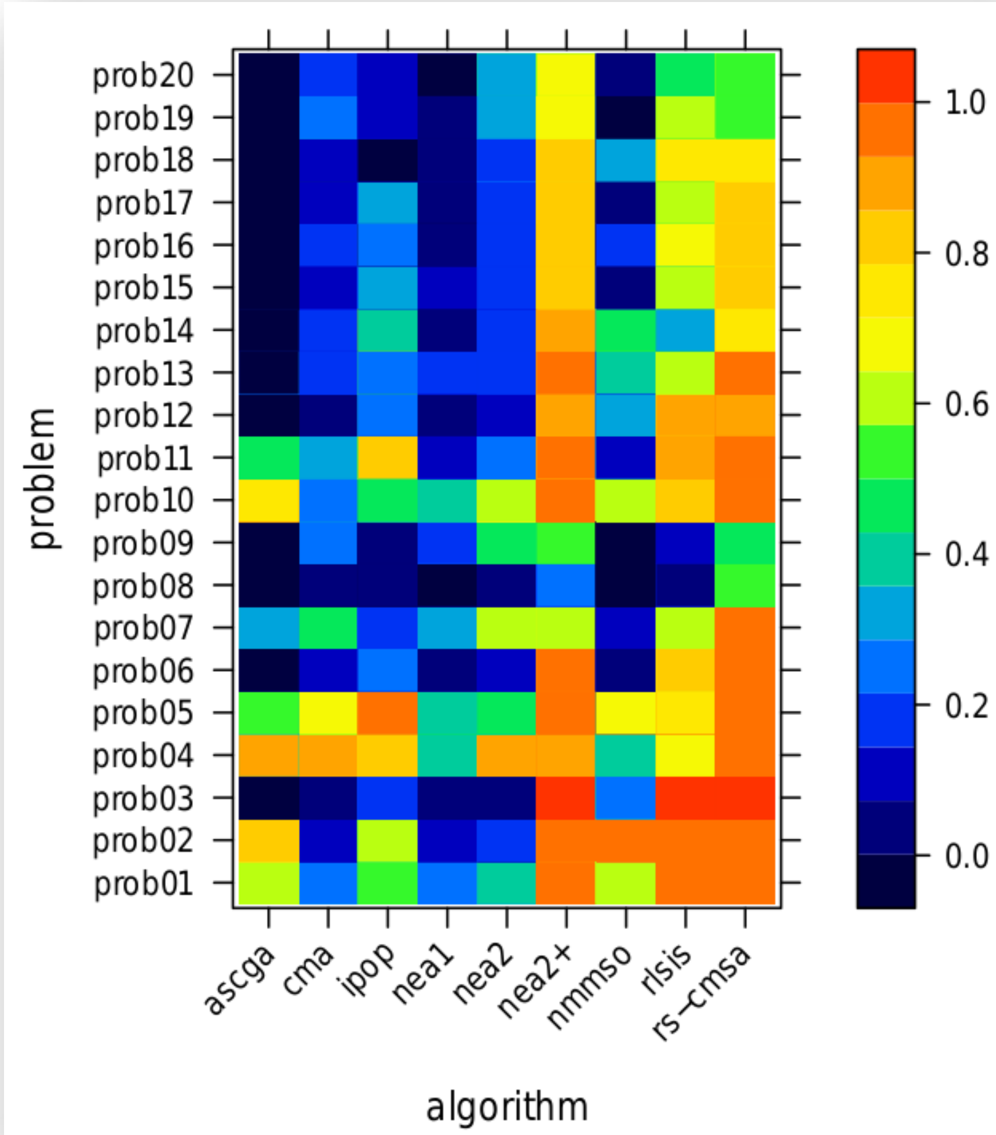
Scenario III: Accuracy level 10^{-3}



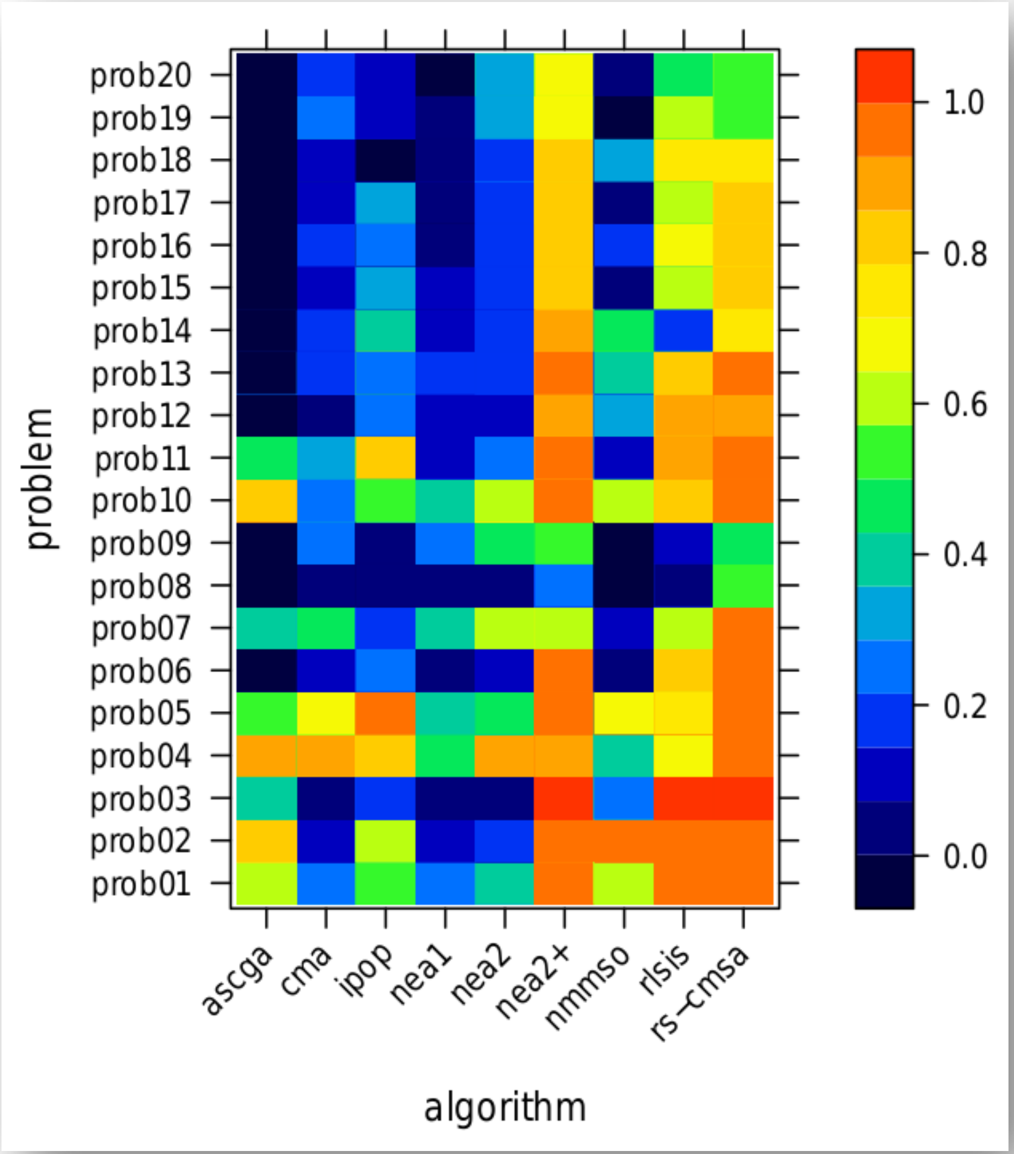
Scenario III: Accuracy level 10^{-4}



Scenario III: Accuracy level 10^{-5}



Scenario III: Overall Performance



GECCO 2016: Overall Performance

Alg.	Sc.I	Rank	Sc.II	Rank	Sc.III	Rank	Mean Rank	Final Rank
ascga	0.349	5	0.065	5	0.236	4	4.666	**5**
nea2+	0.688	4	0.720	3	0.811	2	3.000	**3**
nmmso	0.701	2	0.091	4	0.218	5	3.666	**4**
rlsis	0.698	3	0.799	2	0.663	3	2.666	**2**
rs-cmsa	0.827	1	0.900	1	0.839	1	1.000	**1**

Note: The algorithms have not been fine-tuned for the specific benchmark suite!

GECCO 2016: Winners

2nd
participant

RLSIS

Algorithm:
rlsis: Restarted Local Search with Improved Selection of Starting Points

People:
Simon Wessing

Characteristics:
Algorithm: CMA-ES,
Techniques: initial sampling, restart local search, solution set post-processing.

1st
participant

RS-CMSA-ES

Algorithm:
rs-cmsa-es: Covariance Matrix Self Adaption Evolution Strategy with Repelling Subpopulations

People:
Ali Ahrari, Kalyanmoy Deb and Mike Preuss

Characteristics:
Algorithm: CMA-ES,
Techniques: sub-populations, repelling, solution set post-processing.

3rd
participant

NEA2+

Algorithm:
nea2+: Niching the CMA-ES via Nearest-Better Clustering: First Steps Towards an Improved Algorithm

People:
Mike Preuss

Characteristics:
Algorithm: CMA-ES,
Techniques: Nearest-Better Clustering, solution set post-processing.

Discussion: GECCO 2016

- The competitions gave ***a boost*** to the MMO community
- New ***competitive*** and ***very promising*** approaches
- Key characteristics of the algorithms:
  - **New methodologies**: repelling, restarts, clustering, surrogates, hill-valley approaches, **post-processing**
  - Usage of **local models** to **maintain diversity** and **exploit locally** the neighborhoods
  - **Algorithms**: CMA-ES, GAs, Evolutionary Algorithms, and Multi-swarms

www.epitropakis.co.uk/gecco2017/

GECCO 2017 Competition on Niching Methods for Multimodal Optimization

In this instance of the competition, we continue the successful 2013 and 2015 competitions and rely on an unchanged problem set and similar measures. However, we try out two new aspects regarding the evaluation criteria of the participants:

1. we are including information on the resources (time, function evaluations) needed to find the global optima, not only the fraction of successes within a given time period (number of evaluations), and
2. we take into account the size of the final solution set, and reward small sets that mostly consist of the sought optima only.

Overview

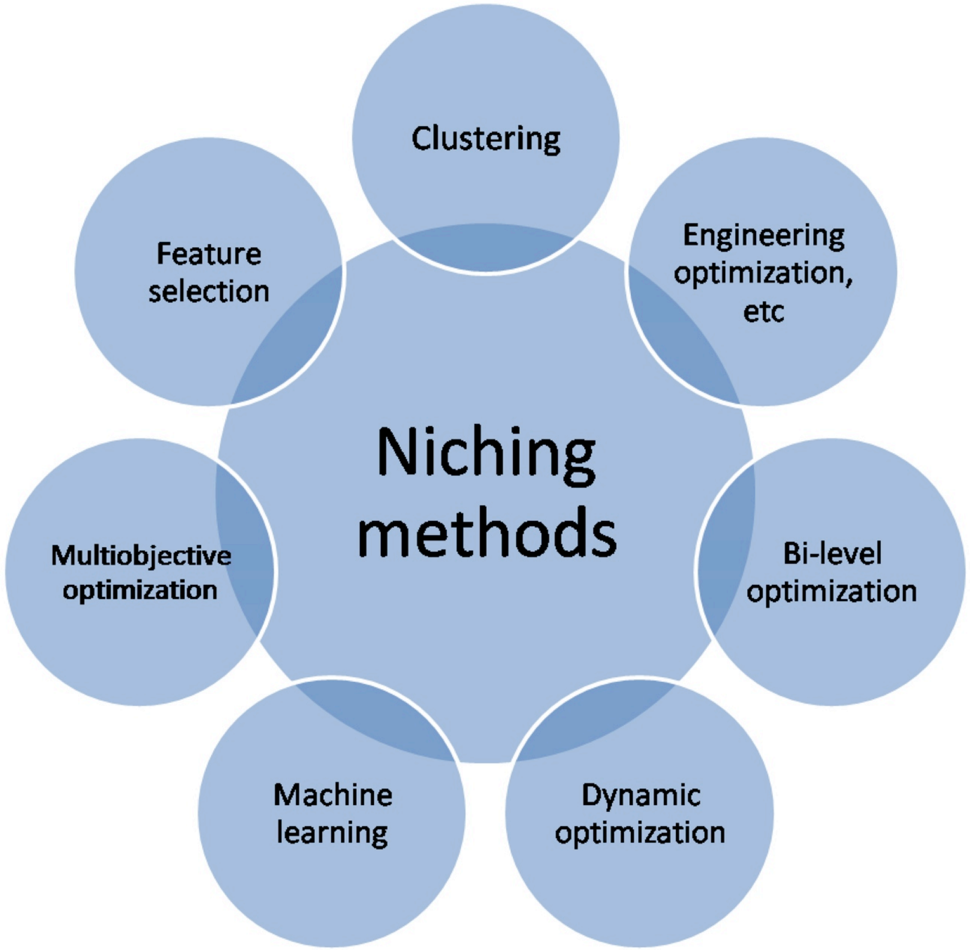
Background photo credit: Berlin Bode-Museum and TV-Tower via photopin (license)

Multi-Modal Optimization

NICHING IN SPECIALIZED TASKS

Niching In Specialized Tasks

- Niching provides ***more effective problem solving*** in a diverse range of tasks
- It also ***benefits from its interaction*** with these areas

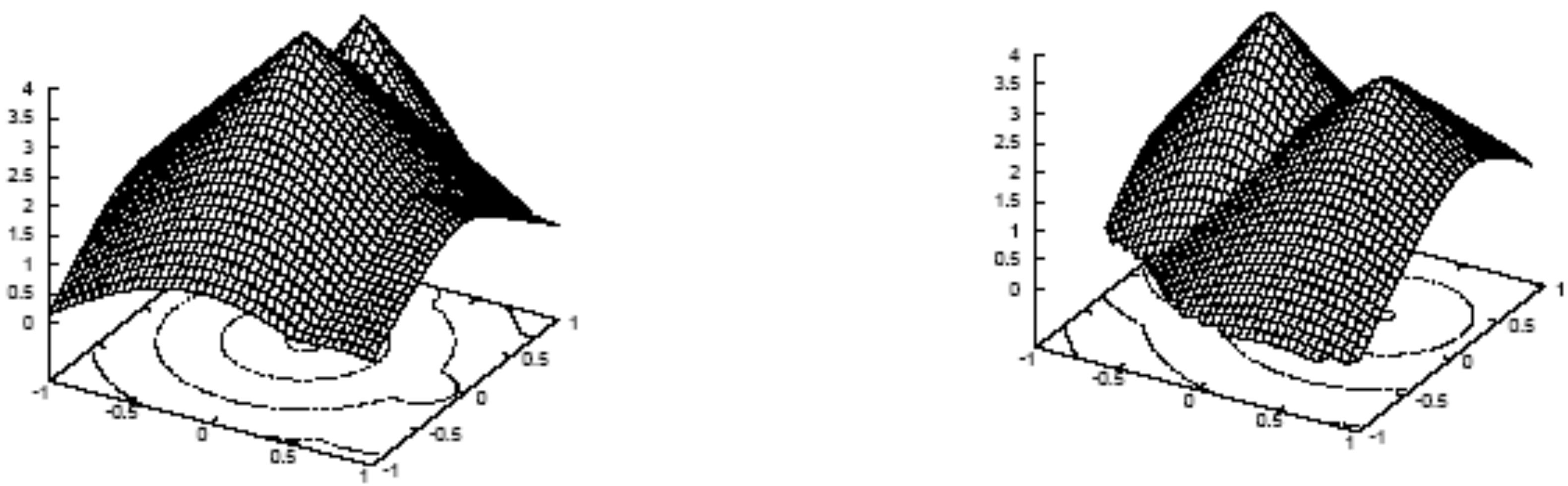


Multi-Modal Optimization

NICHING IN DYNAMIC OPTIMIZATION

SPSO for tracking optima

- In a dynamic environment the goal is to **track** as closely as possible the **dynamically changing optima**
- A useful **strategy to ensure good tracking** of the global optimum **is to maintain multiple species** at all the optima found so far, regardless whether they are globally or locally optimal
- By **maintaining individual species** at each local optimum, it **helps tremendously** in case when such a local optimum turns into a global optimum



X. Li, J. Branke, and T. Blackwell, “Particle swarm with speciation and adaptation in a dynamic environment,” in Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, GECCO ’06. New York, NY, USA: ACM, 2006, pp. 51–58.

Niching in Dynamic Environments

- **Vector-based PSO**: utilize ***directional information*** provided by the particles in a swarm to ***adaptively form niches in parallel*** to track multiple dynamically changing optima
  - I. Schoeman and A. Engelbrecht, “Niching for dynamic environments using particle swarm optimization,” in *SEAL*, 2006, vol. 4247, pp. 134–141.

- **rSPSO**: simple ***regression*** method with ***Speciation-based PSO*** to ***speed up local convergence*** and ***to estimate and predict the positions*** of the changing ***optima***
  - S. Bird and X. Li, *Computational Intelligence in Expensive Optimization Problems*. Springer, 2010, Improving Local Convergence in Particle Swarms by Fitness Approximation Using Regression, pp. 265–293.
  - S. Bird and X. Li, “Using regression to improve local convergence,” in *Evolutionary Computation, 2007. CEC 2007. IEEE Congress on*, Sept 2007, pp. 592–599.

- **Multi-population niching based algorithms**
  - D. Parrott and X. Li, “Locating and tracking multiple dynamic optima by a particle swarm model using speciation,” *IEEE Trans. on Evol. Comput.*, vol. 10, no. 4, pp. 440–458, August 2006.
  - T. Blackwell and J. Branke, “Multi-swarms, exclusion, and anti- convergence in dynamic environments,” *Evolutionary Computation, IEEE Transactions on*, vol. 10, no. 4, pp. 459–472, 2006.
  - X.Li, J.Branke, and T.Blackwell, “Particle swarm with speciation and adaptation in a dynamic environment,” in *Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation*, ser. GECCO ’06. New York, NY, USA: ACM, 2006, pp. 51–58.
  - T. Blackwell, J. Branke, and X. Li, *Swarm Intelligence: Introduction and Applications*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, ch. Particle Swarms for Dynamic Optimization Problems, pp. 193–217.

Multi-Modal Optimization

NICHING IN MULTI-OBJECTIVE OPTIMIZATION

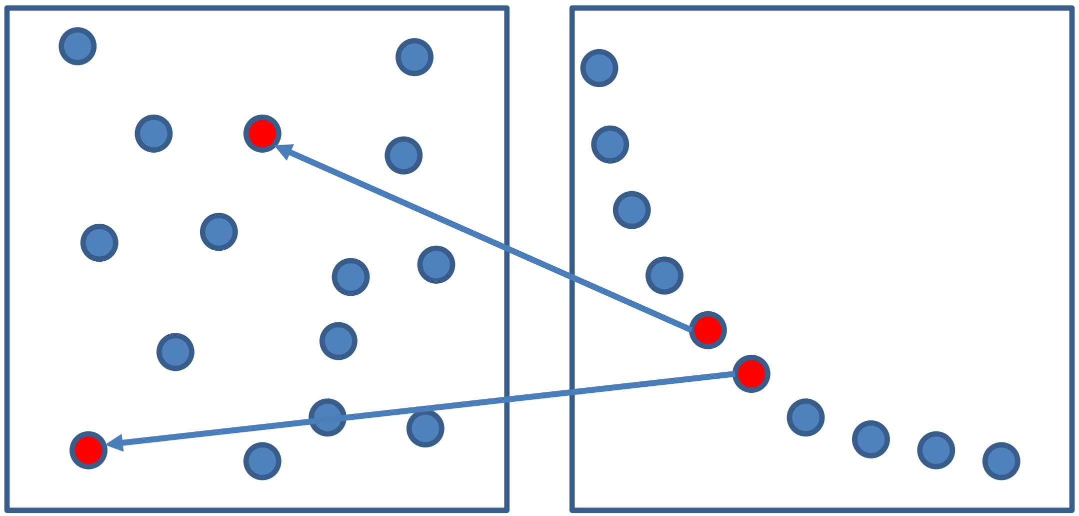
EMO solution diversity

- Although ***diversity maintenance*** is a much common issue in any population-based meta-heuristics, it is possible to use ***niching methods for maintaining solution diversity***
  - Early example is the **Niched-Pareto GA** (NGPA) (Horn, et al., 1994) , which is a multi-objective GA using a variant of fitness sharing to maintain Pareto solution diversity in the objective space
  - Another example is the **crowding distance** metric used **in NSGA-II** (Deb, et al., 2002)
- Much attention has been given to ***maintaining solution diversity in the objective space***
- ***However***, little attention has been given to how to ***maintain solution diversity in the decision space***

J. Horn, N. Nafpliotis, and D. E. Goldberg, "A Niched Pareto Genetic Algorithm for Multiobjective Optimization," in *Proc. of the First IEEE Conference on Evolutionary Computation,* vol. 1. IEEE Service Center, 1994, pp. 82–87.
K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, "A fast and elitist multiobjective genetic algorithm: Nsga-ii," Evolutionary Computation, IEEE Transactions on, vol. 6, no. 2, pp. 182–197, Apr 2002.

Diversity in **both spaces**

- A MOEA (e.g., MOEA Niching-CMA) can produce a much ***more diverse set of efficient solutions*** (i.e., solutions in the **decision space**), ***without sacrificing objective space diversity*** (Shir, et. al. 2009)



Decision space **Objective space**

An example where two solutions that are close in the objective space but their corresponding points in the decision space are further apart

O. M. Shir, M. Preuss, B. Naujoks, and M. Emmerich, "Enhancing decision space diversity in evolutionary multiobjective algorithms," in Proceedings of the 5th International Conference on Evolutionary Multi-Criterion Optimization, ser. EMO '09. Berlin, Heidelberg: Springer- Verlag, 2009, pp. 95–109.

Omni-Optimizer

- Allows **degeneration of NSGA-II** into a single objective multimodal optimization method (i.e., a **niching method**)
- A ***variable space crowding distance*** metric is used to ***encourage distant solutions*** in the decision space to ***remain in the population***
- ***Distant solutions*** with similar or equal objective function values ***will survive***
- **Omni-Optimizer** can ***degenerate to a niching method*** for ***single/multi-objective multi-modal optimization***, capable of finding multiple Pareto-optimal fronts

K. Deb and S. Tiwari, "Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization." European Journal of Operational Research, vol. 185, no. 3, pp. 1062–1087, 2008.

Multi-Modal Optimization

NICHING FOR CLUSTERING AND MACHINE LEARNING

Clustering

- Aim: to ***group data points*** into clusters, such that
  - ***points in each cluster*** have a ***high degree of similarity***
  - ***points in different clusters*** have a ***high degree of dissimilarity***
  - A ***similarity metric*** is often based on some ***distance measure*** between these data points
- Both ***clustering*** and ***niching*** share some ***common features:***
  - ***data points*** seen as ***individuals***
  - ***clusters*** as ***niches***
- ***Clustering methods*** can be used to do ***niching***, and vice versa.

Clustering examples

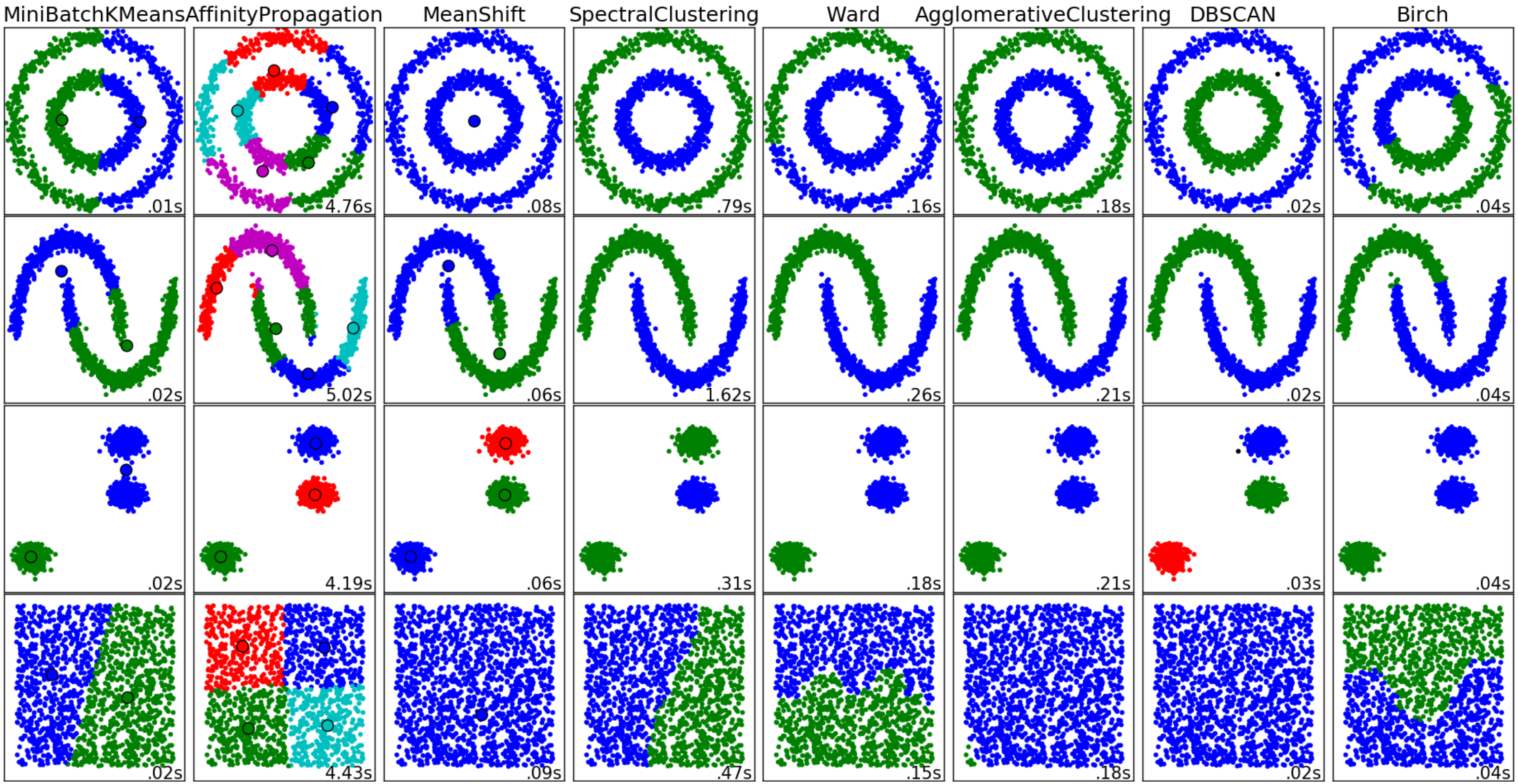


Figure from: http://scikit-learn.org

Clustering for Niching

- ***Clustering methods*** (k-means, NBC, etc) can be used ***to sub-divide the population into clusters*** (or niches)
  - Identification of species/niches
  - Species conserving, topological species conservation
- ***Exploit*** each of ***the niches accordingly***
  - CMA-ES (NEA2), GAs (SCGA, TSC), DE/PSO
- Other ***characteristic examples*** include:
  - Clustering-based niching methods based on dynamic niche sharing, dynamic niche clustering, and dynamic fitness sharing
  - Automatically estimate clustering parameters (such as k in k-means)

X. Yin and N. Germay, “A fast genetic algorithm with sharing scheme using cluster analysis methods in multi-modal function optimization,” in the International Conference on Artificial Neural Networks and Genetic Algorithms, 1993, pp. 450–457.

Clustering for Niching (references)

- X. Yin and N. Germay, “A fast genetic algorithm with sharing scheme using cluster analysis methods in multi-modal function optimization,” in the International Conference on Artificial Neural Networks and Genetic Algorithms, 1993, pp. 450–457.
- M. Preuss. "Niching the CMA-ES via nearest-better clustering." In *Proceedings of the 12th annual conference companion on Genetic and evolutionary computation* (GECCO ’10). ACM, New York, NY, USA, pp. 1711-1718, 2010
- A. Della Cioppa, C. De Stefano, and A. Marcelli, “Where are the niches? dynamic fitness sharing,” *Evolutionary Computation, IEEE Transactions on*, vol. 11, no. 4, pp. 453–465, Aug 2007.
- B. L. Miller and M. J. Shaw, “Genetic algorithms with dynamic niche sharing for multimodal function optimization,” in *Proceedings of the 1996 IEEE International Conference on Evolutionary Computation*, May 1996, pp. 786–791.
- J. Gan and K. Warwick, “Dynamic niche clustering: a fuzzy variable radius niching technique for multimodal optimisation in gas,” in *Proc. of the 2001 Congress on Evolutionary Computation*. IEEE Press, 2001, pp. 215–222.
- J.-P. Li, M. E. Balazs, G. T. Parks, and P. J. Clarkson, “A species conserving genetic algorithm for multimodal function optimization,” *Evol. Comput.*, vol. 10, no. 3, pp. 207–234, 2002.
- C. Stoean, M. Preuss, R. Stoean, and D. Dumitrescu, “Multimodal optimization by means of a topological species conservation algorithm,” *Evolutionary Computation, IEEE Transactions on*, vol. 14, no. 6, pp. 842–864, Dec 2010.
- D. K. Tasoulis, V. P. Plagianakos, and M. N. Vrahatis, “Clustering in evolutionary algorithms to efficiently compute simultaneously local and global minima,” in *2005 IEEE Congress on Evolutionary Computation*, vol. 2, Sept 2005, pp. 1847–1854 Vol. 2.
- V. P. Plagianakos, “Unsupervised clustering and multi-optima evolutionary search,” in *2014 IEEE Congress on Evolutionary Computation (CEC)*, July 2014, pp. 2383–2390.

Niching for clustering (I)

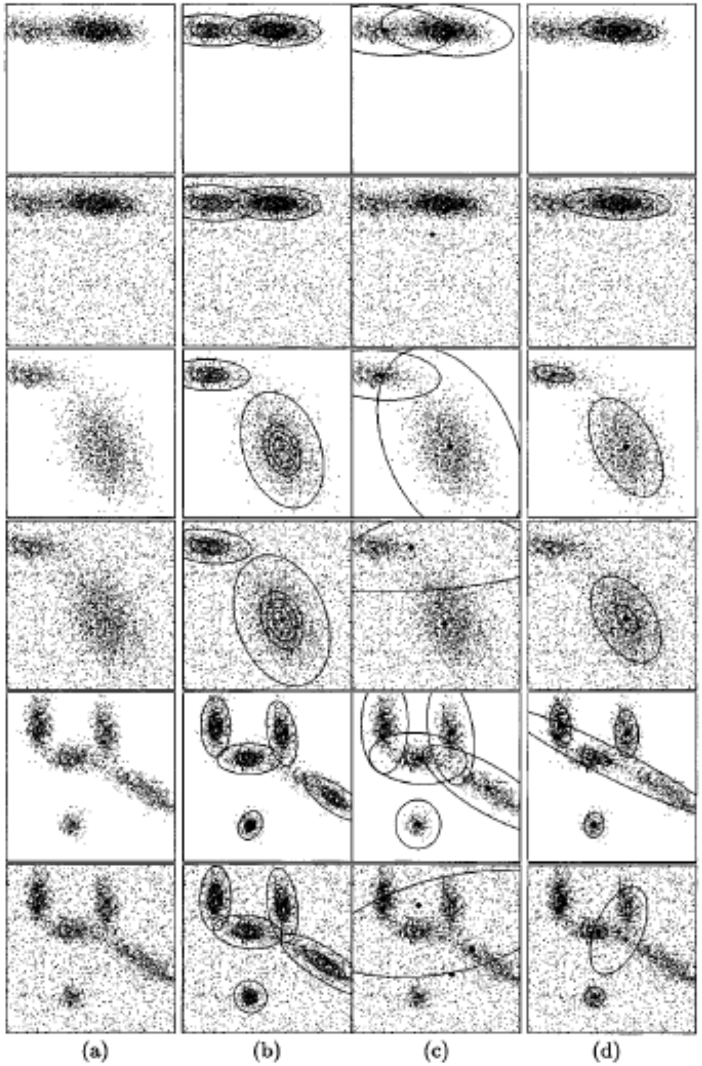
- A **clustering** problem can be **formulated as a multi-modal optimization** problem, and be handled by a niching method.
- We can define a **density-based fitness function** that would reach a ***maximum at every good cluster center***.
- The value of the fitness function will be high for points falling within the boundary of a cluster, and low for points falling outside of the cluster.
- The results in Nasraoui, et al. (2005) suggested that the ***niching approach for clustering*** to be ***less prone than non-niching techniques*** to premature convergence, noise, and initialization.

O. Nasraoui, E. Leon, and R. Krishnapuram, "Unsupervised niche clustering: Discovering an unknown number of clusters in noisy data sets," in Evolutionary Computation in Data Mining, ser. Studies in Fuzziness and Soft Computing, A. Ghosh and L. Jain, Eds. Springer Berlin Heidelberg, 2005, vol. 163, pp. 157–188.

Niching for clustering (II)

Nasraoui, *et al.* (2005) proposed an ***Unsupervised Niche Clustering algorithm*** *(*UNC) and evaluated its performance under different conditions related to cluster size, density, noise contamination, orientation, and number of clusters. Their results were presented on the 9 noisy data sets (see the figure on the right):

(a) original data set;
(b) results of UNC;
(c) results of K-means with pre-specified correct *c* (the number of clusters);
(d) results of PCM with pre-specified correct *c*;



O. Nasraoui, E. Leon, and R. Krishnapuram, "Unsupervised niche clustering: Discovering an unknown number of clusters in noisy data sets," in Evolutionary Computation in Data Mining, 2005, vol. 163, pp. 157–188.

Feature Selection

- The aim of feature selection is ***to choose features that allow us to discriminate patterns*** belonging to different classes
- Feature selection algorithms are generally classified into
  - **wrapper methods** make use of a ***learning classifier***'*s* performance ***to evaluate the suitability of the feature subset***
  - **filter methods** treat the selection of feature ***subsets as a pre-processing step***, independent from the learning classifier.

Niching for Feature Selection

- Optimal ***subset of features*** might ***not be unique***
  - Merit for ***obtaining all*** such optimal ***subsets*** before the final choice
- Different ***optimal subsets*** of features are considered as ***different optima on a multi-modal fitness landscape*** (searched by niching methods)
- Example ***representation of a subset*** of the selected features: ***binary string***
  - **1** indicates that the *i*-th feature is ***included*** in the subset, otherwise (**0**) the feature is ***excluded***
- Evaluate the ***goodness of a subset***:
  - the binary string is fed into a ***learning classifier*** (e.g., neural network)
- ***Fitness function*** takes into account:
  - the ***classifier accuracy*** term and
  - the ***penalty*** for selecting a ***large number of features***

F. Brill, D. Brown, and W. N. Martin, “Fast generic selection of features for neural network classifiers,” Neural Networks, IEEE Transactions on, vol. 3, no. 2, pp. 324–328, Mar 1992.

Niching for Machine Learning

- Machine Learning (ML) plays an ***increasingly important role in data analytics*** these days:
  - ***predict by learning from data***
- Many ***real-world problems*** are often too large and complex **to be *solved by a single machine learning*** model
- An effective approach may be to employ an ***ensemble of learning models***, each specializing in solving a subtask of a much larger problem.
- ***Meta-heuristic algorithms*** can be used to ***evolve a population of ML models***
  - e.g., an ensemble of neural networks, or a set of knowledge rules.

Y. Liu and X. Yao, "Simultaneous training of negatively correlated neural networks in an ensemble," IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics), vol. 29, no. 6, pp. 716–725, Dec 1999.
M. L. Wong and K. S. Leung, Data Mining Using Grammar-Based Genetic Programming and Applications. Norwell, MA, USA: Kluwer Academic Publishers, 2000.

Evolving neural network ensembles

- ***Niching techniques*** (Speciation) used ***to evolve a diverse but accurate set of specialist modules***, which can be then combined to perform learning tasks
- Evolutionary Ensembles with Negative Correlation Learning (***EENCL***)[Liu *et al*. (2000)] ***automatically determine the number of individual neural networks in an ensemble***
- ***Motivation***: A population contains more information than a single individual
- **Fitness sharing** was adopted to ***promote diversity in the ensemble***:
  - If one training example is learnt correctly by ***n*** individual neural networks, then each of these ***n*** neural networks receives a fitness value 1/***n***, and the remaining neural networks in the ensemble receive zero fitness
  - This procedure is repeated for all examples in the training set
  - The final fitness of an individual is determined by summing up its fitness values over all training examples

Y. Liu, X. Yao, and T. Higuchi, “Evolutionary ensembles with negative correlation learning,” IEEE Transactions on Evolutionary Computation, vol. 4, no. 4, pp. 380–387, Nov 2000.

Learning multiple rules from data

- In ***data mining***: meta-heuristics can be used to ***extract knowledge such as rules*** and use these rules to solve classification problems
  - the ***Michigan approach***: where each individual encodes a single rule, and
  - the ***Pittsburgh approach*** where each individual represents multiple rules, i.e., a rule set.
- Since it is often difficult to capture the knowledge of a data set by a single rule, ***multiple rules are often required***
- **Niching methods** can be used to ***evolve multiple different good individuals*** that are required ***to produce a rule set***:
  - See for example, in the idea of "*token competition*" [Wong and Leung (2000)]

M. L. Wong and K. S. Leung, Data Mining Using Grammar-Based Genetic Programming and Applications. Norwell, MA, USA: Kluwer Academic Publishers, 2000.
K. Tan, Q. Yu, C. Heng, and T. Lee, "Evolutionary computing for knowledge discovery in medical diagnosis," Artificial Intelligence in Medicine, vol. 27, pp. 129–154, 2003.

Multi-Modal Optimization

IEEE CIS TASK FORCE ON MULTI-MODAL OPTIMIZATION

IEEE CIS Taskforce on MMO

- **The key objective** is to **promote research on multi-modal optimization**, including its development, education and understanding of sub topic areas of multi-modal optimization. Further info: http://www.epitropakis.co.uk/ieee-mmo/
- **Current chair**: Michael G. Epitropakis (Lancaster University, UK).
- **Vice-Chairs**: Andries Engelbrecht (University of Pretoria, South Africa), and Xiaodong Li (RMIT University, Australia).
- **Members**: Carlos A. Coello Coello, Kalyanmoy Deb, Andries Engelbrecht, Michael G. Epitropakis, Jonathan Fieldsend, Jian-Ping Li, Xiaodong Li, Jonathan Mwaura, Konstantinos Parsopoulos, Vassilis Plagianakos, Mike Preuss, Bruno Sareni, Ofer M. Shir, Patrick Siarry, P. N. Suganthan, Michael N. Vrahatis, Simon Wessing, Xin Yao.

IEEE CIS Taskforce on MMO

IEEE Computational Intelligence Society
MIMICKING NATURE FOR PROBLEM SOLVING

- **Past and planned activities:**
  - **IEEE CEC** 2010, 2013, 2015, 2016 and 2017 **special sessions** and/or **competitions** on “Niching Methods for Multimodal Optimization”.
  - **GECCO** 2016, 2017 **competitions** on “Niching Methods for Multimodal Optimization”.
  - International **Workshop** on "Advances in Multimodal Optimization", **PPSN** 2014, **PPSN** 2016.
  - **Tutorials** at **WCCI** 2016, **PPSN** 2014, **CEC** 2017.
  - More activities to come soon…
    - A repository for related material, publications and source code.

Summary

- **Niching methods** have been studied for the past few decades, and **now experience a revival**, as more people from diverse backgrounds find its relevance in their own disciplinary areas.
- **Niching methods** can be **developed using other meta-heuristics**, apart from evolutionary algorithms.
- **Niching** has its **application in many problem solving domains**, e.g., dynamic optimization and multi-objective optimization.
- **A good starting point for new comers**: several **survey papers** are available, plus recently a **new book** by Mike Preuss.
- **Many open research questions and challenges to be addressed**.
- **Many** possible **real-world applications** of niching methods.

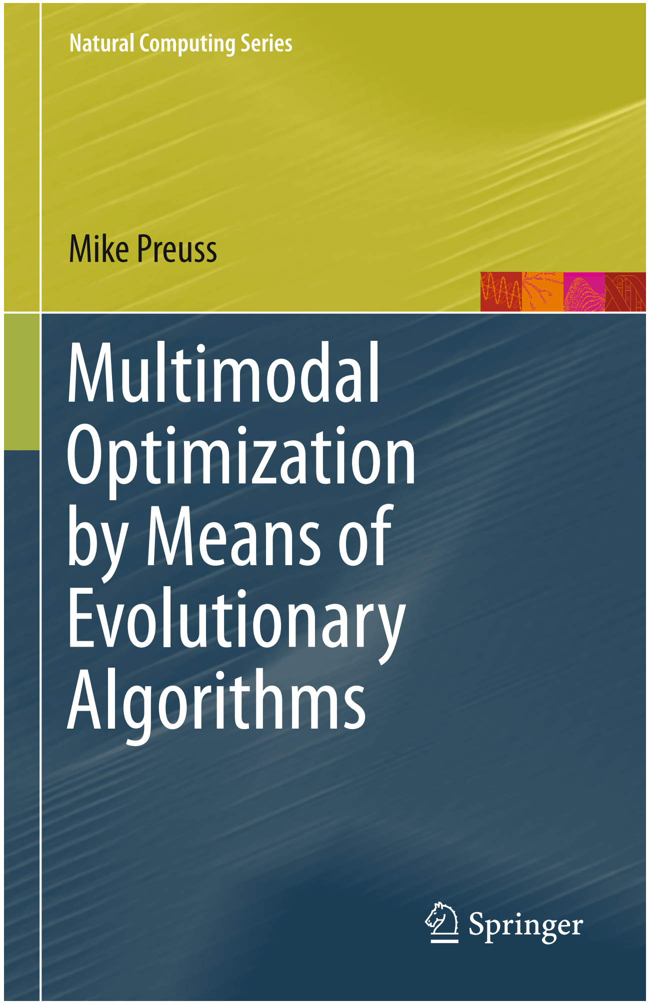
S. Das, S. Maity, B.-Y. Qu, and P. Suganthan, “Real-parameter evolutionary multimodal optimization - a survey of the state-of-the-art,” Swarm and Evolutionary Computation, vol. 1, pp. 71–88, June 2011.
O. Shir, “Niching in evolutionary algorithms,” Handbook of Natural Computing: Theory, Experiments, and Applications, pp. 1035–1069, 2012.
X. Li, “Developing niching algorithms in particle swarm optimization,” in Handbook of Swarm Intelligence, ser. Adaptation, Learning, and Optimization, B. Panigrahi, Y. Shi, and M.-H. Lim, Eds. Springer Berlin Heidelberg, 2011, vol. 8, pp. 67–88.
M. Preuss, Multimodal Optimization by Means of Evolutionary Algorithms, ser. Natural Computing Series. Springer International Publishing, 2016.

New book on MMO!!!

- Describes state of the art in algorithms, measures and test problems
- Approaches multimodal optimization algorithms via model-based simulation and statistics
- Valuable for practitioners with real-world black-box problems

DOI:10.1007/978-3-319-07407-8

http://www.springer.com/gp/book/9783319074061



Recent Survey on MMO @ IEEE TEVC

Seeking Multiple Solutions: An Updated Survey on Niching Methods and Their Applications

Xiaodong Li, Michael G. Epitropakis, Kalyanmoy Deb, Andries Engelbrecht

***Abstract*—Multi-Modal Optimization (MMO) aiming to locate multiple optimal (or near-optimal) solutions in a single simulation run has practical relevance to problem solving across many fields. Population-based meta-heuristics have been shown particularly effective in solving MMO problems, if equipped with specifically-designed diversity-preserving mechanisms, commonly known as *niching* methods. This paper provides an updated survey on niching methods. The paper first revisits the fundamental concepts about niching and its most representative schemes, then reviews the most recent development of niching methods, including novel and hybrid methods, performance measures, and benchmarks for their assessment. Furthermore, the paper surveys previous attempts at leveraging the capabilities of niching to facilitate various optimization tasks (e.g., multi-objective and dynamic optimization) and machine learning tasks (e.g., clustering, feature selection, and learning ensembles). A list of successful applications of niching methods to real-world problems is presented to demonstrate the capabilities of niching methods in providing solutions that are difficult for other optimization methods to offer. The significant practical value of niching methods is clearly exemplified through these applications. Finally, the paper poses challenges and research questions on niching that are yet to be appropriately addressed. Providing answers to these questions is crucial before we can bring more fruitful benefits of niching to real-world problem solving.**

***Index Terms*—Niching methods, Multi-modal optimization, Meta-heuristics, Multi-solution methods, Evolutionary computation, Swarm intelligence.**

The goal of locating multiple optimal solutions in a single run by niching methods contrasts sharply with the goal of a classic optimization method [2], which usually starts from an initial single point and iteratively improving it, before arriving at one final solution. Since it is not guaranteed that starting at different initial points will arrive at different solutions with multiple runs, classic optimization methods are not suited for the purpose of locating multiple solutions. This goal is also different from the usual single-optimum seeking mechanism employed by a standard meta-heuristic method. In literature, sometimes "*multi-modal optimization*" also refers to seeking a single optimum on a *multi-modal* fitness landscape. To avoid this confusion and to be more precise, in this paper we also refer to niching methods as "*multi-solution*" methods.

Classic niching methods, including *fitness sharing* [3] and *crowding methods* [4], were developed in the early 70s and 80s. In subsequent years, many niching methods have been proposed. Some representative examples include *deterministic crowding* [5], *derating* [6], *restricted tournament selection* [7], *parallelization* [8], *clustering* [9], *stretching and deflation* [10], [11], and *speciation* [12], [13]. Initially, niching methods were developed for Evolutionary Algorithms (EAs). However, recently niching methods were also developed for other meta-heuristic optimization algorithms [14], such as Evolu-

X. Li; M. Epitropakis; K. Deb; A. Engelbrecht, "Seeking Multiple Solutions: an Updated Survey on Niching Methods and Their Applications," in ***IEEE Transactions on Evolutionary Computation***, doi: 10.1109/TEVC.2016.2638437

Thank you!

Questions?

Closely follow

IEEE CIS Taskforce on MMO

activities!

References

- Deb K, Gulati S. "Design of truss-structures for minimum weight using genetic algorithms," *Finite Elements Anal Des* 2001;37:447–65.
- G.-C. Luh and C.-Y. Lin, "Optimal design of truss-structures using particle swarm optimization," *Computers and Structures*, vol. 89, no.23-24, pp. 2221 – 2232, Dec. 2011.
- S. W. Mahfoud. *Niching Methods for Genetic Algorithms*. PhD thesis, Department of General Engineering, University of Illinois at Urbana-Champaign, Urbana, IL, 1995.
- K. A. de Jong. *An Analysis of the Behavior of a Class of Genetic Adaptive Systems*. PhD thesis, Department of Computer and Communication Sciences, University of Michigan, Ann Arbor, MI, 1975.
- G. Harick. Finding multi-modal solutions using restricted tournament selection. In *Proceedings of the Sixth International Conference on Genetic Algorithms*(ICGA-95), pages 24–31, 1997.
- A. Petrowski. A clearing procedure as a niching method for genetic algorithms. In *Proceedings of Third IEEE International Conference on Evolutionary Computation*(ICEC'96), pages 798–803. Piscataway, NJ:IEEE Press, 1996.
- J.-P. Li, M. E.B., G. T.P., and C. P.J. A species conserving genetic algorithm for multi-modal function optimization. In *Evolutionary Computation*, volume 10(3), pages 207–234, 2002.

References

- M. Epitropakis, D. Tasoulis, N. Pavlidis, V. Plagianakos, and M. Vrahatis, “Enhancing differential evolution utilizing proximity-based mutation operators,” *IEEE Trans on Evol. Comput.*, vol. 15, no. 1, pp. 99–119, Feb 2011.
- X. Li, “Niching without niching parameters: Particle swarm optimization using a ring topology,” *IEEE Trans. on Evol. Comput.*, vol. 14, no. 1, pp. 150 – 169, February 2010.
- J. Horn, N. Nafpliotis, and D. E. Goldberg, “A Niched Pareto Genetic Algorithm for Multiobjective Optimization,” in *Proc. of the First IEEE Conference on Evolutionary Computation, IEEE World Congress on Computational Intelligence*, vol. 1. Piscataway, New Jersey: IEEE Service Center, 1994, pp. 82–87.
- K. Deb, A. Pratap, S. Agarwal, and T. Meyarivan, “A fast and elitist multiobjective genetic algorithm: Nsga-ii,” Evolutionary Computation, IEEE Transactions on, vol. 6, no. 2, pp. 182–197, Apr 2002.
- X. Li, J. Branke, and T. Blackwell, “Particle swarm with speciation and adaptation in a dynamic environment,” in Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, ser. GECCO ’06. New York, NY, USA: ACM, 2006, pp. 51–58.
- D. Parrott and X. Li, “Locating and tracking multiple dynamic optima by a particle swarm model using speciation,” IEEE Trans. on Evol. Comput., vol. 10, no. 4, pp. 440–458, August 2006.
- G. Singh and K. Deb, “Comparisons of multi-modal optimization algorithms based on evolutionary algorithms,” in Proc. of the Genetic and Evolutionary Computation Conference 2006 (GECCO’06), Washington, USA, 2006, pp. 1305 – 1312.
- D. E. Goldberg and J. Richardson, “Genetic algorithms with sharing for multimodal function optimization,” in Proc. of the Second International Conference on Genetic Algorithms, J. Grefenstette, Ed., 1987, pp. 41–49.
- X. Li, A. Engelbrecht, and M. Epitropakis, “Benchmark functions for cec’2013 special session and competition on niching methods for multimodal function optimization,” Technical Report, Evolutionary Computation and Machine Learning Group, RMIT University, 2013.

References

- K. Deb and S. Tiwari, “Omni-optimizer: A generic evolutionary algorithm for single and multi-objective optimization.” European Journal of Operational Research, vol. 185, no. 3, pp. 1062–1087, 2008.
- X. Li, “Developing niching algorithms in particle swarm optimization,” in *Handbook of Swarm Intelligence*, ser. Adaptation, Learning, and Optimization, B. Panigrahi, Y. Shi, and M.-H. Lim, Eds. Springer Berlin Heidelberg, 2011, vol. 8, pp. 67–88.
- S. Das, S. Maity, B.-Y. Qu, and P. Suganthan, “Real-parameter evolutionary multimodal optimization - a survey of the state-of-the-art,” Swarm and Evolutionary Computation, vol. 1, pp. 71–88, June 2011.
- O. Shir, “Niching in evolutionary algorithms,” Handbook of Natural Computing: Theory, Experiments, and Applications, pp. 1035–1069, 2012.