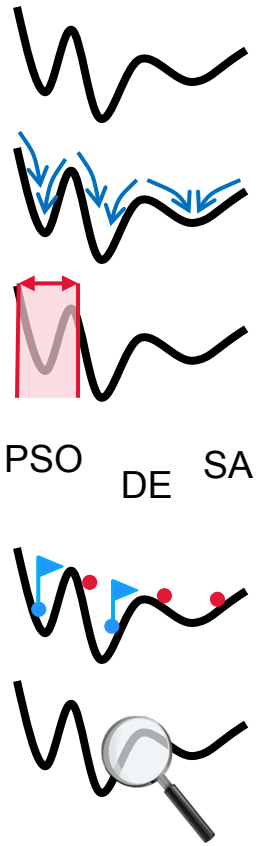
From Thresheld Convergence to Leaders and Followers

Stephen Chen and James Montgomery

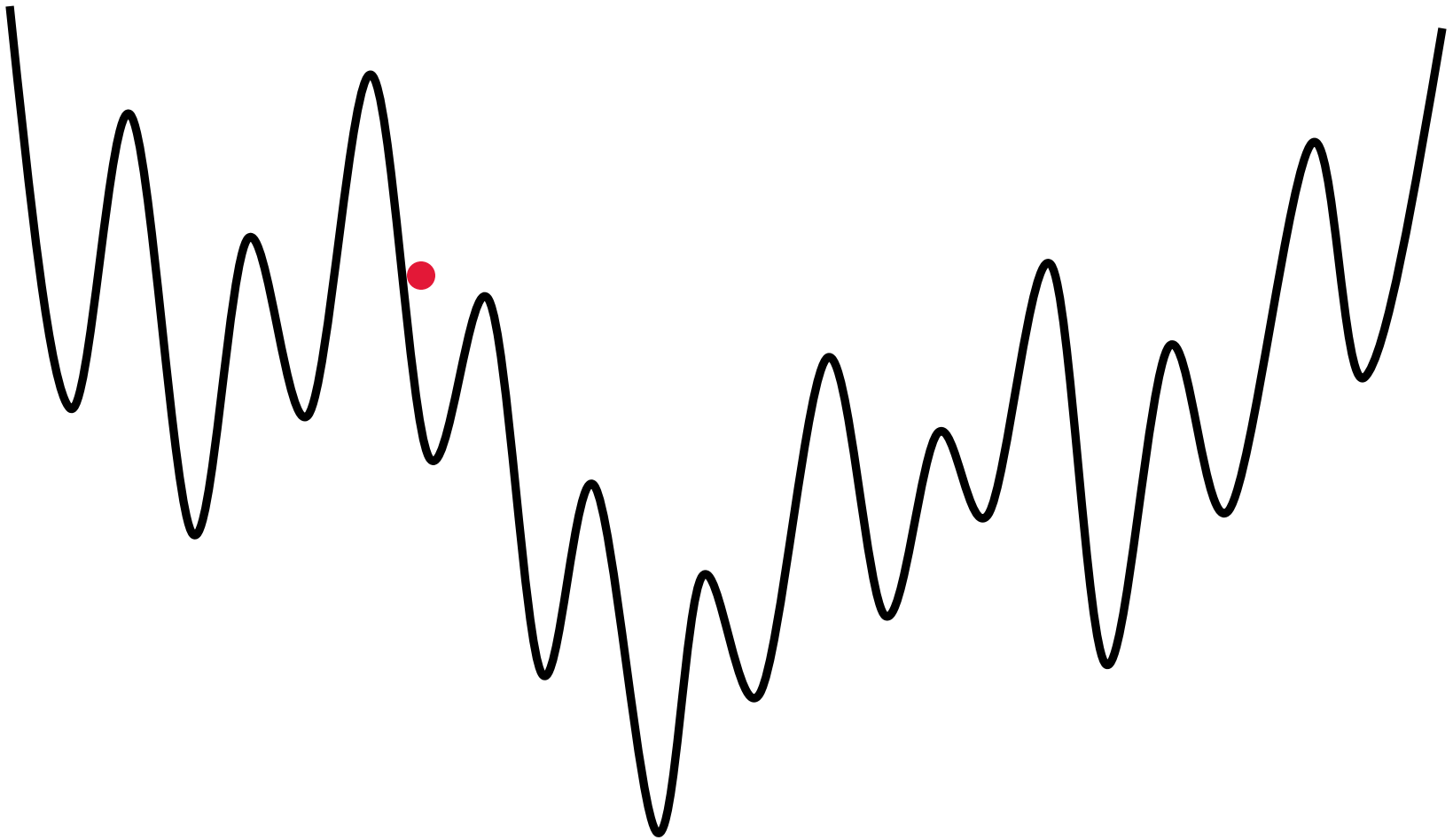
Overview

- Optimization in multi-modal search spaces
- A study of attraction basins
- An introduction to thresheld convergence
- Applications of thresheld convergence
- An introduction to leaders and followers
- Conclusions about search in multi-modal spaces
- Current and future work



Optimization in Multi-Modal Search Spaces

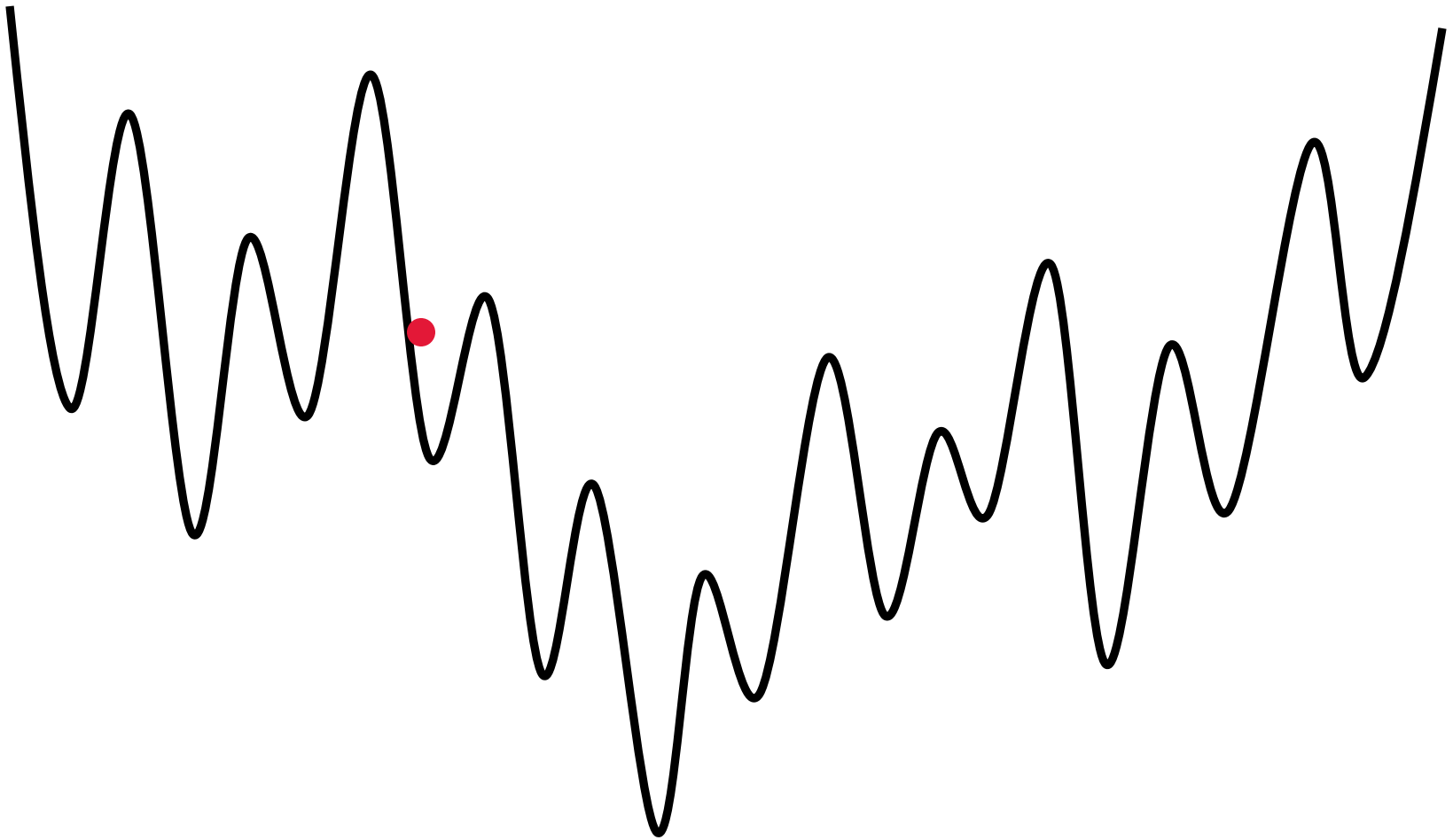
The Basics – Hill Climbing/Descent



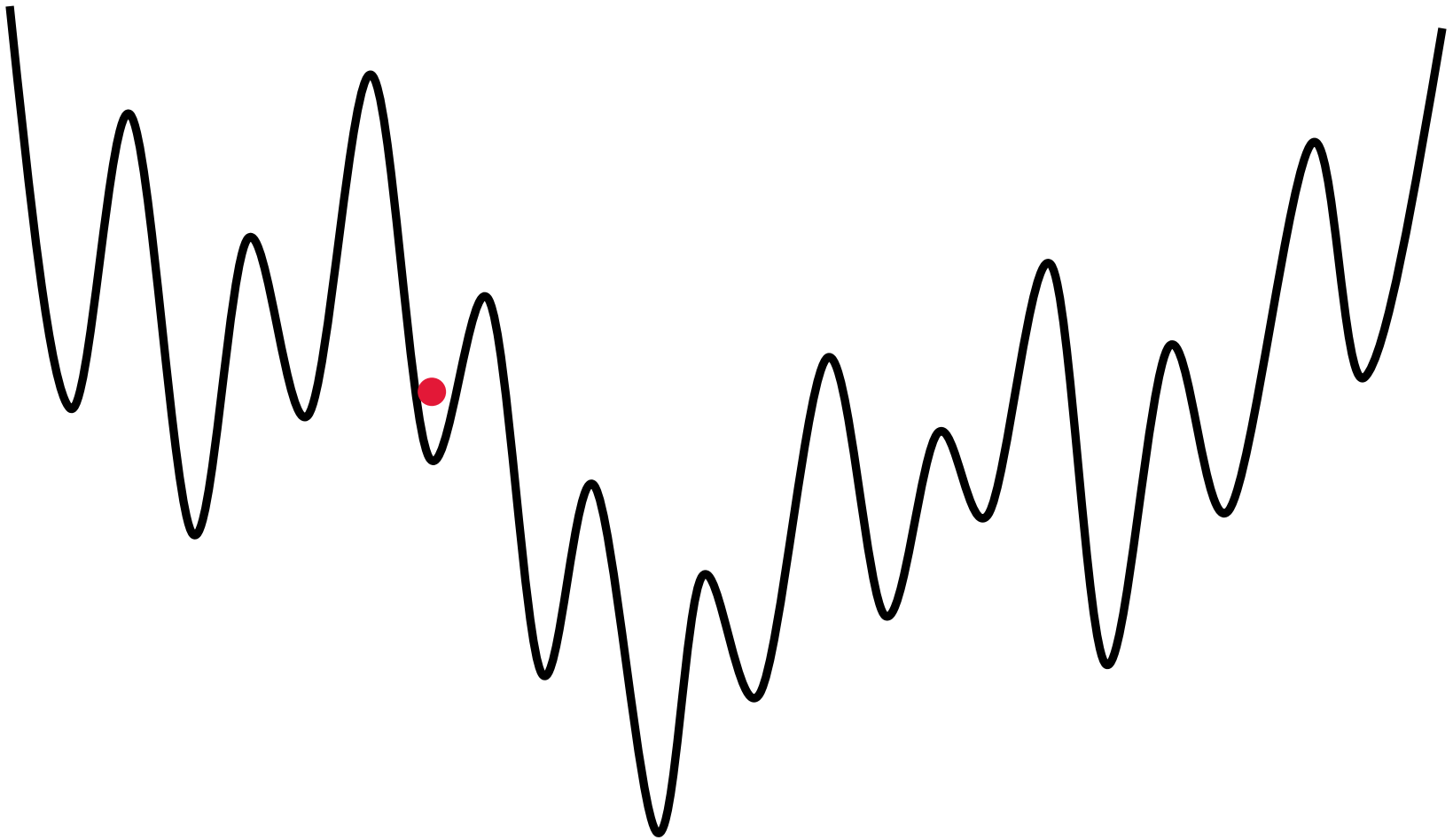
UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

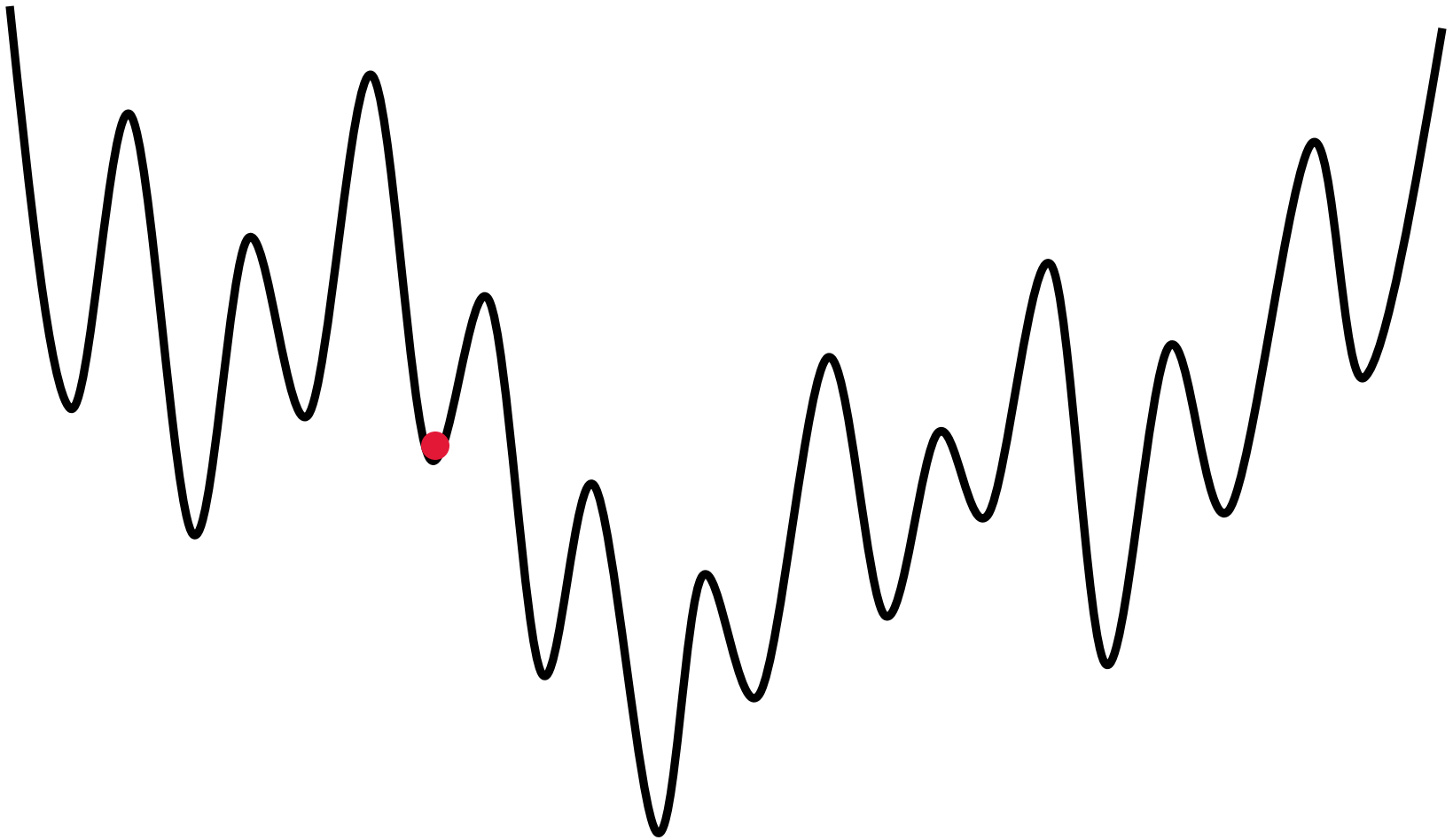
The Basics – Hill Climbing/Descent



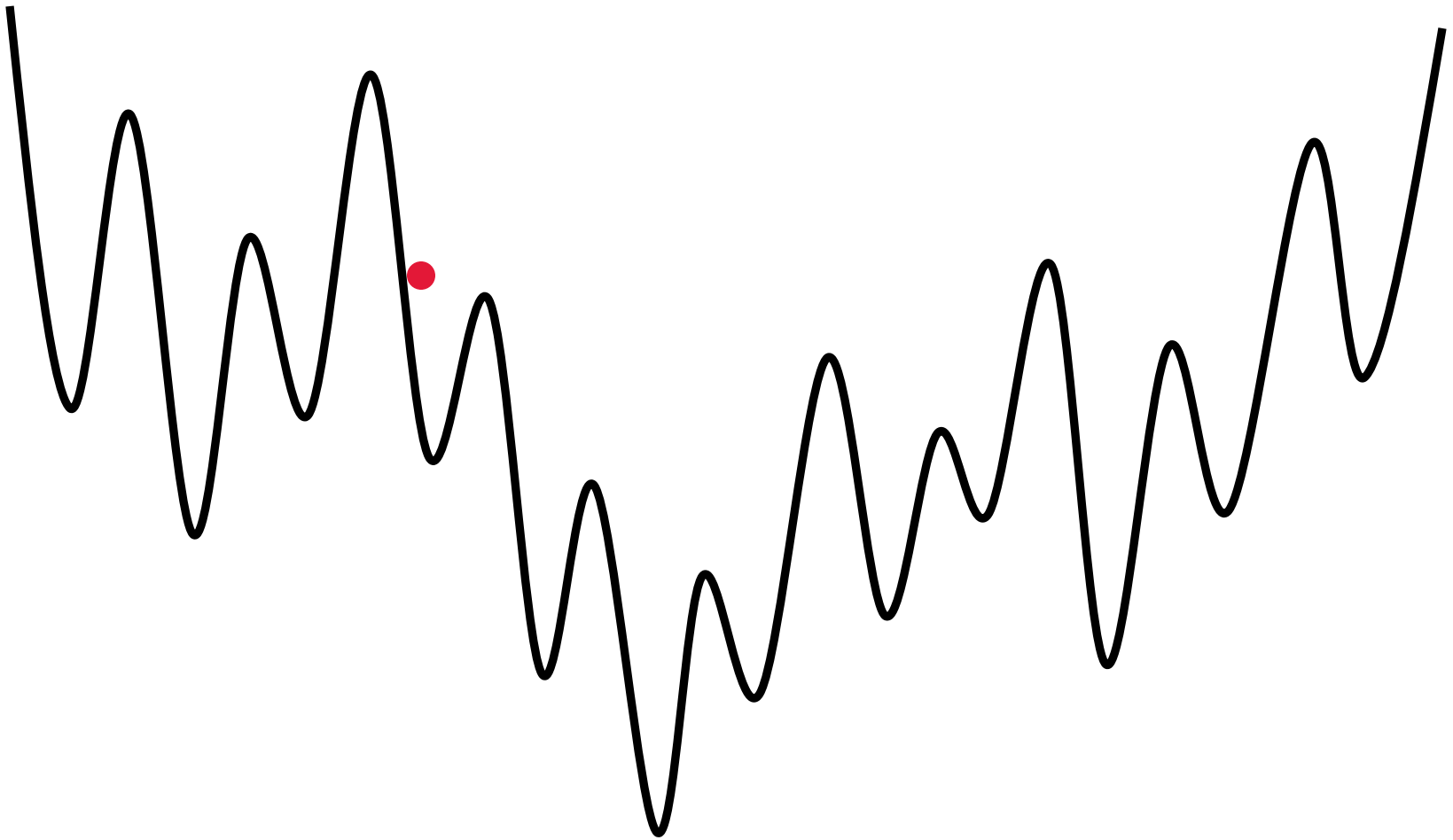
The Basics – Hill Climbing/Descent



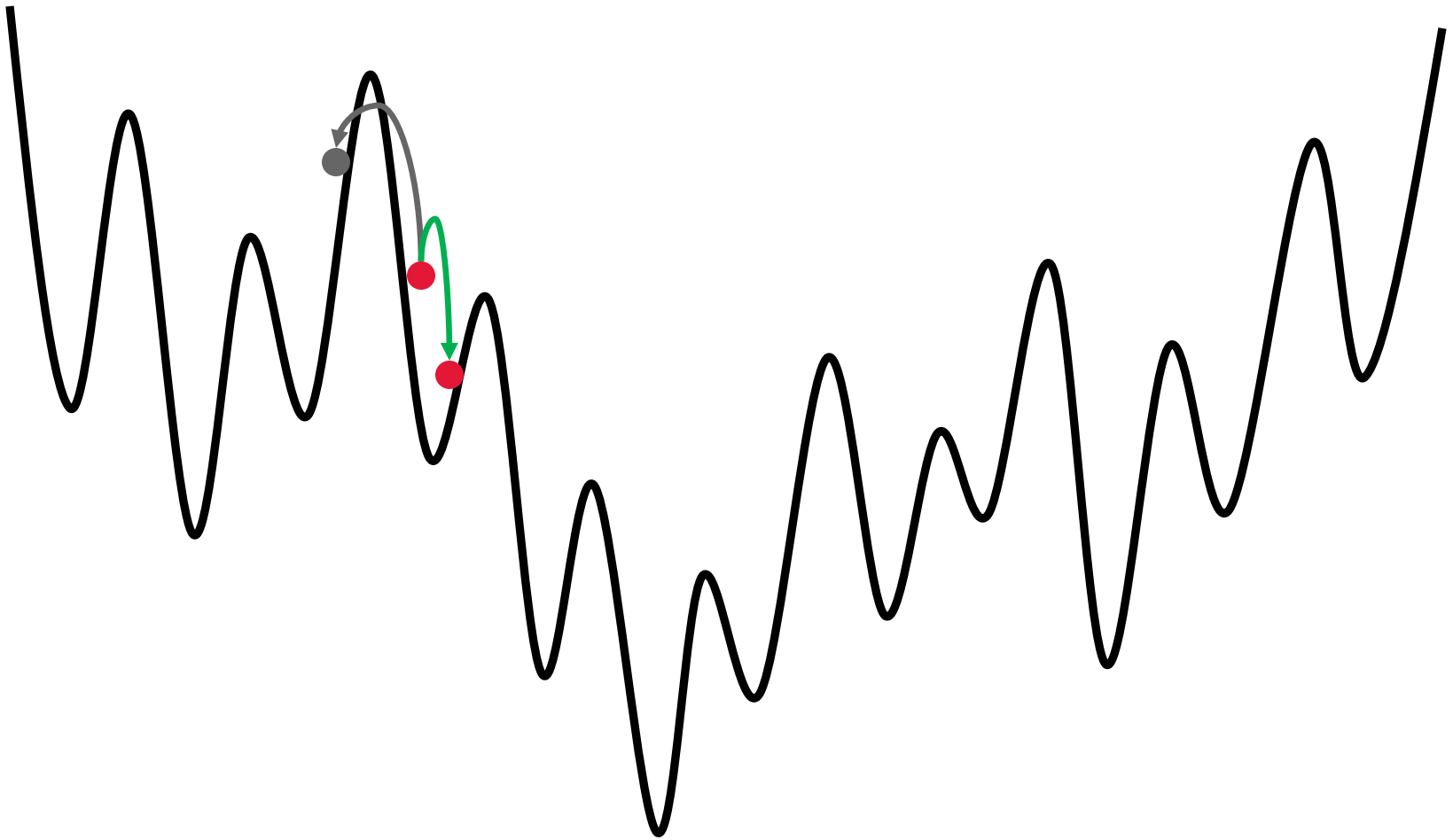
The Basics – Hill Climbing/Descent



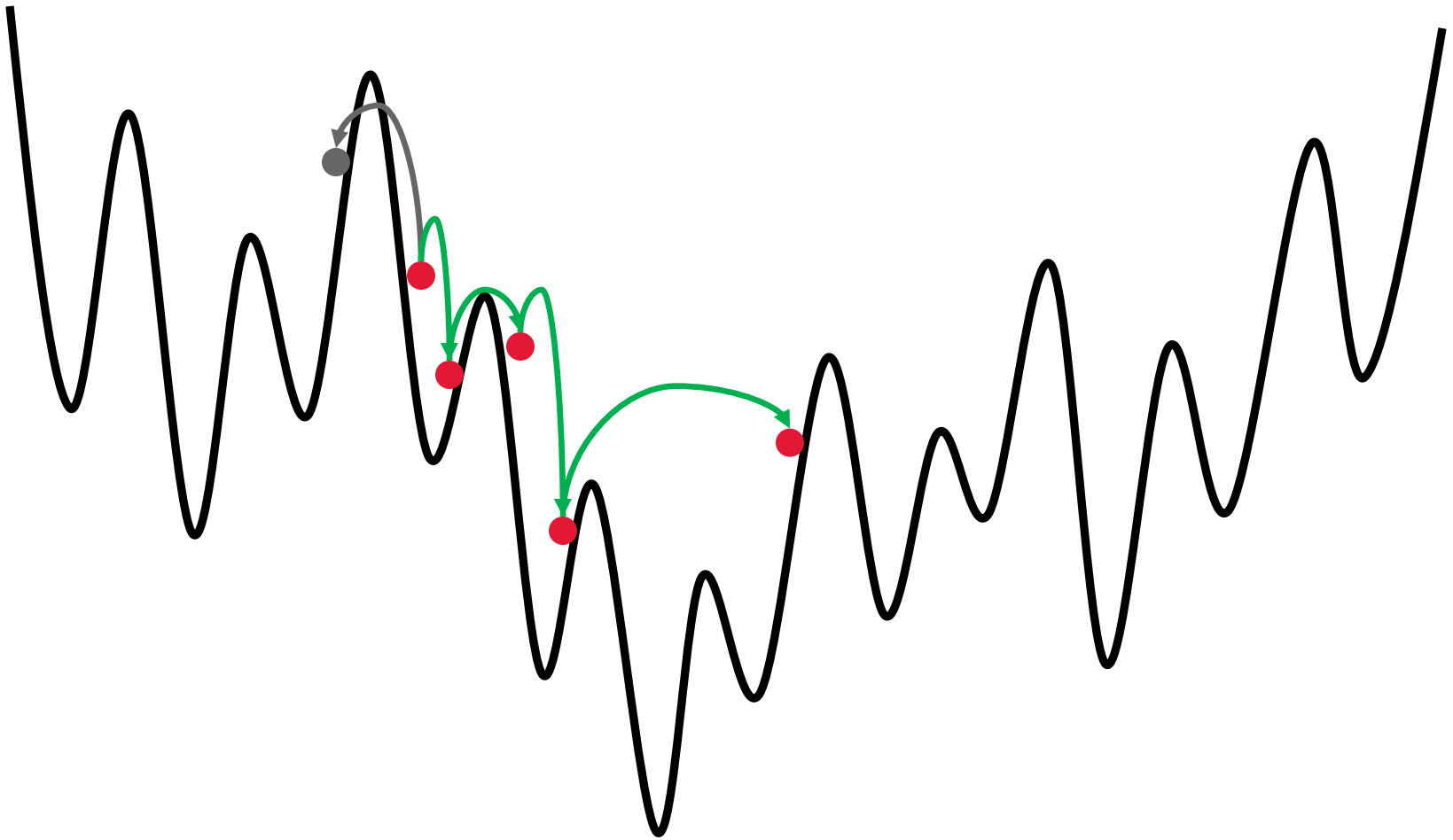
The Basics – Simulated Annealing



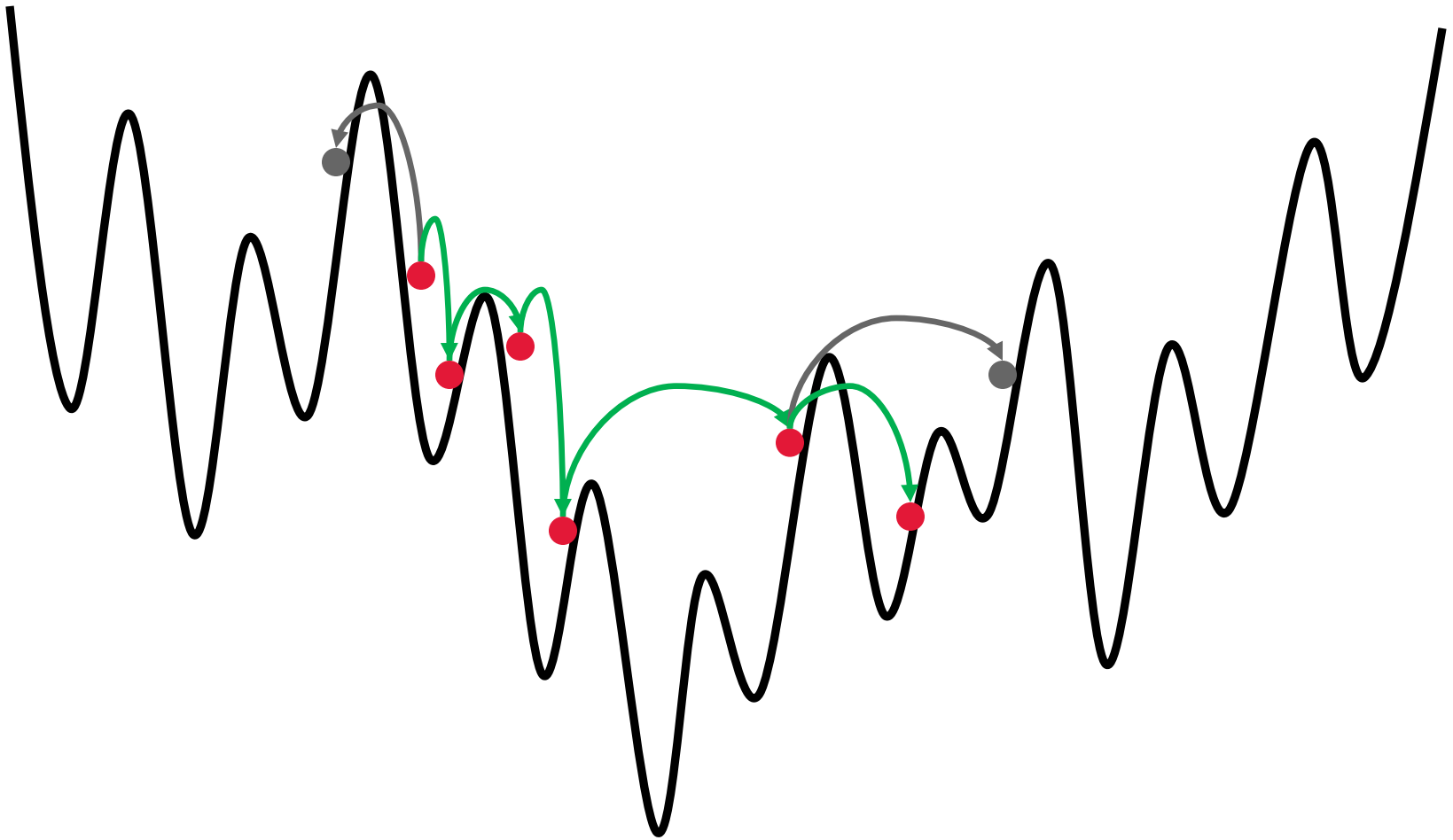
The Basics – Simulated Annealing



The Basics – Simulated Annealing



The Basics – Simulated Annealing



UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

Escape from Local Optima

- The definition of fitness landscapes, multi-modality, and local optima is often based on Hill Climbing/Descent
- This perspective leads to the mindset of needing to escape local optima
- Key example: Simulated Annealing

- Are other perspectives possible?

Rethinking Multi-Modal Optimization

- Stop thinking about local optima
  - Escaping from local optima
  - Finding multiple local optima
- Start thinking about attraction basins
  - Exploration → search among different attraction basins
  - Exploitation → search within the same attraction basin
- Distinct two-phase process
  - Find the fittest attraction basin
  - Find its local optimum

Two Phase Process for Metaheuristics

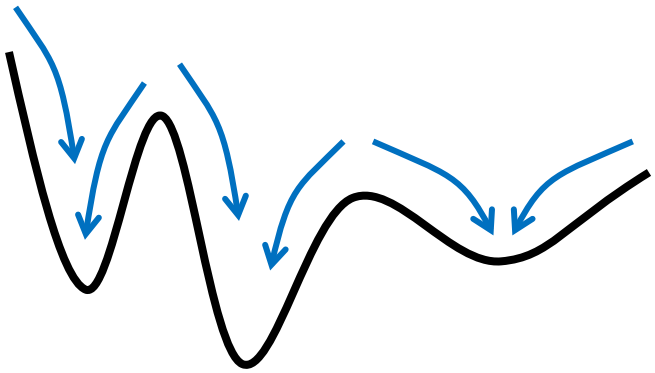
Phase 1 – Exploration

- Find the fittest attraction basin
  - (Not necessarily the fittest search point)

Phase 2 – Exploitation

- Stay within the fittest attraction basin (no further exploration)

Attraction Basins



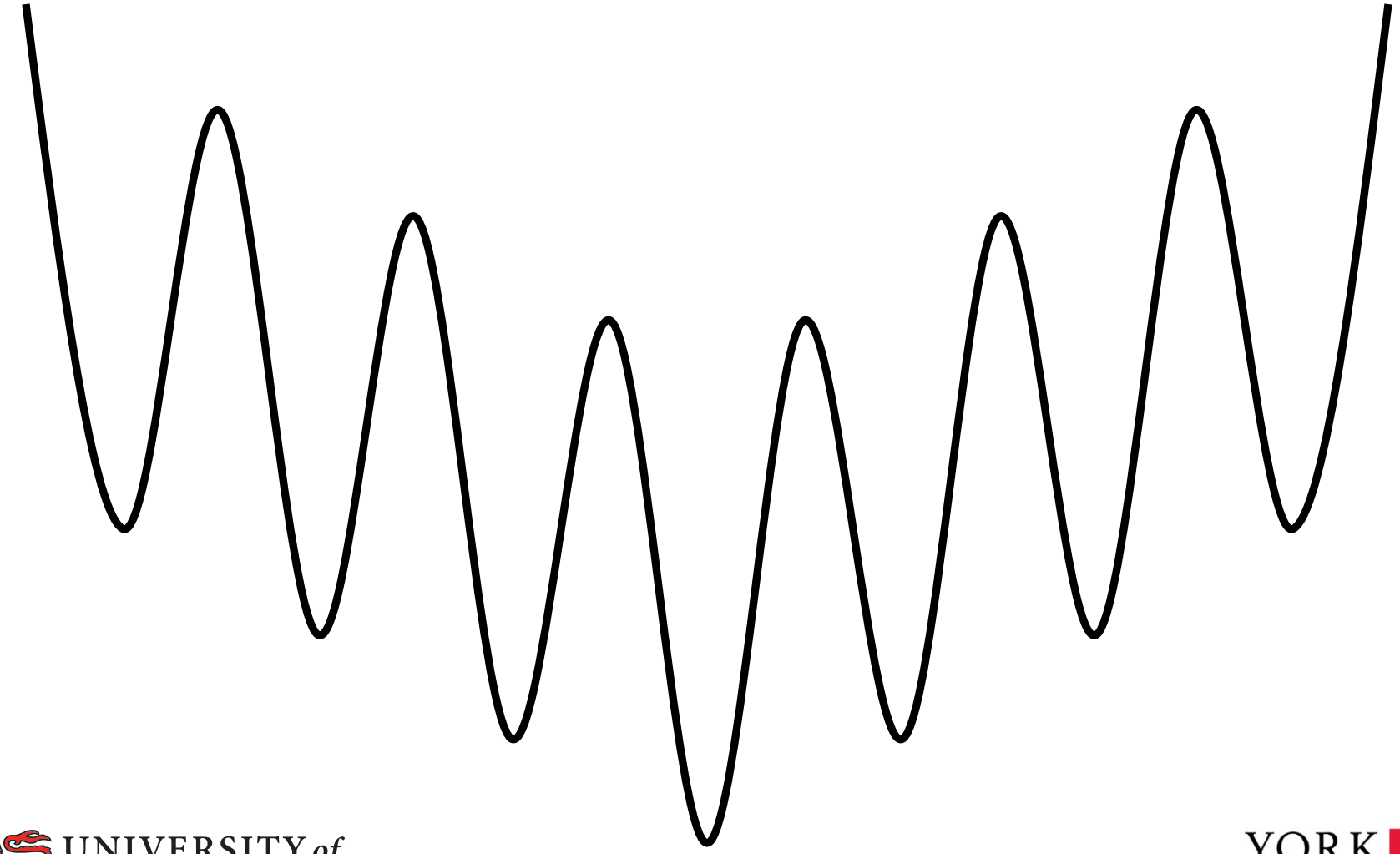
UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

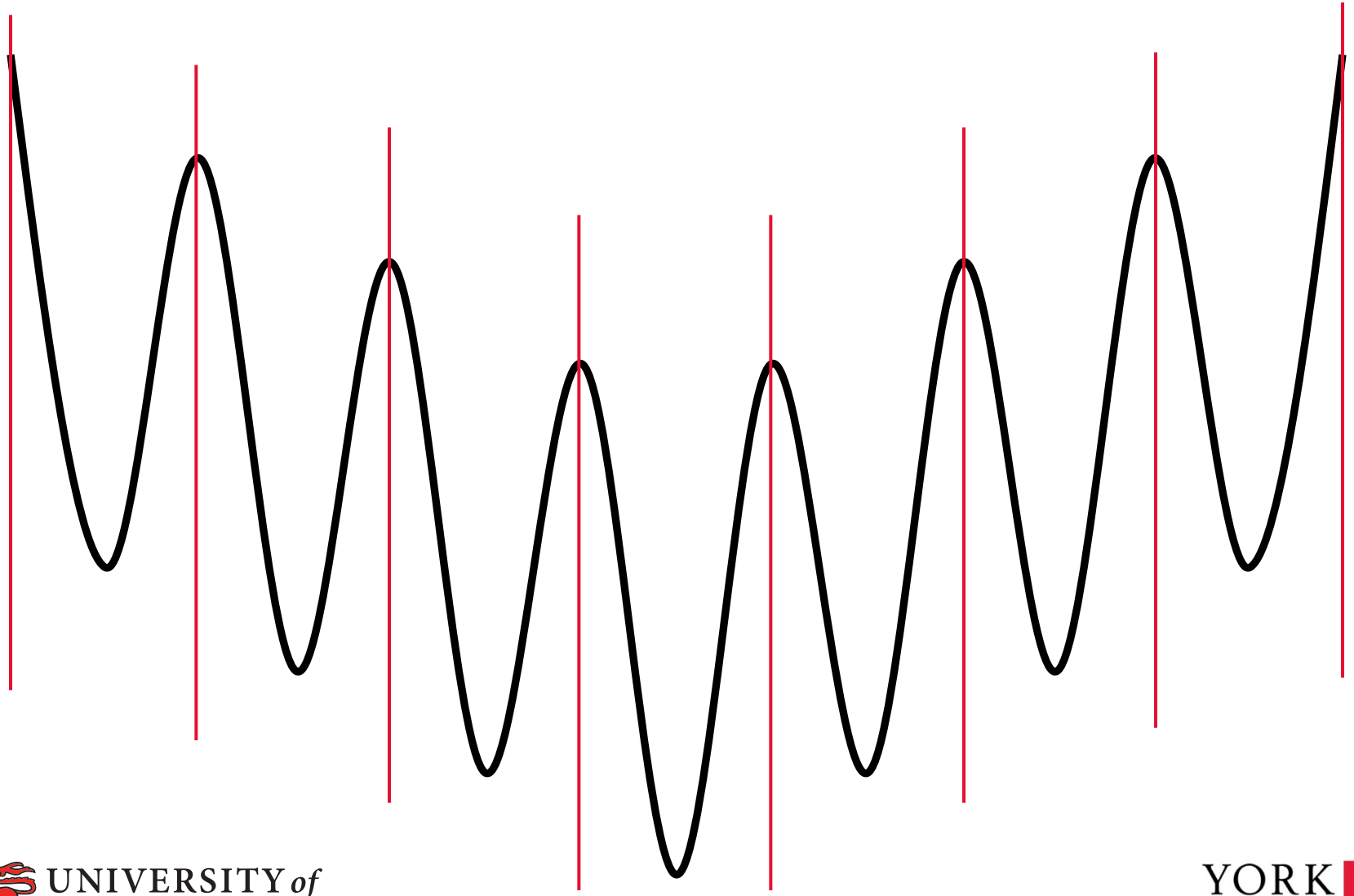
Definitions

- An attraction basin is all of the points around a local optimum that lead to that optimum when greedy local search is used
- The fitness of the an attraction basin is the fitness of its local optimum

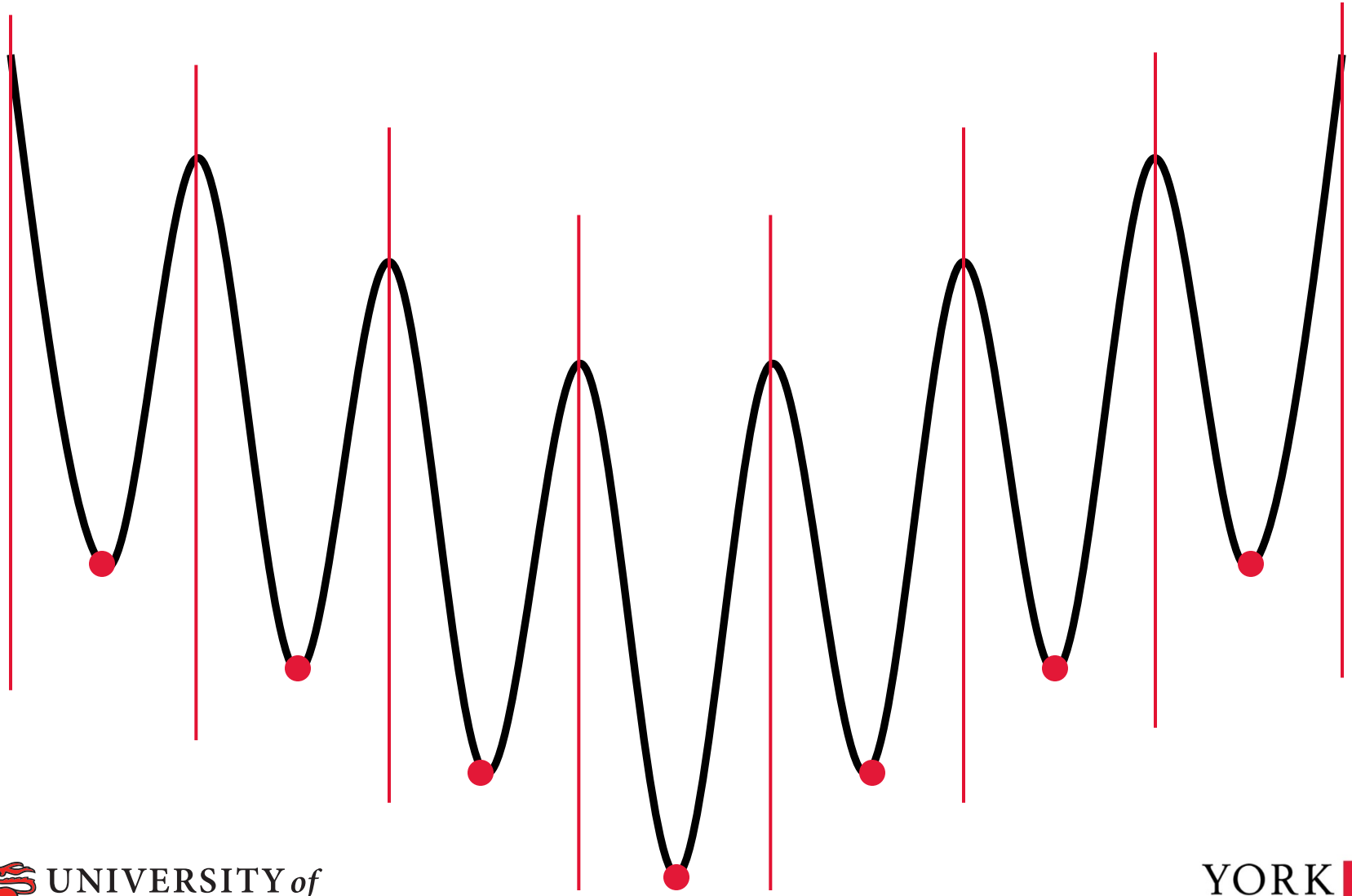
Attraction Basins



Attraction Basins



Attraction Basins



UNIVERSITY of TASMANIA

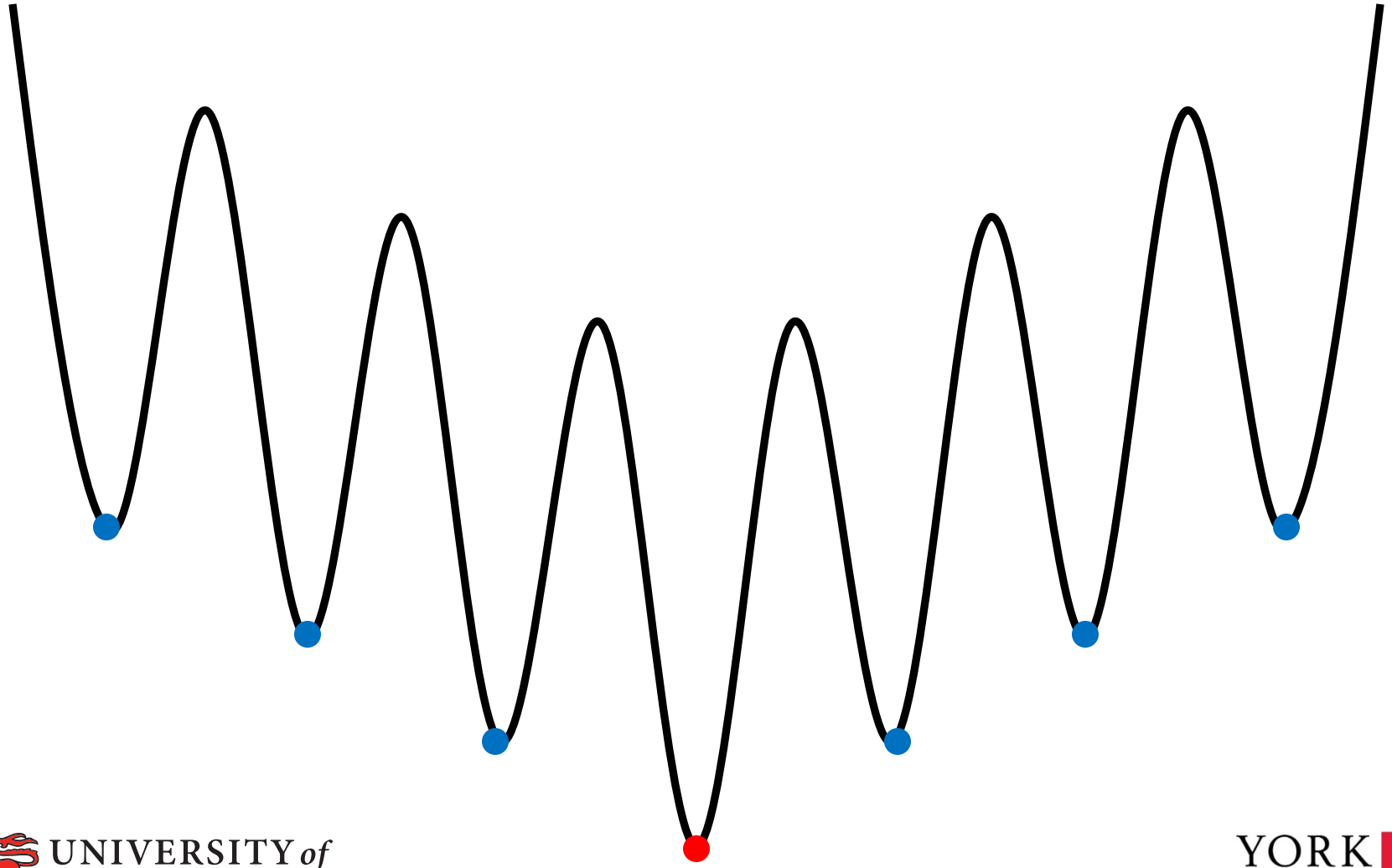
YORK UNIVERSITÉ UNIVERSITY

A Two-Phase Search Process

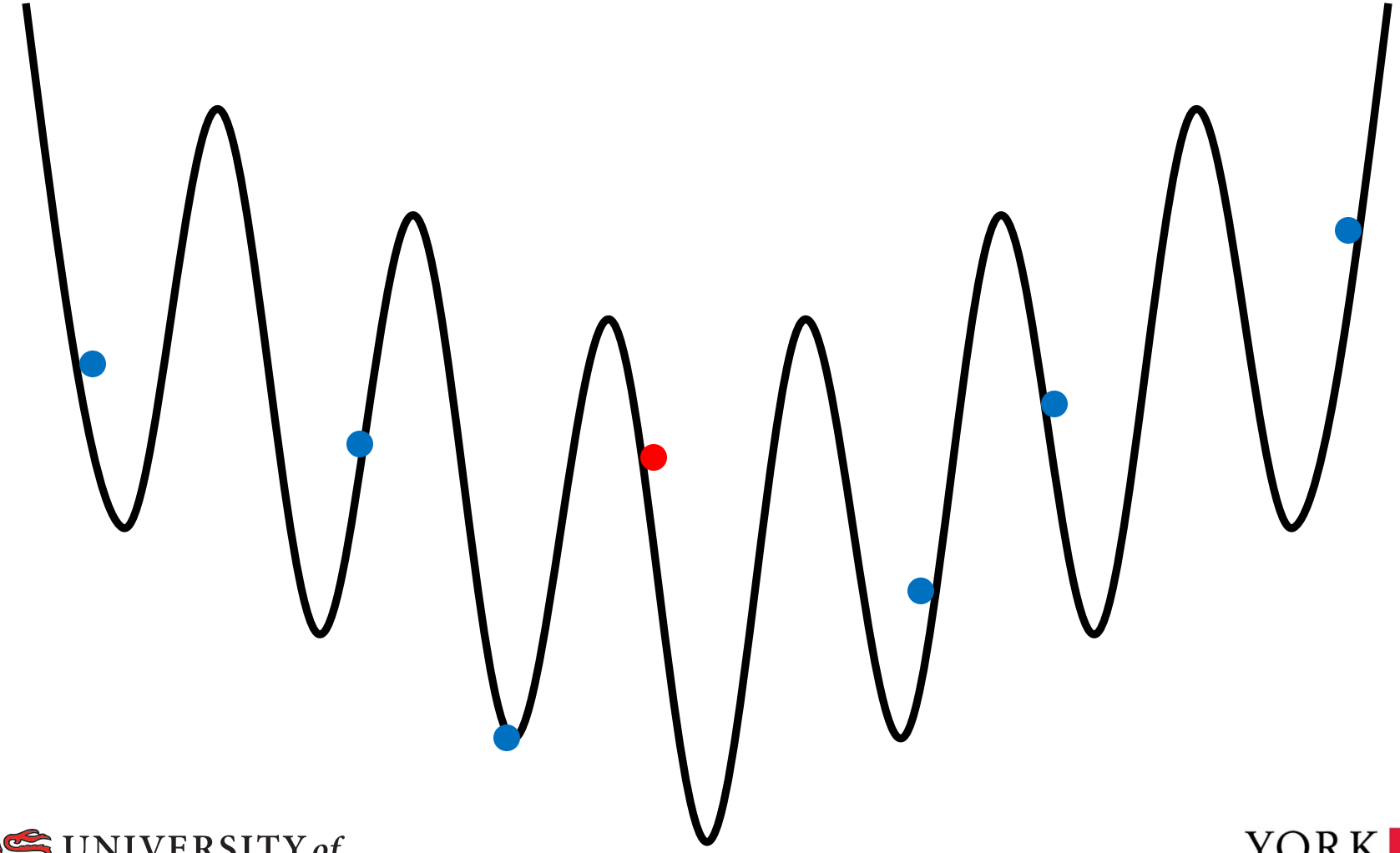
The global optimum is the local optimum in the fittest attraction basin

1. Find the fittest attraction basin (exploration)
2. Find the local optimum in this attraction basin (exploitation)

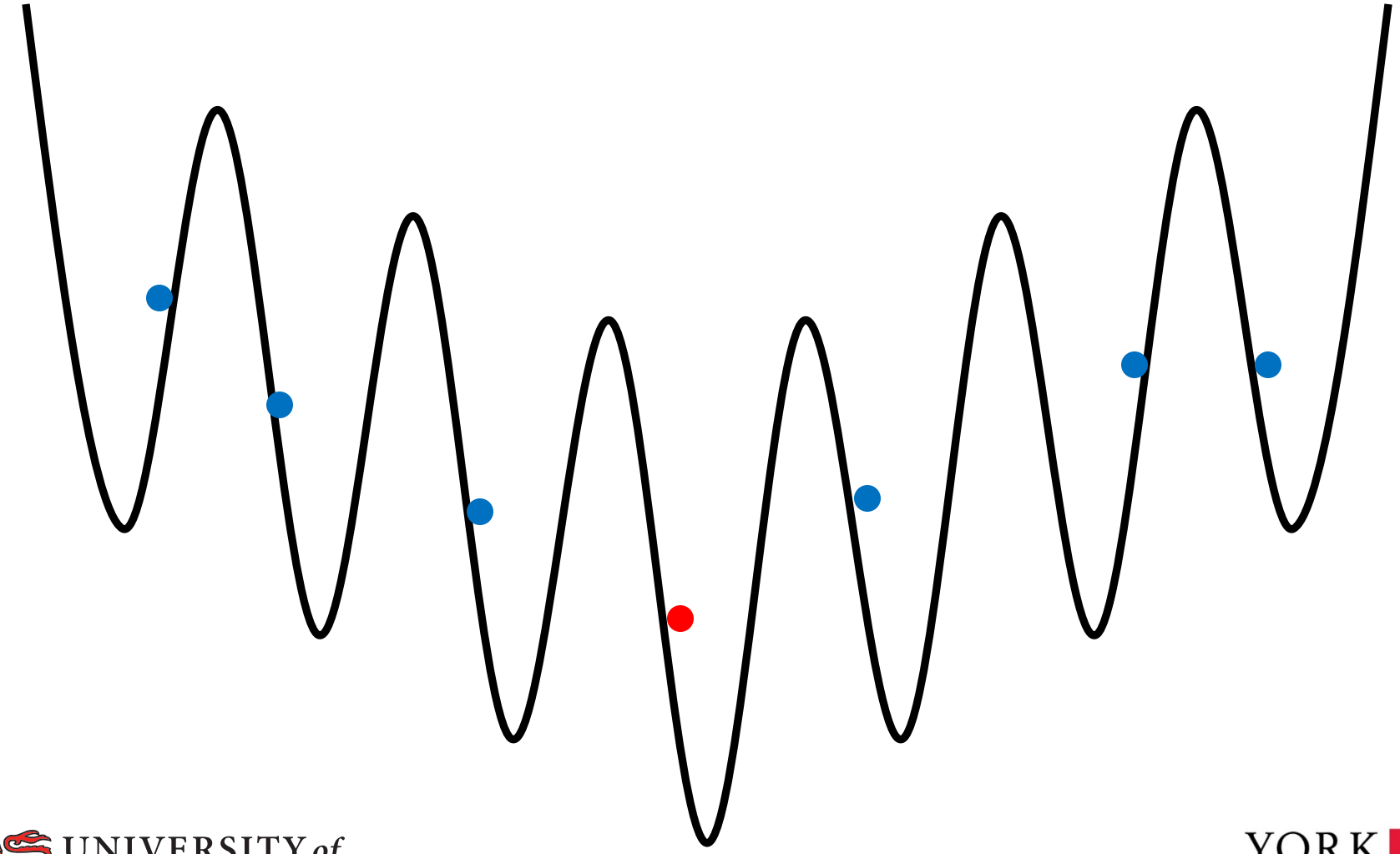
Finding the Fittest Attraction Basin (Trivial Case)



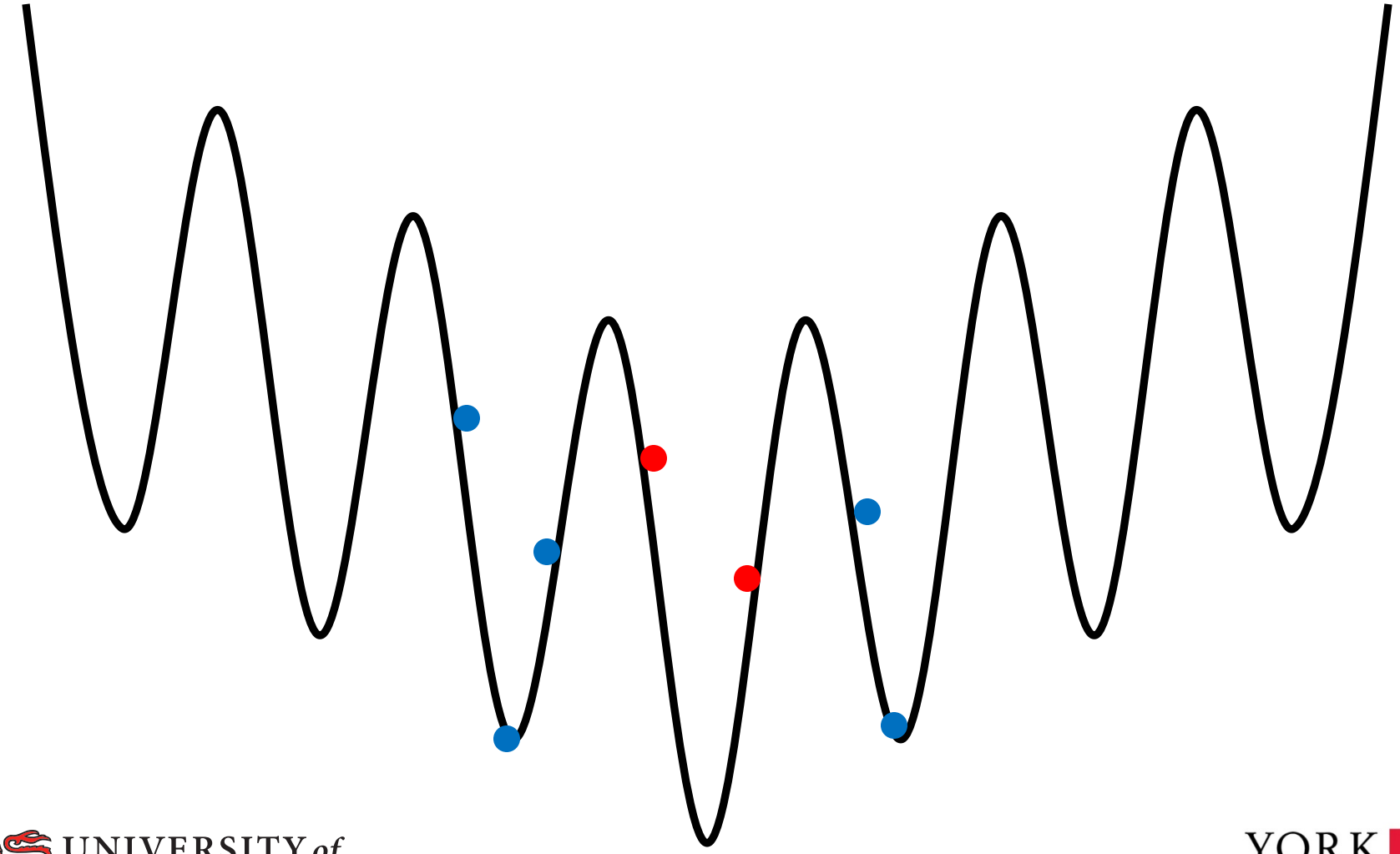
Finding the Fittest Attraction Basin (Hard Case)



Finding the Fittest Attraction Basin (Typical Case – Early)



Finding the Fittest Attraction Basin (Typical Case – Late)



A Detailed Study with Rastrigin

Attraction basins of known size, shape, and location

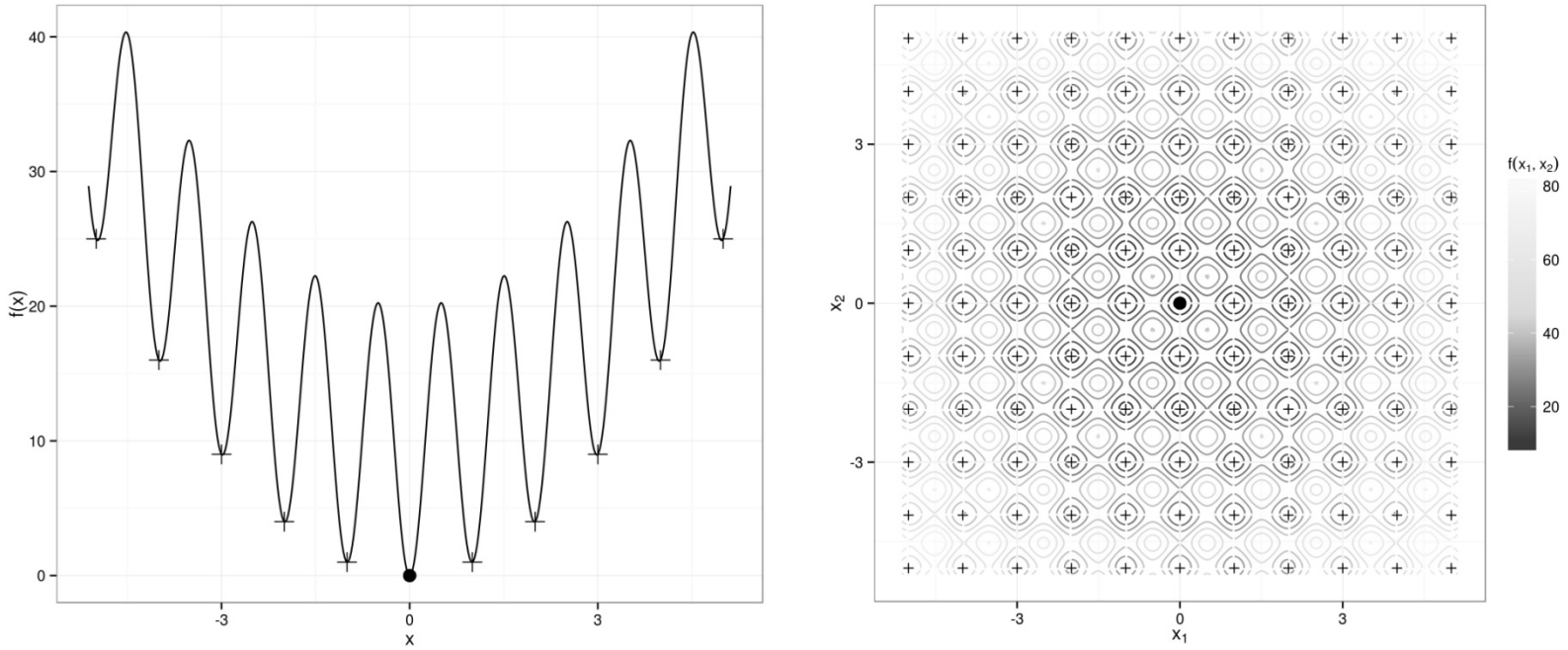


Image from Y. Gonzalez-Fernandez and S. Chen, “Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information”, IEEE CEC, 2015.

UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

A Detailed Study with Rastrigin

- Focus on exploration
  - The ability to identify the fitter attraction basin based on the fitness of random sample solutions
- Demonstrate the negative effects of concurrent exploration and exploitation
  - Comparing only random solutions is the Easy Case
  - Comparing random solutions against a local optimum is the Hard Case

A Detailed Study with Rastrigin

Success of exploration goes down with exploitation

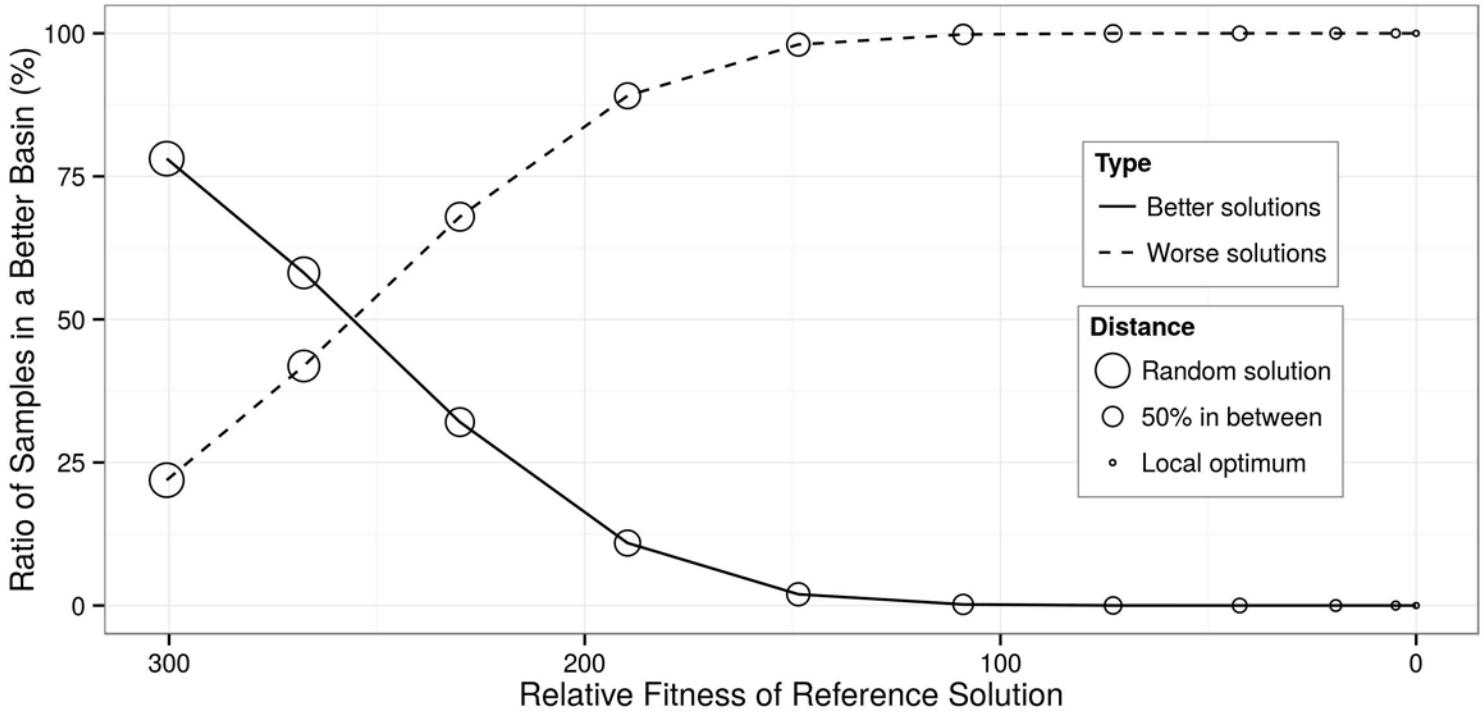
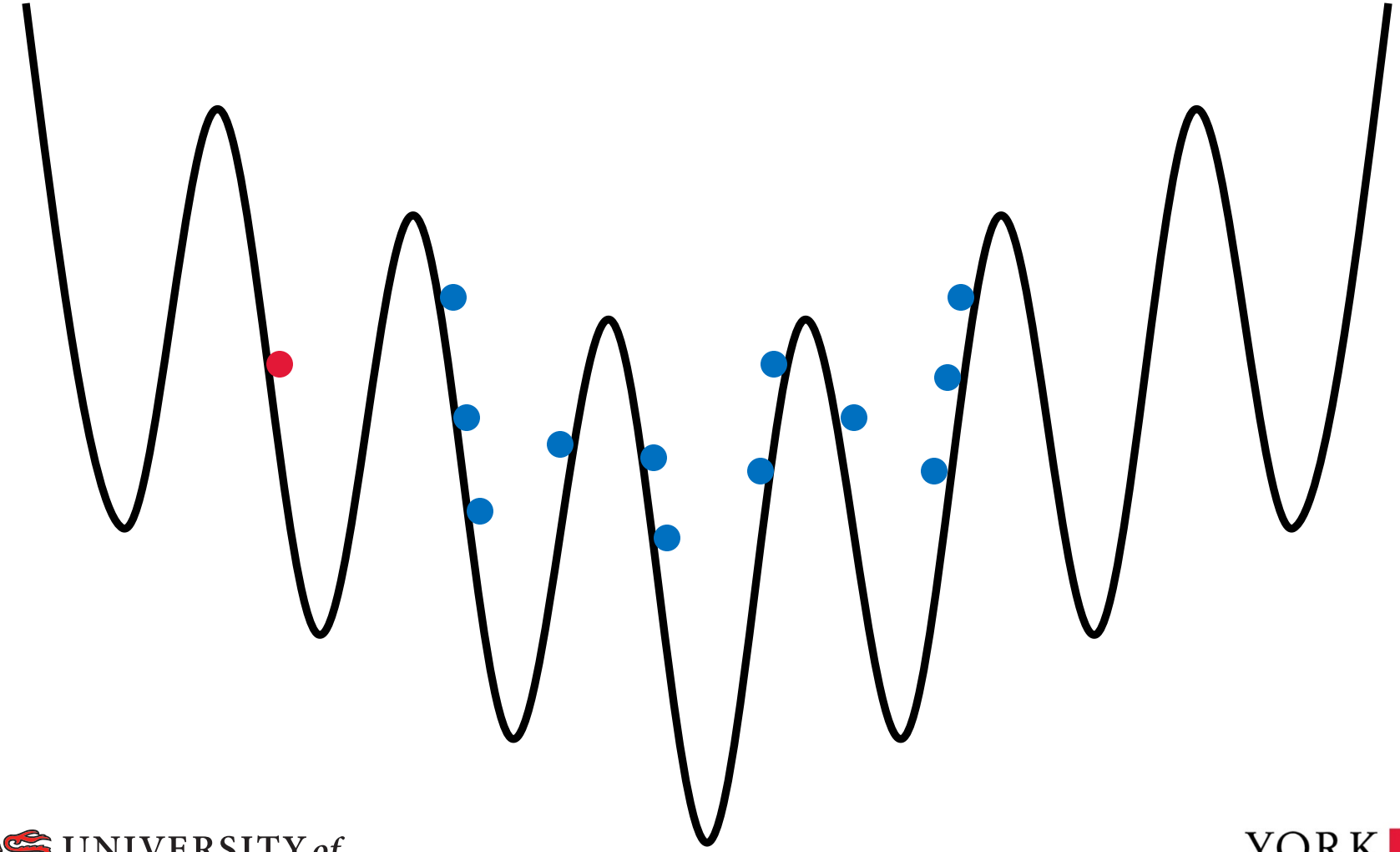
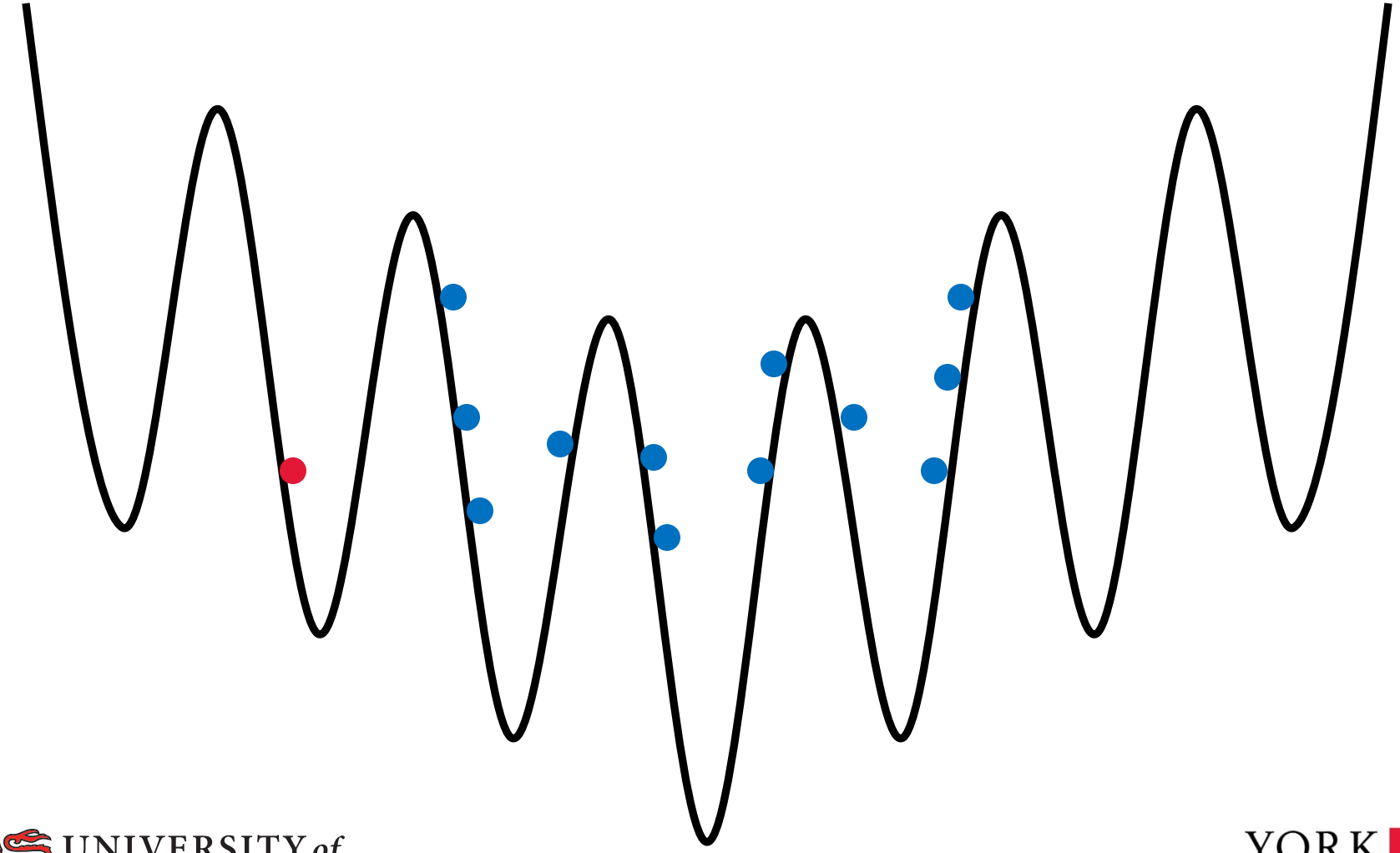


Image from Y. Gonzalez-Fernandez and S. Chen, "Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information", IEEE CEC, 2015.

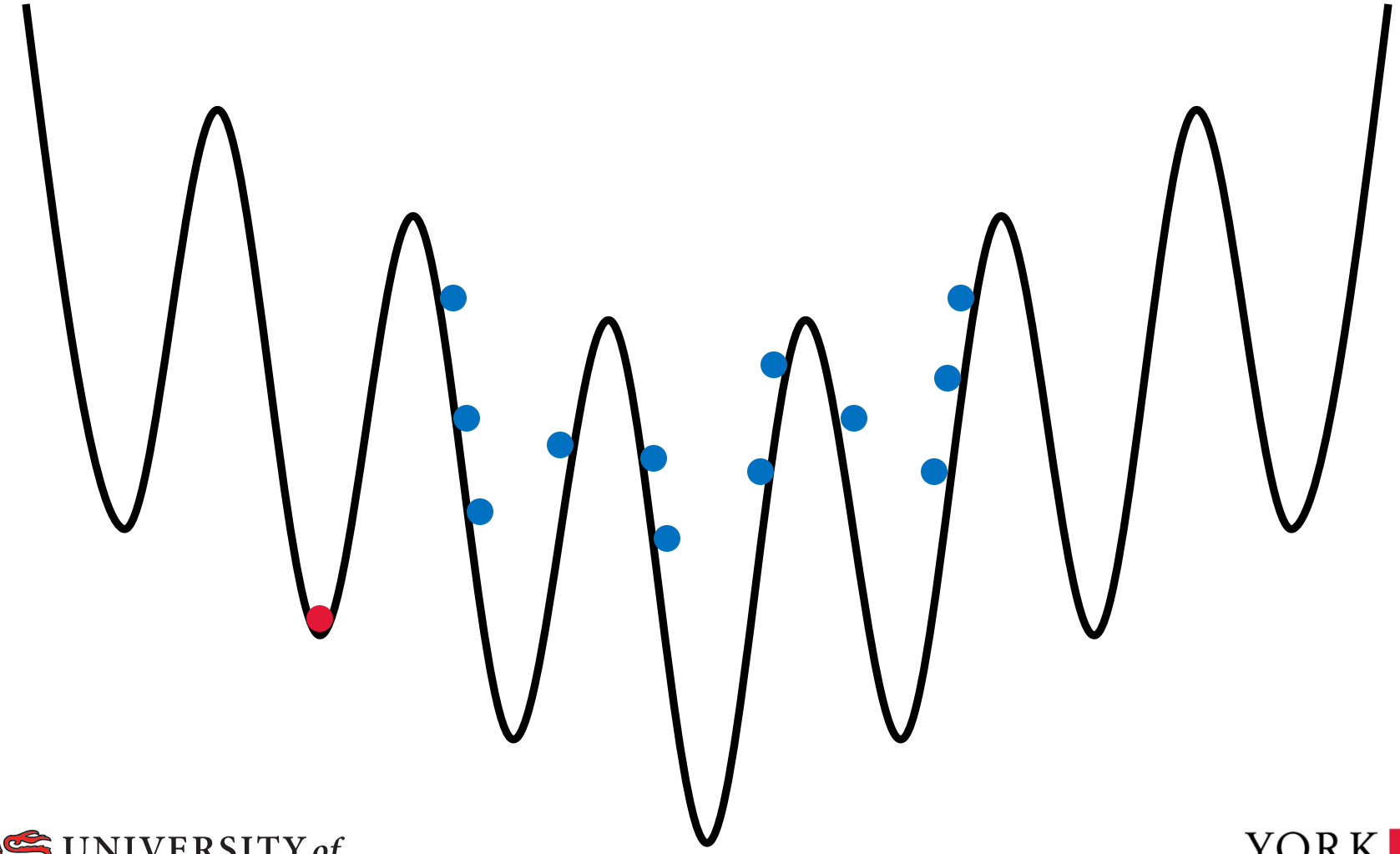
A Detailed Study with Rastrigin – Initial Random Solution



A Detailed Study with Rastrigin – Partially Optimized Solution



A Detailed Study with Rastrigin – Locally Optimal Solution



A Detailed Study with Rastrigin

50% optimization makes exploration almost impossible

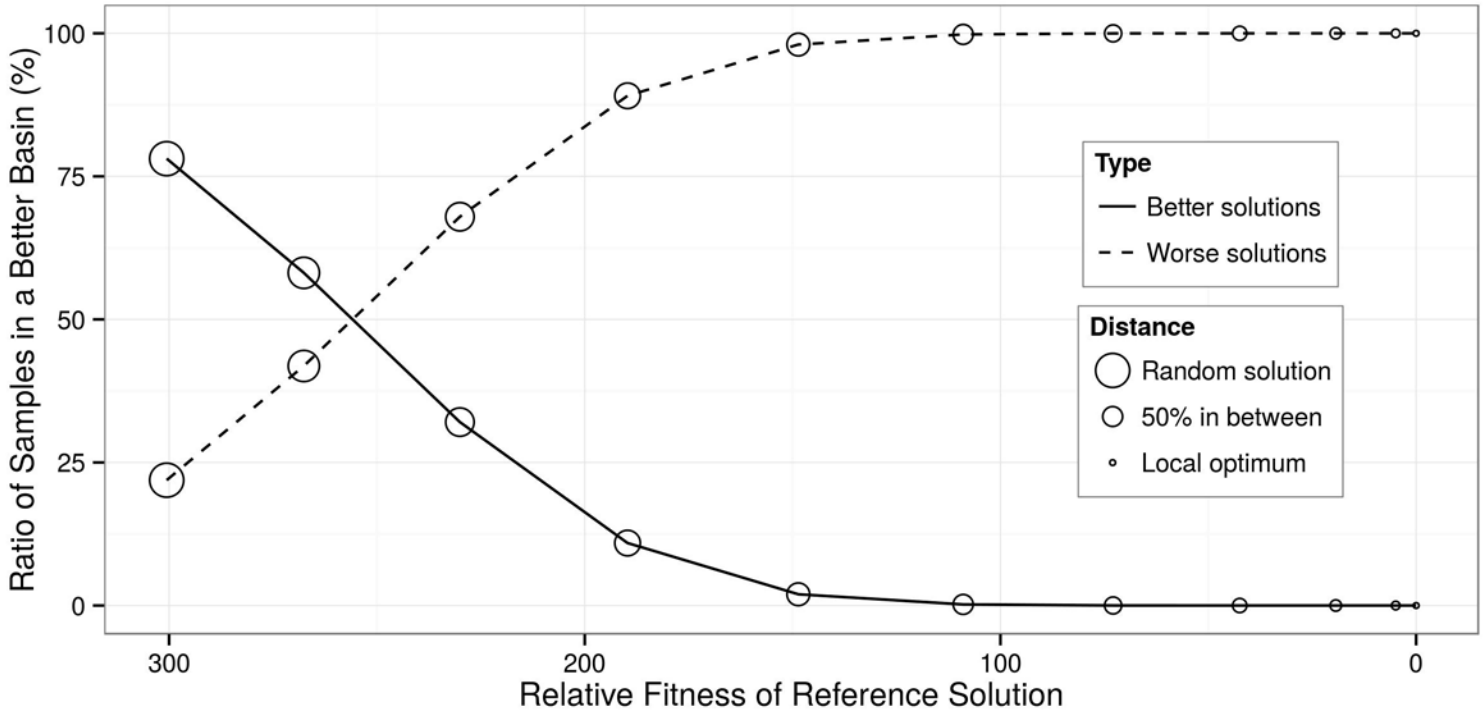
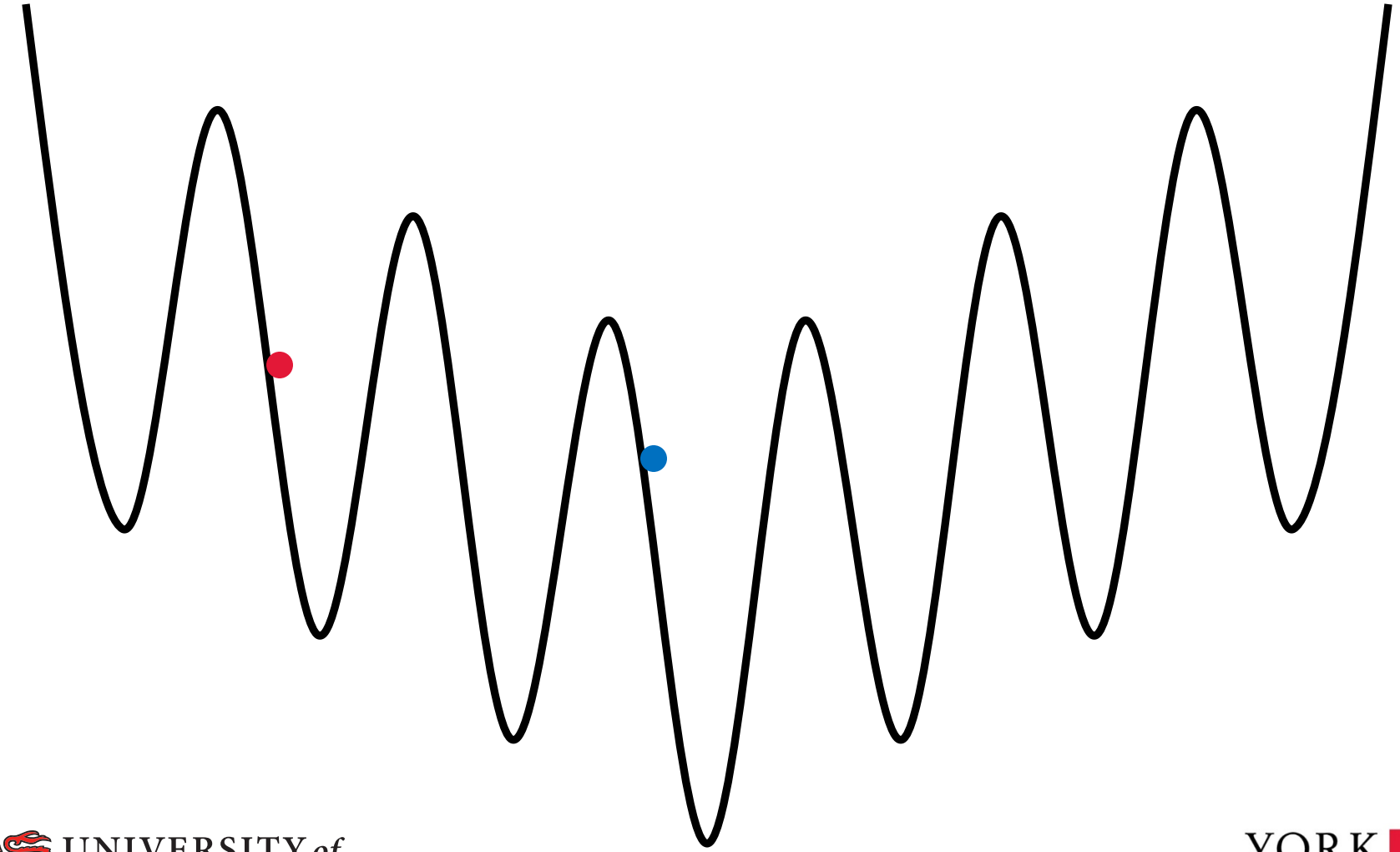
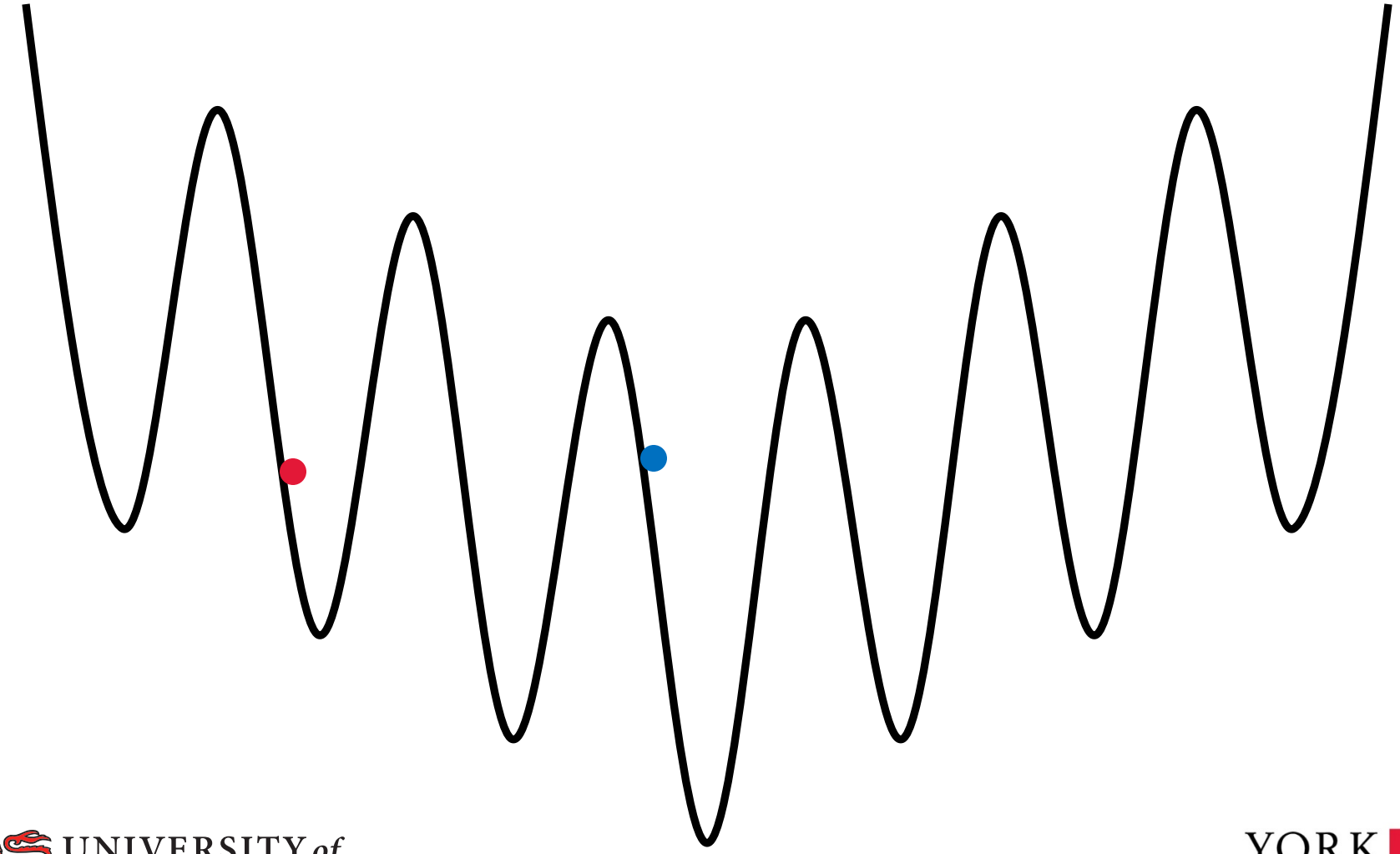


Image from Y. Gonzalez-Fernandez and S. Chen, "Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information", IEEE CEC, 2015.

Concurrent Exploration and Exploitation – From Successful Exploration



Concurrent Exploration and Exploitation – To Unsuccessful Exploration



A Detailed Study with Rastrigin

PSO and DE perform local optimization very early

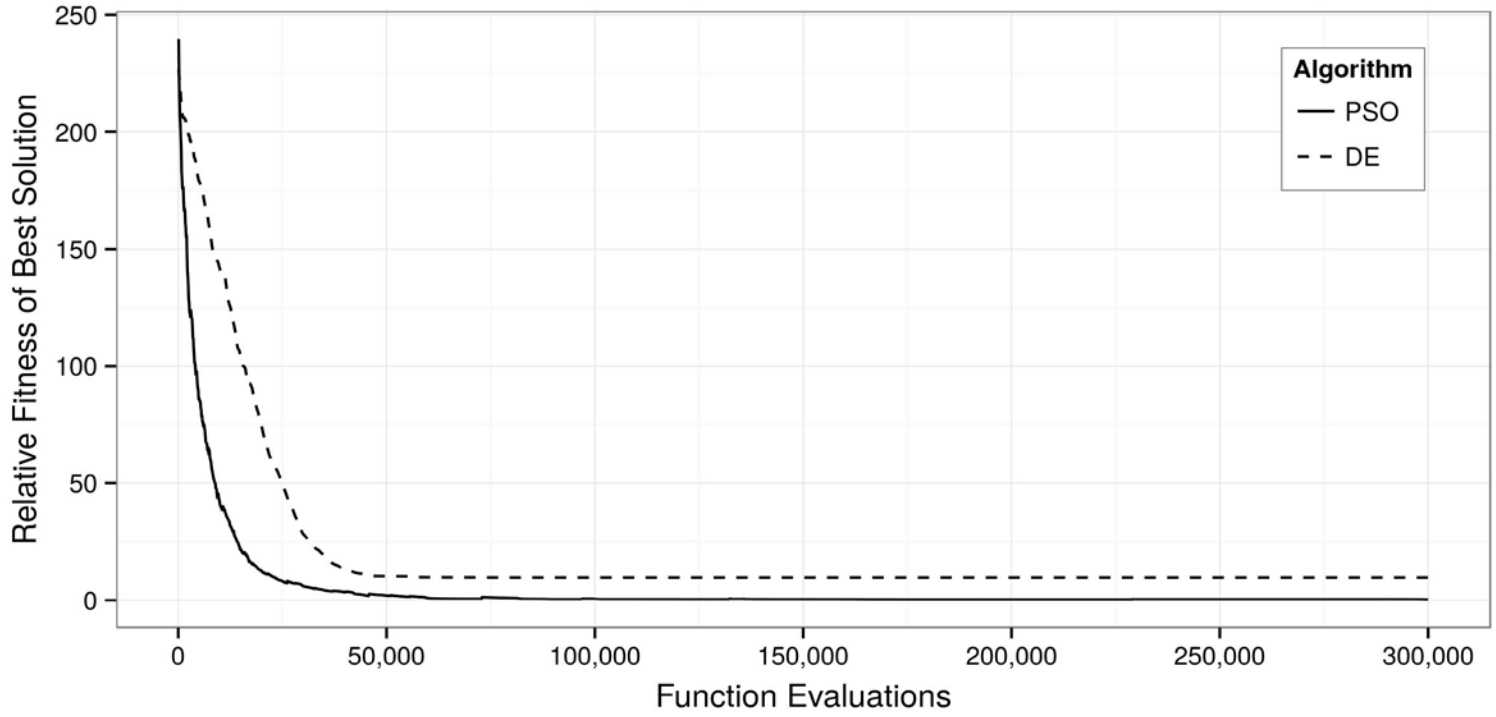


Image from Y. Gonzalez-Fernandez and S. Chen, "Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information", IEEE CEC, 2015.

A Detailed Study with Rastrigin

PSO and DE stall quickly – no more exploration

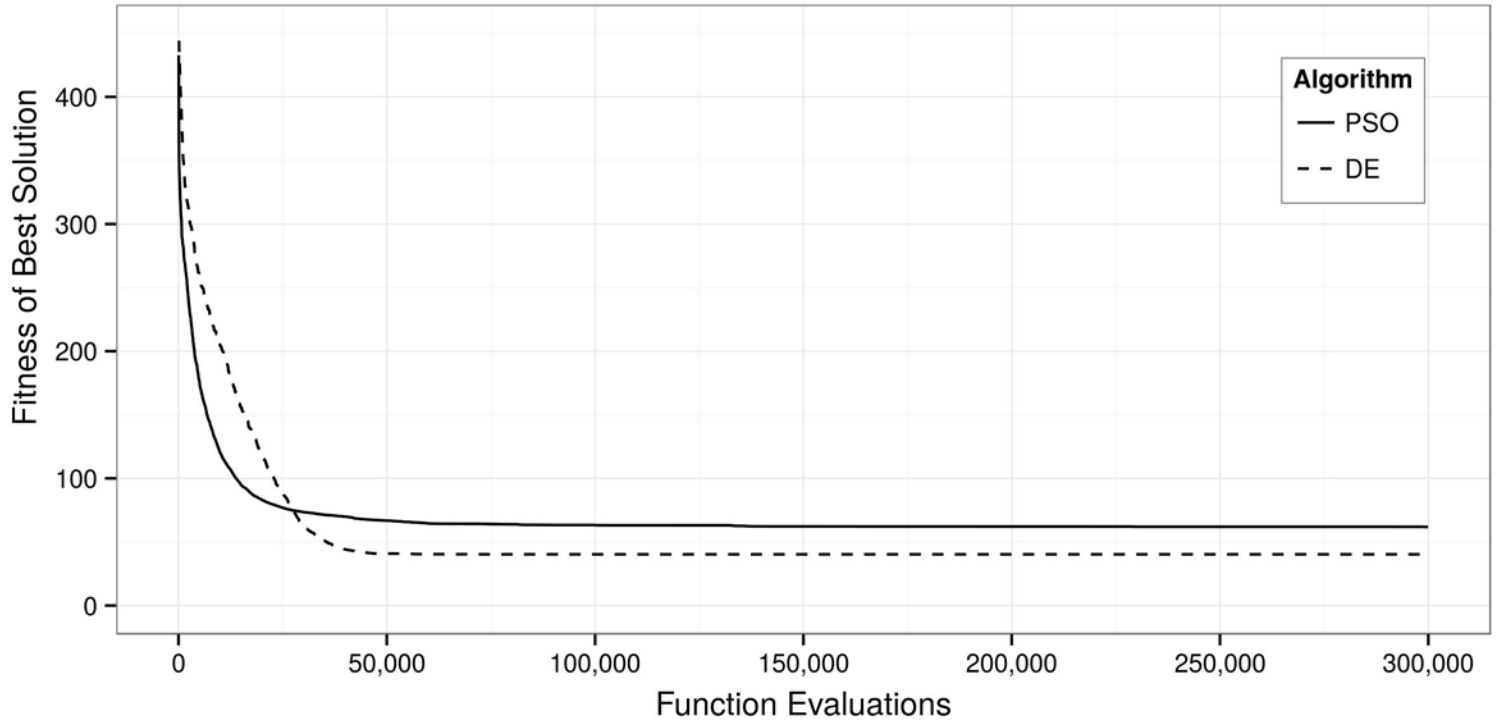
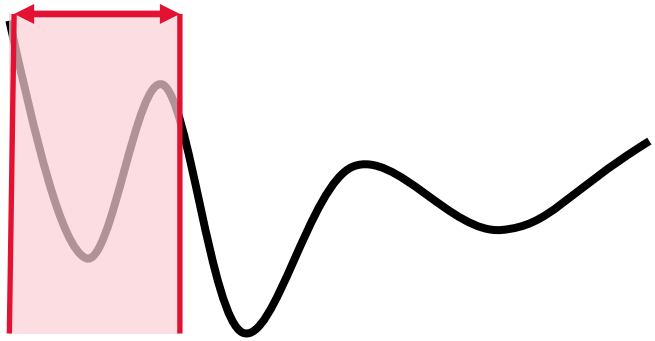


Image from Y. Gonzalez-Fernandez and S. Chen, "Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information", IEEE CEC, 2015.

Thresheld Convergence

an introduction



UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

Optimization in Multi-Modal Search Spaces

Phase 1: Find the fittest attraction basin

- Concurrent exploration and exploitation interferes with exploration
- Eliminate exploitation during this phase

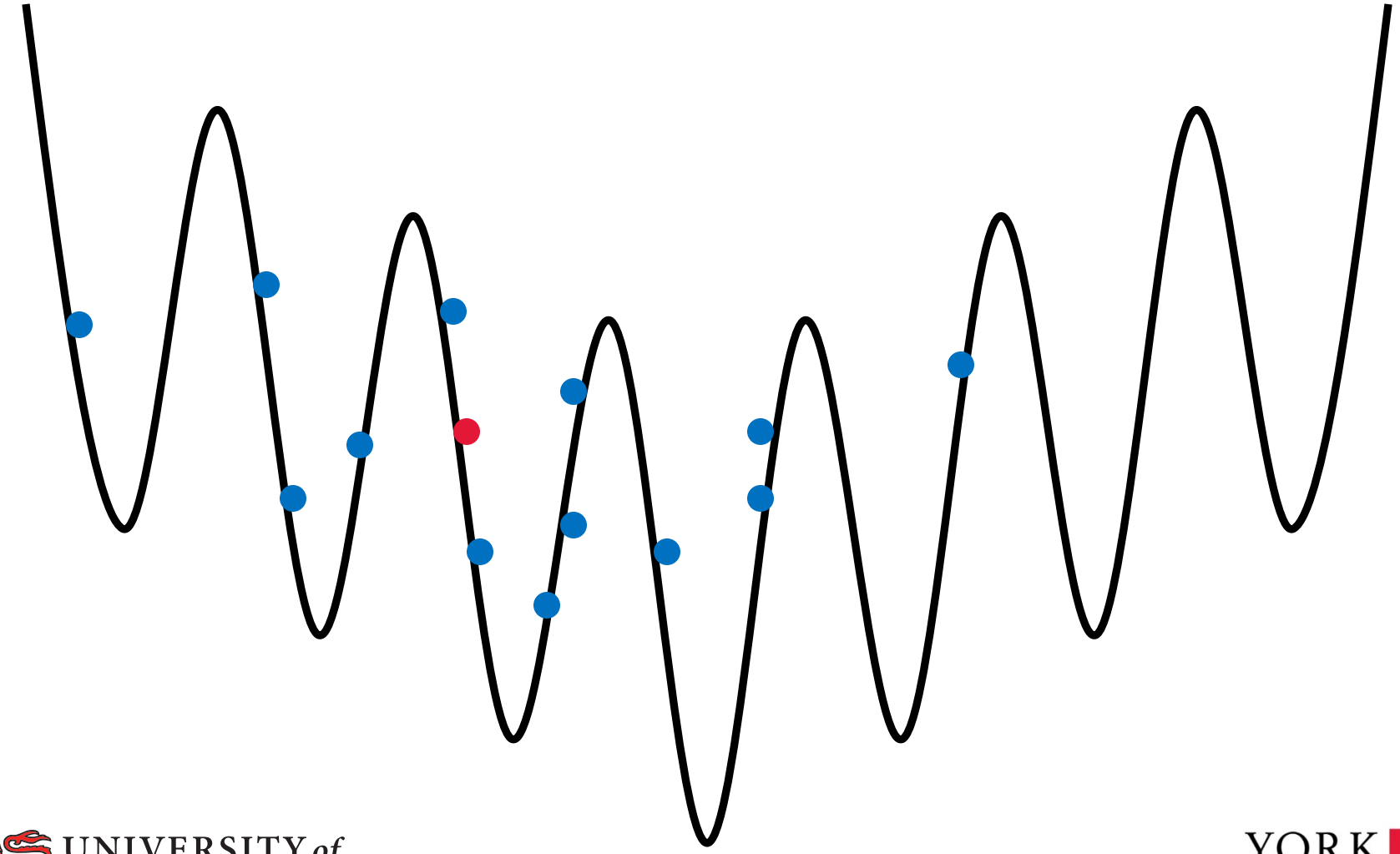
Thresheld convergence

- Convergence is “held” back by a threshold function
- Small, exploitative moves which are smaller than the threshold value are disallowed
- Create an initial phase of pure exploration

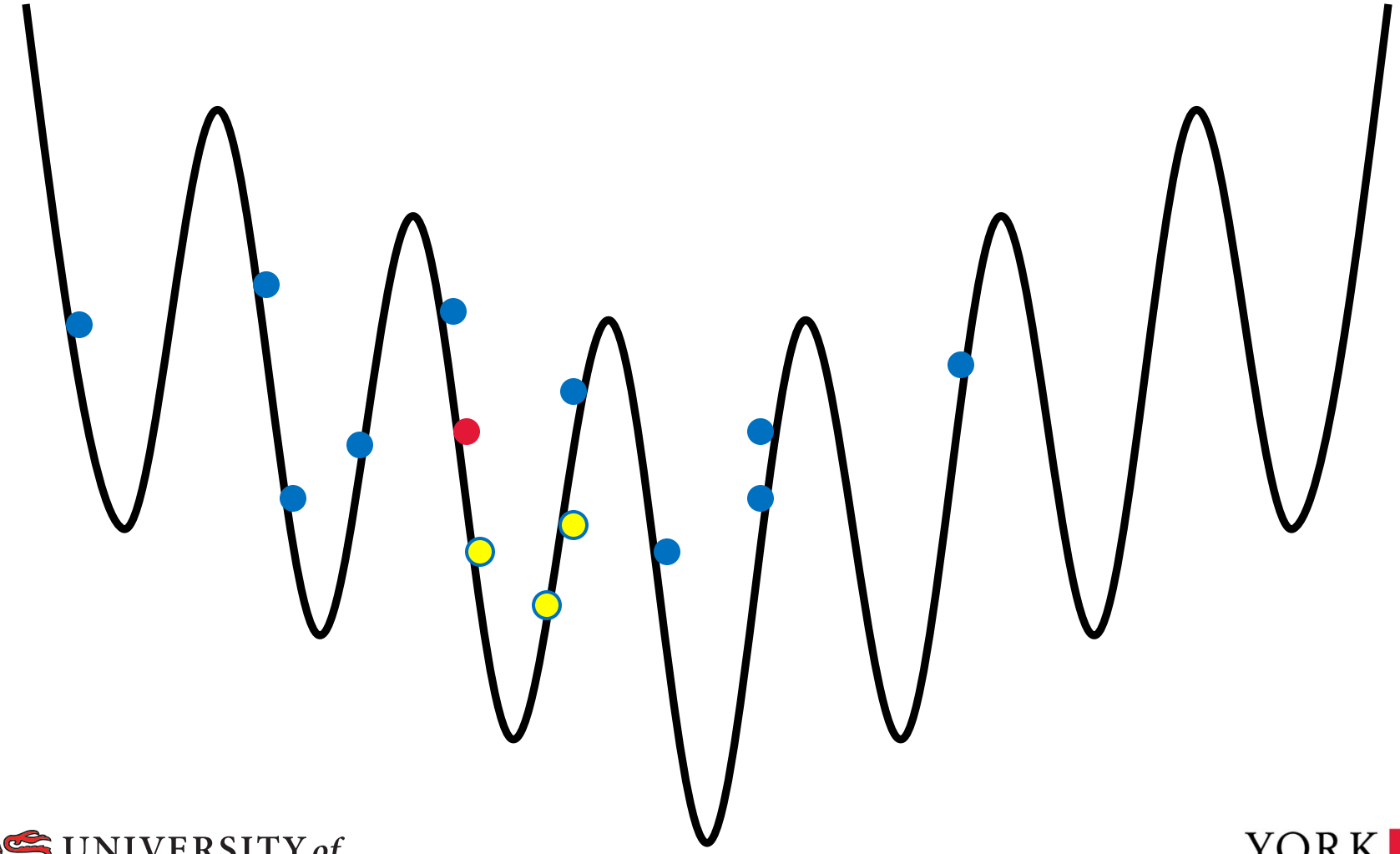
Working Definitions

- A new search point in a new attraction basin is exploration
- A new search point in the same attraction basin is exploitation
- The first search point in a new attraction basin is likely to have a random/average fitness
- Exploitation within an existing attraction basin leads to solutions of better-than-random/above-average fitness

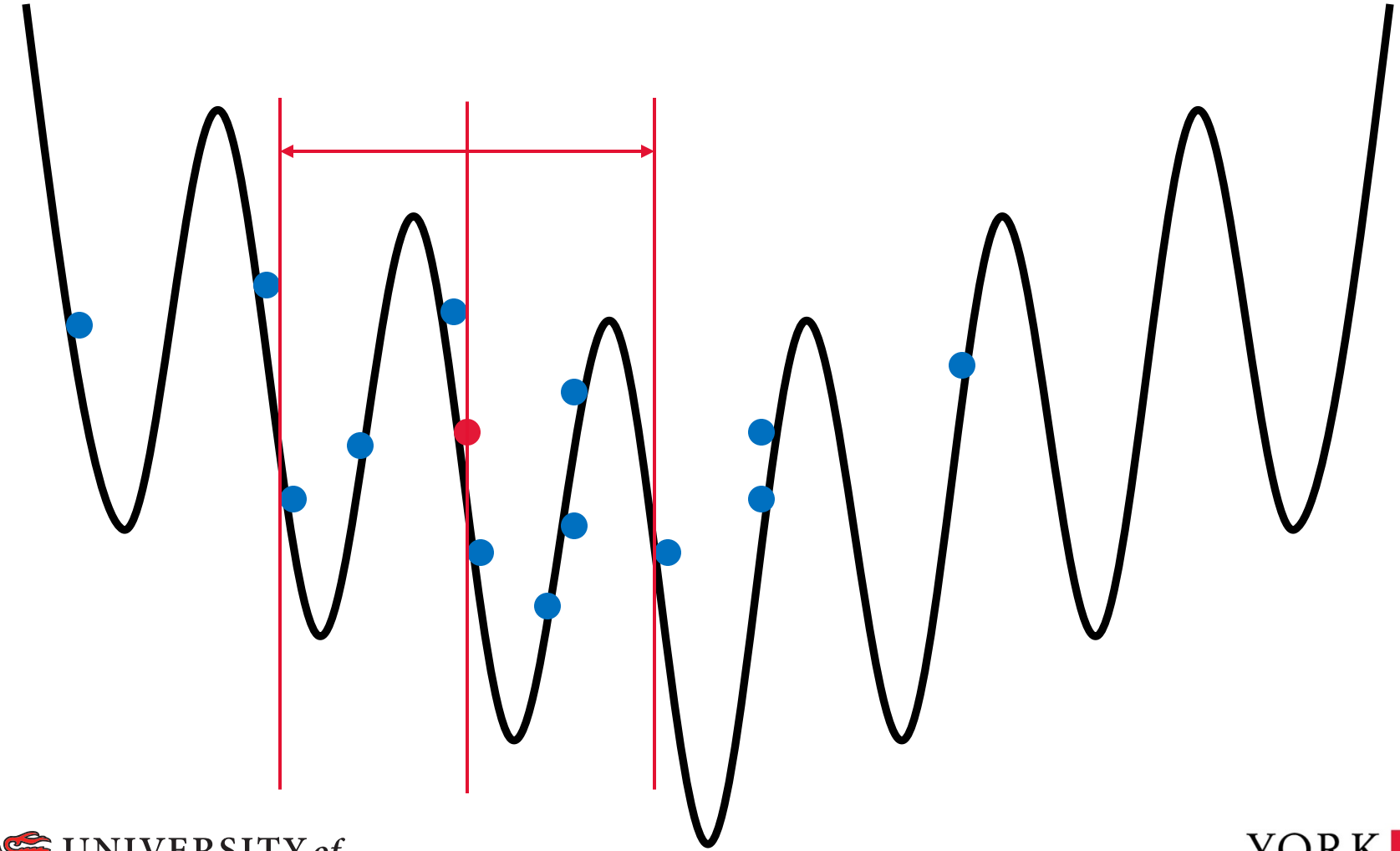
Normal Search – No Minimum Step



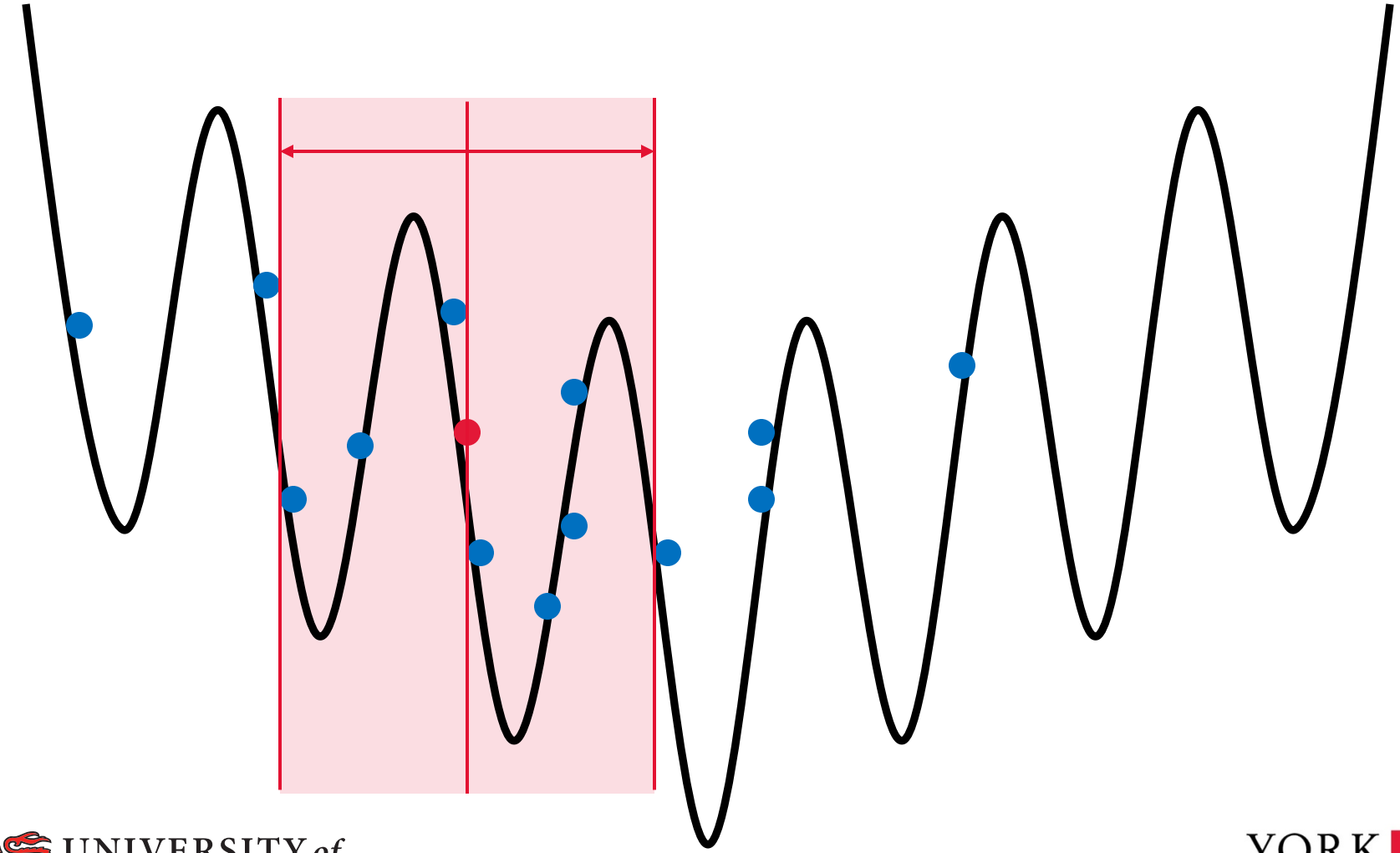
Normal Search – Exploitation Possible



Thresheld Convergence – Minimum Step



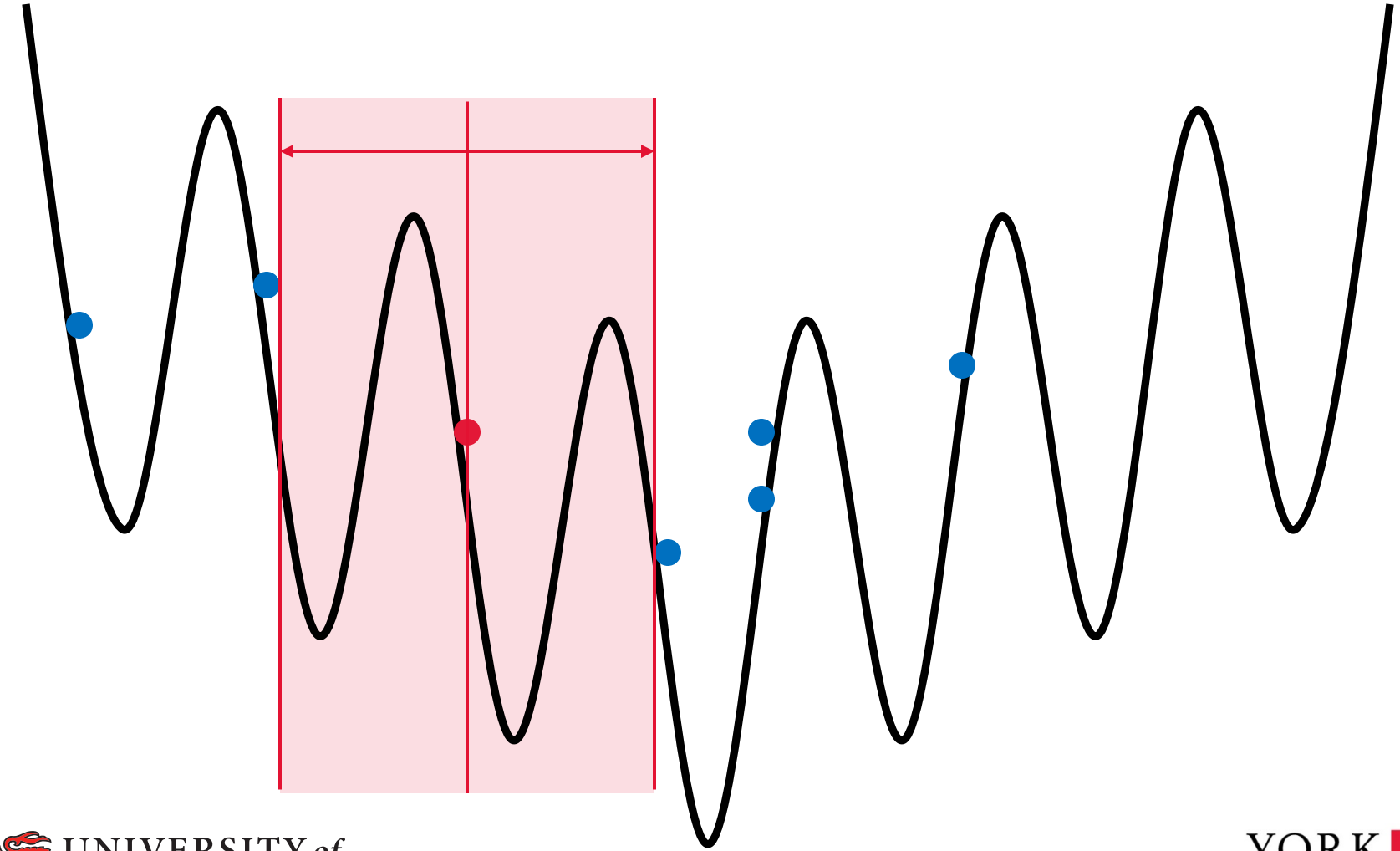
Thresheld Convergence – No Exploitation



UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

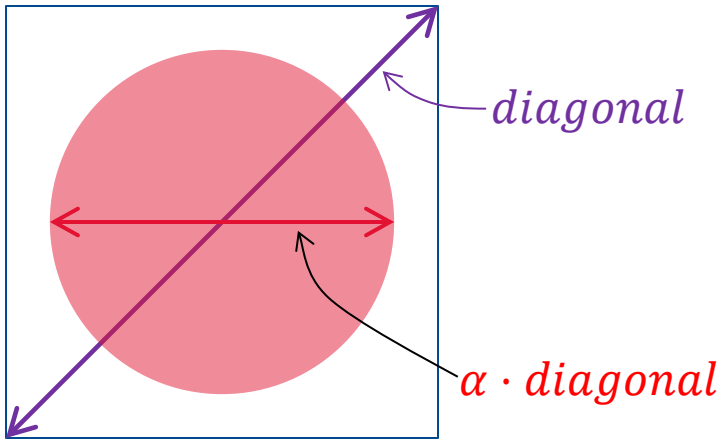
Thresheld Convergence – Exploration Only



A Simple Threshold Function

- Early applications used a polynomial decay function
- In effect, this is exhaustive search to find the size of attraction basins

$$threshold = (\alpha \cdot diagonal) \cdot \left(\frac{FEs - i}{FEs}\right)^{\gamma}$$



Applications of Thresheld Convergence

PSO, DE, and SA

UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

PSO with Thresheld Convergence

- Original update

  if f(**x**) < f(*pbes*t)

  pbest = **x**;

  end if

- New update

  if f(**x**) < f(*pbest*)

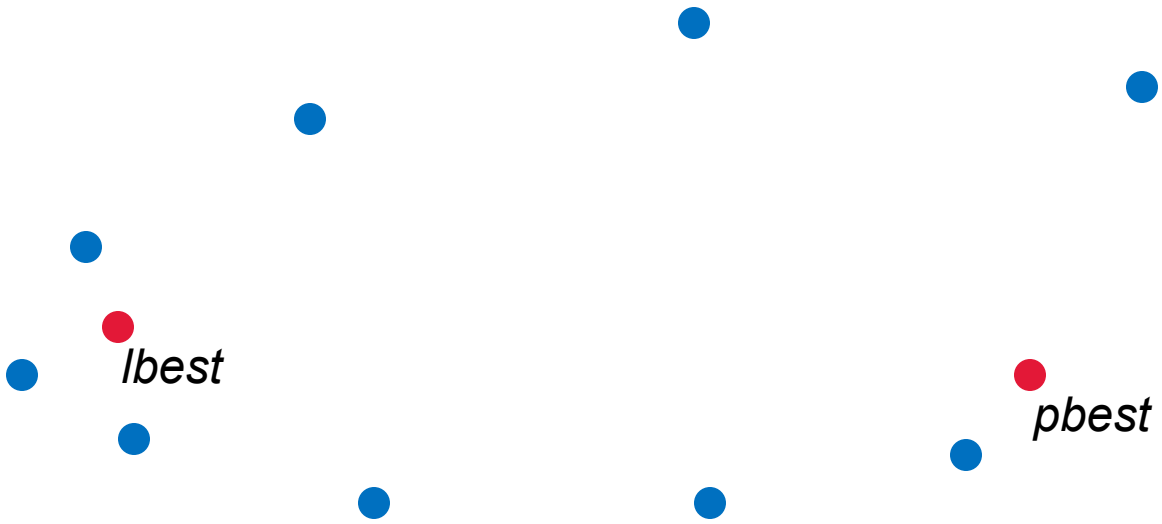
  AND distance (**x**, *pbest*) > *threshold*

  AND distance (**x**, *lbest*) > *threshold*

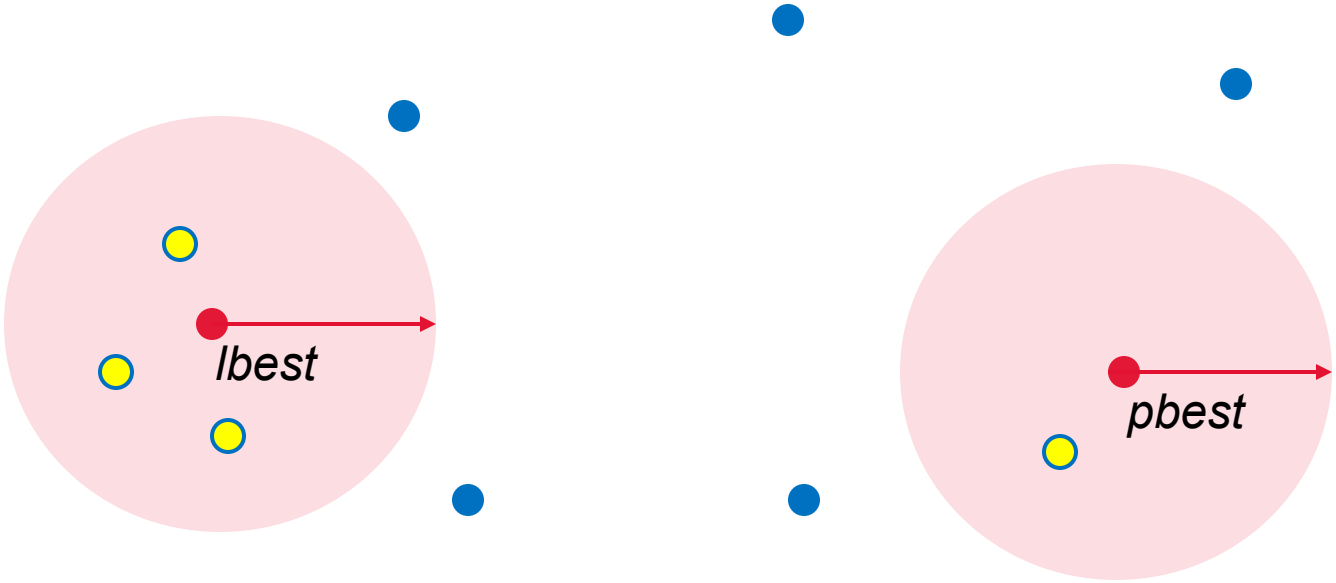
  pbest = **x**;

  end if

PSO with Thresheld Convergence



PSO with Thresheld Convergence



PSO with Thresheld Convergence

- Eliminating search points within the threshold wastes FEs
- Despite inefficiency, performance improves in multi-modal search spaces
- Successful proof of concept

DE with Thresheld Convergence

- Modified update

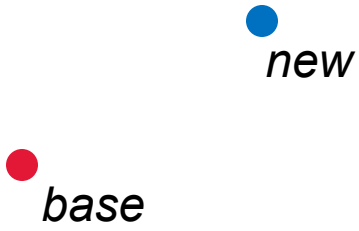
if ||*new* - *base*|| < *threshold*

direction = Normalize(*new* – *base*)

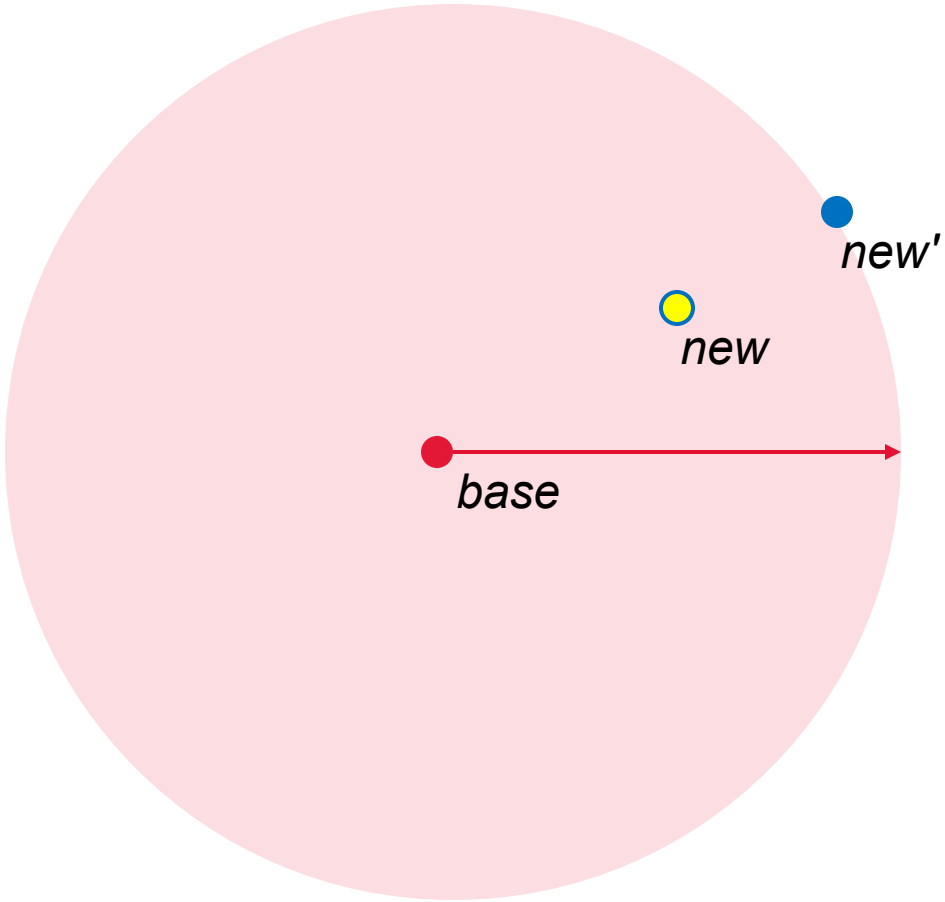
new′ = *base* + *threshold* · direction

end if

DE with Thresheld Convergence



DE with Thresheld Convergence

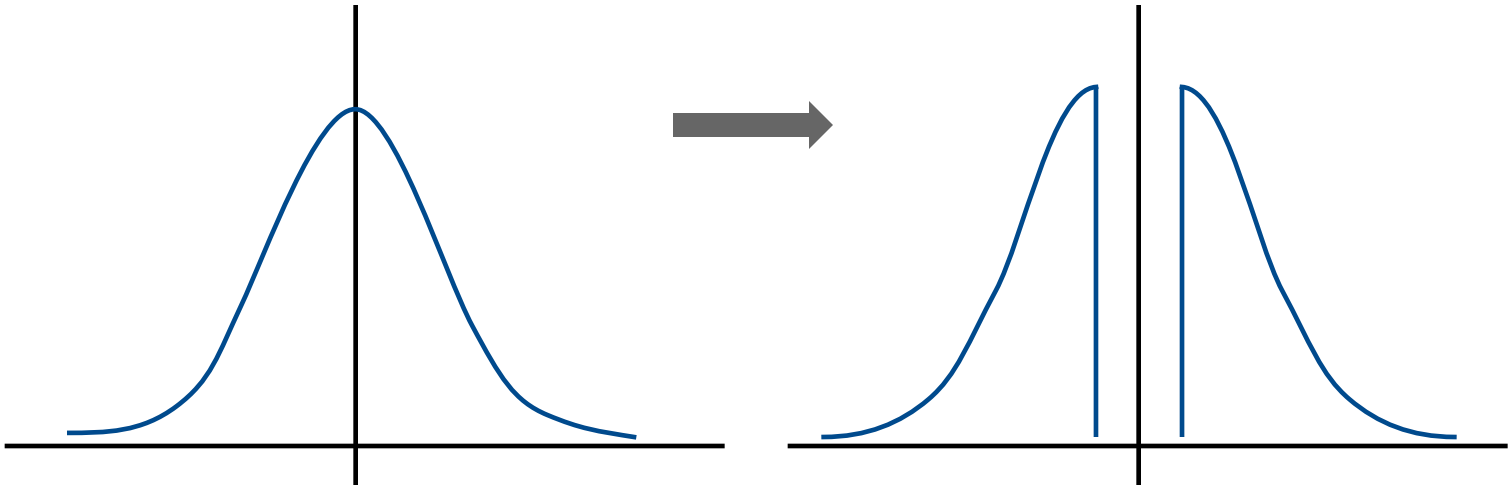


DE with Thresheld Convergence

- No wasted FEs
- Minimum step size both improves exploration (in multi-modal search spaces) and prevents premature convergence (in all search spaces)
- Improved performance across a broad range of benchmark functions

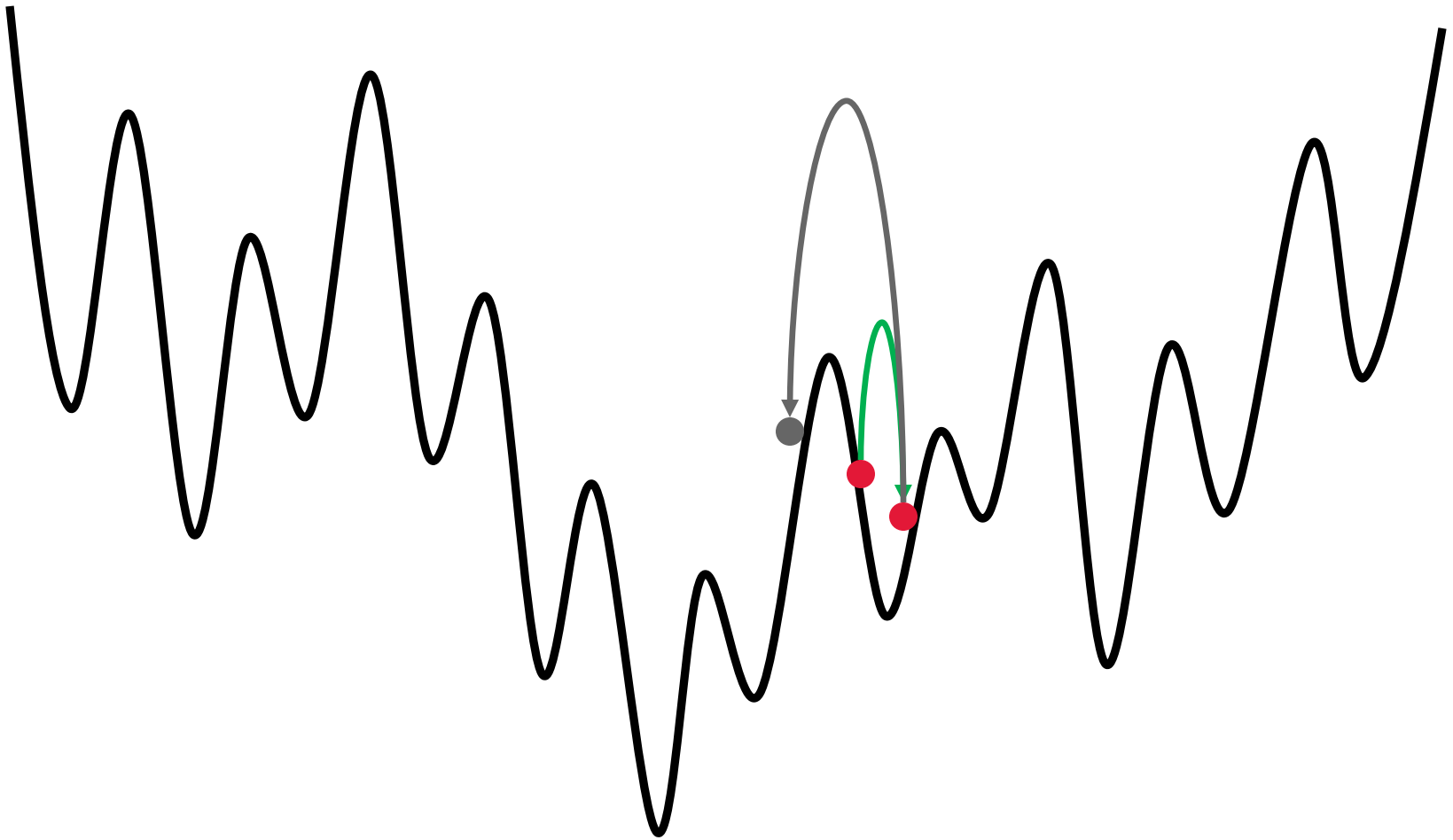
SA with Thresheld Convergence

Create “hollow” sampling distribution

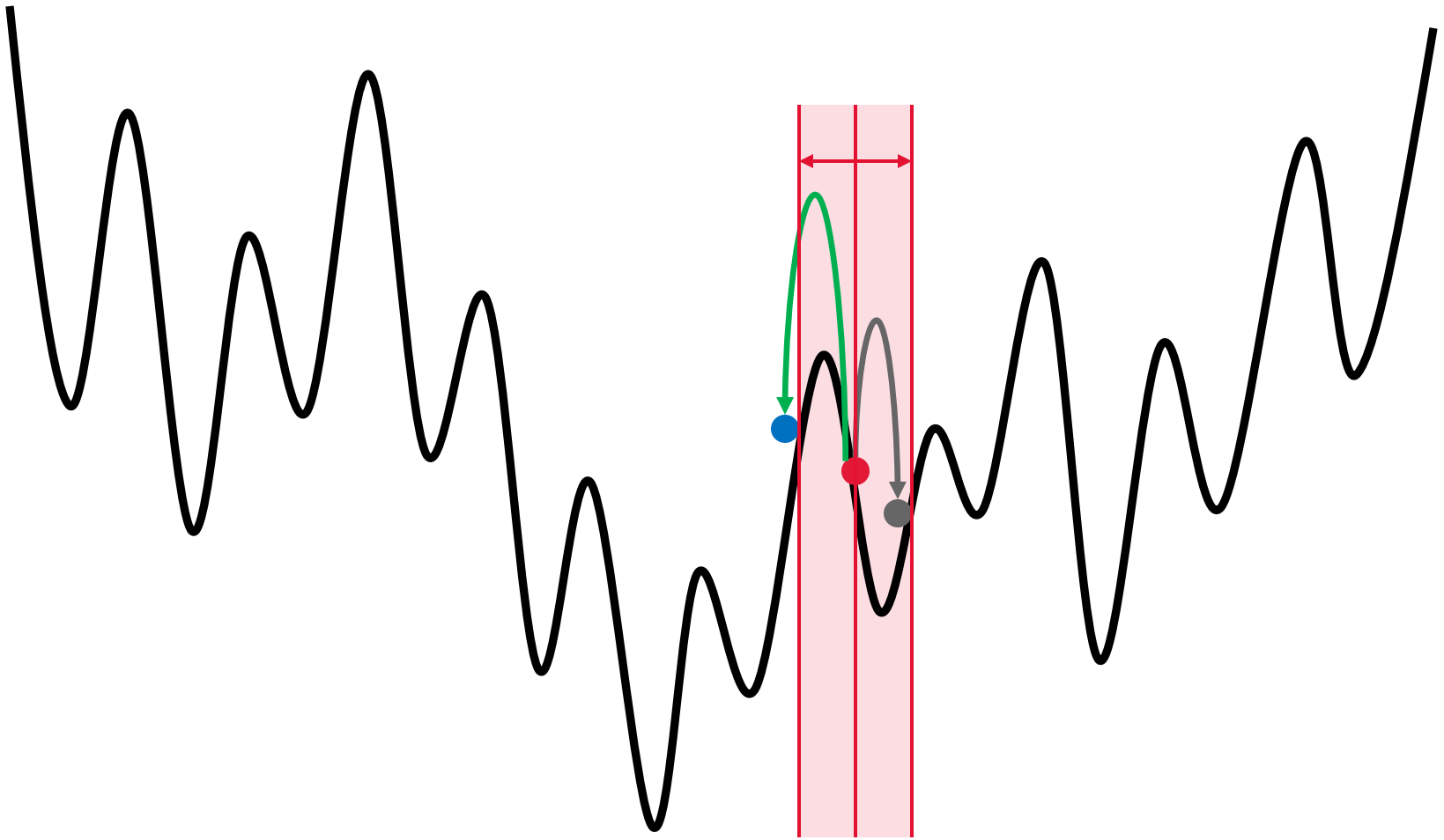


Images from S. Chen, C. Xudiera, and J. Montgomery, “Simulated Annealing with Thresheld Convergence”, IEEE CEC, 2012.

Original Simulated Annealing



SA with Thresheld Convergence



UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

SA with Thresheld Convergence

- Local search/optimization makes escaping from the current attraction basin more difficult
- With threshold, search becomes almost random
  - Concurrent exploration and exploitation is required for useful convergence in a memory-less search system
  - E.g. simulated annealing like a real physical system, genetic algorithms like a real biological system
- SA with thresheld convergence requires a memory

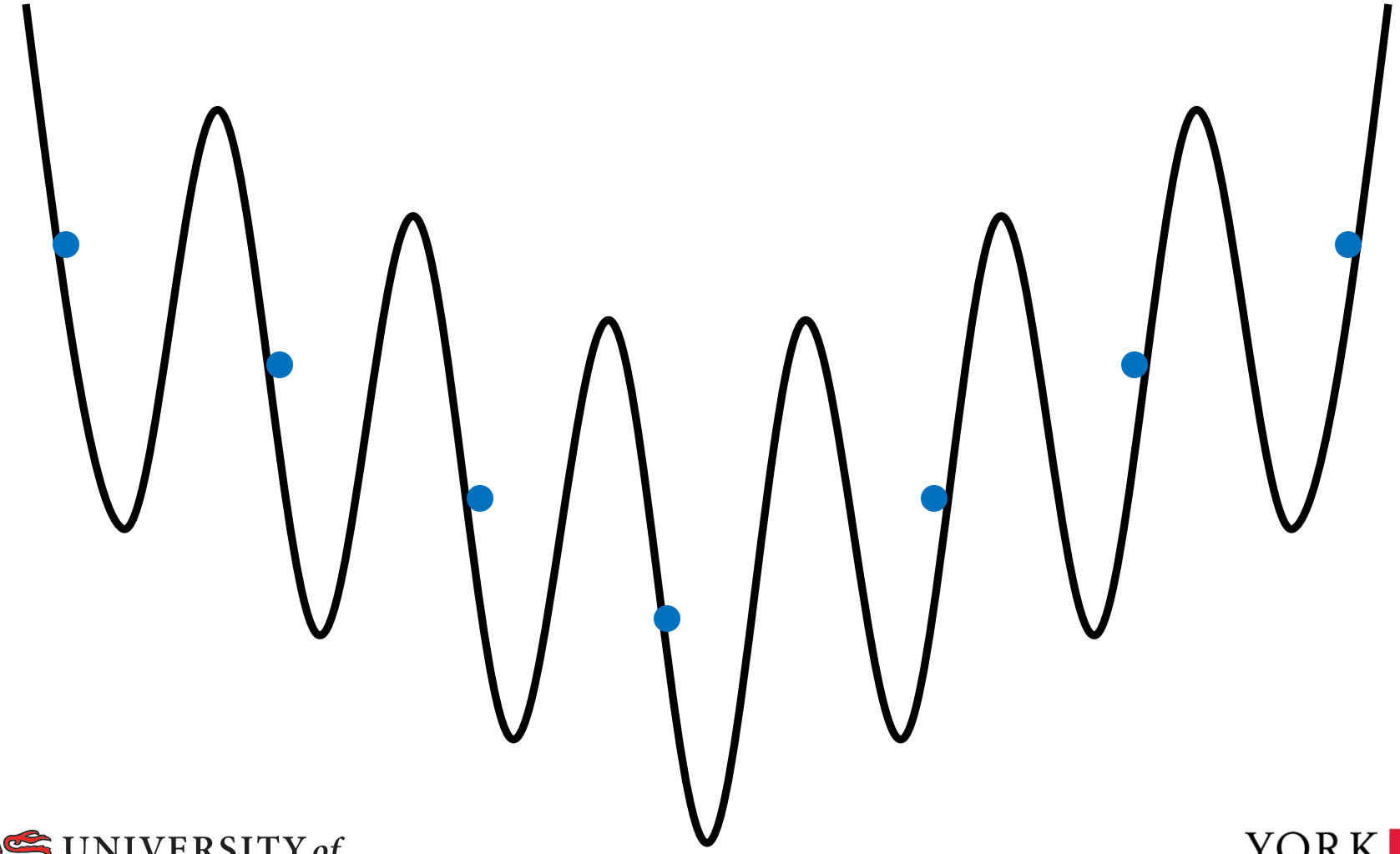
Early Applications of Thresheld Convergence

- Many metaheuristics allow (even require) concurrent exploration and exploitation
- Adding minimum step size **reduces**, but does not **eliminate** concurrent exploration and exploitation
- Experiments on benchmark problems led to improved results in the targeted multi-modal search spaces
- Improvements limited by the simple, preliminary threshold function

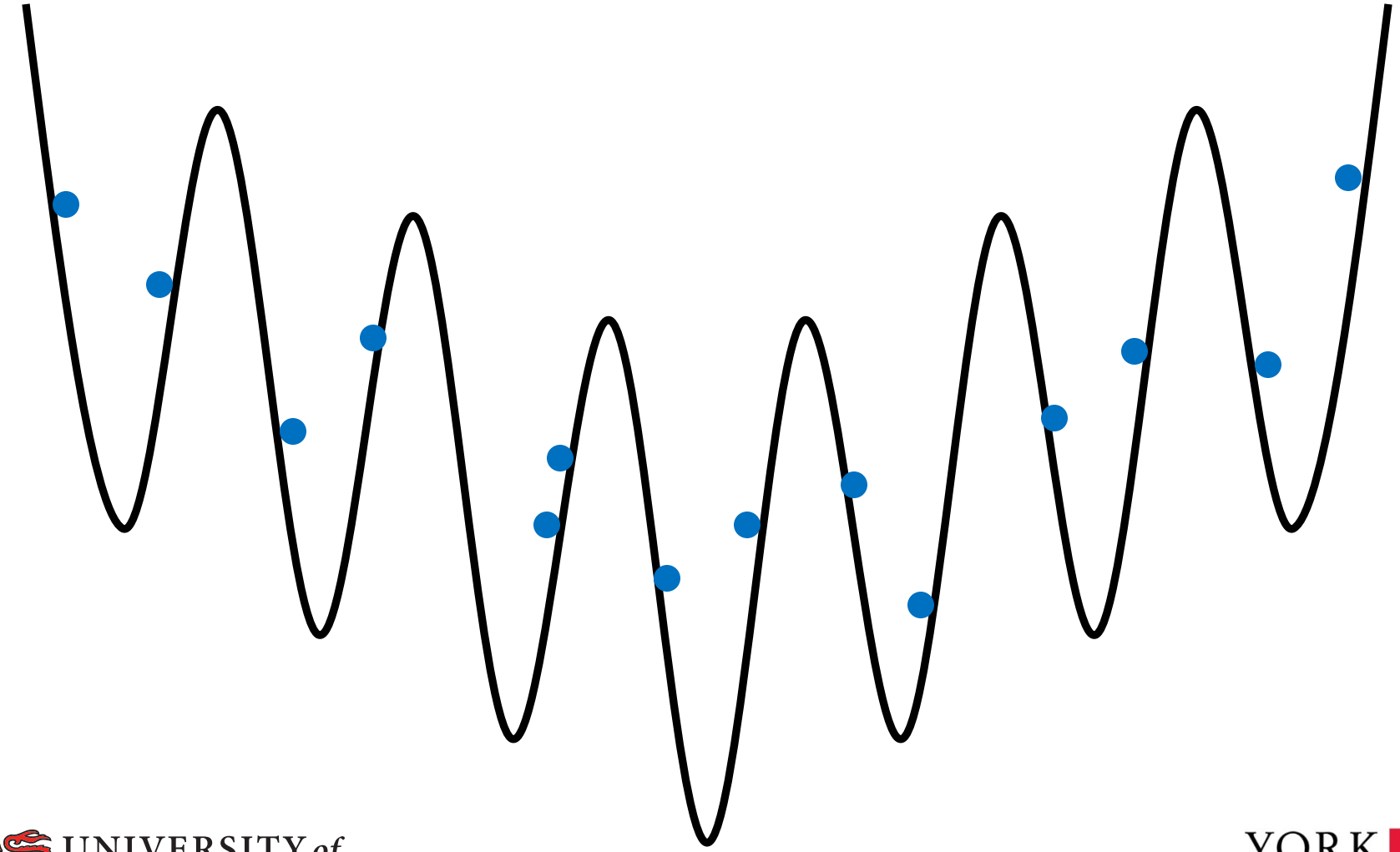
Limitations of Thresheld Convergence

- Ideal case assumes same relative fitness for each sample solution
- Random variations naturally occur in the relative fitness of sample solutions
- Random variations lead to random search
  - Relative fitness improves over time
  - Even when threshold size is larger than attraction basins
  - Even when no exploitation is possible
- Exploration with thresheld convergence still stalls

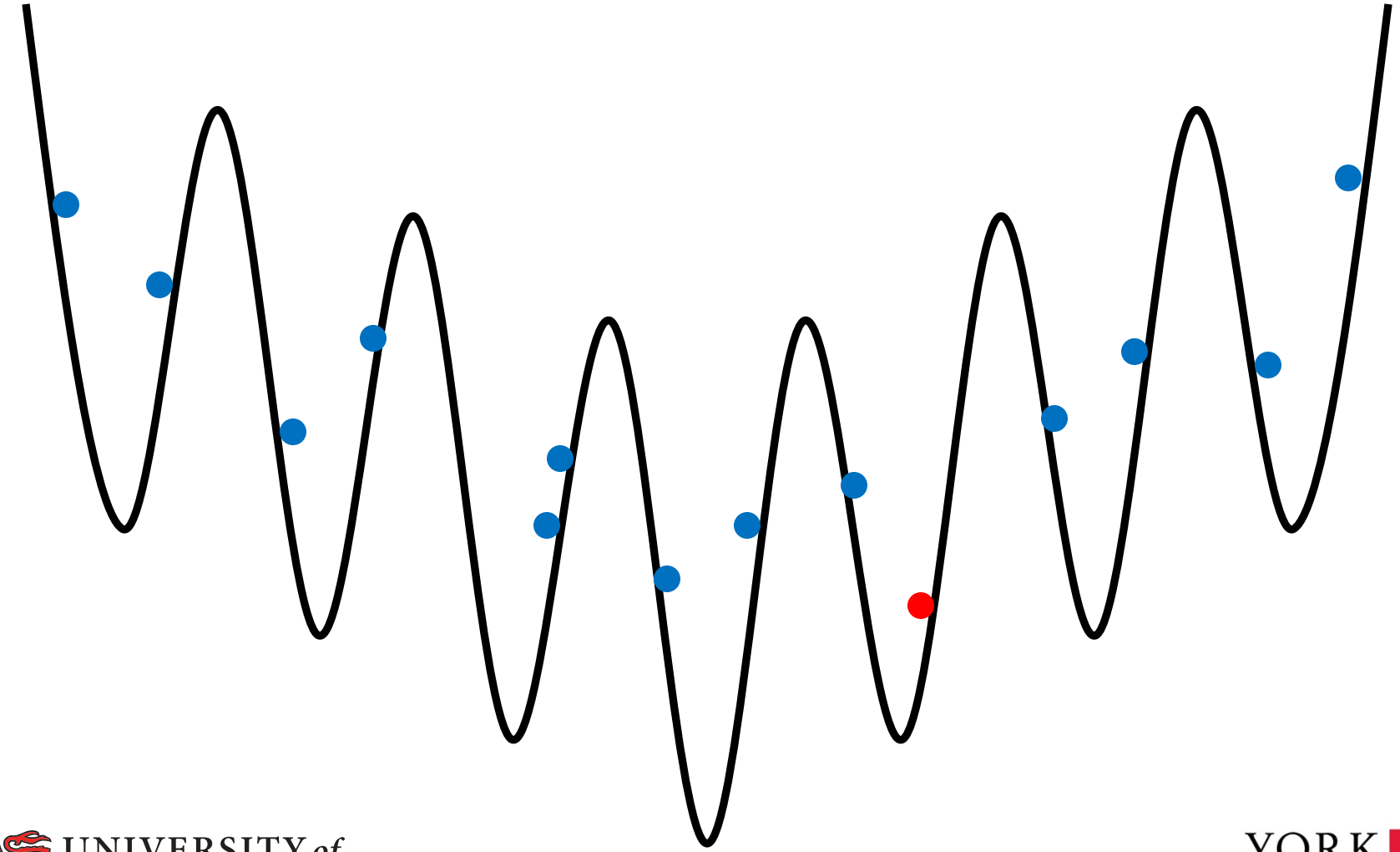
Ideal Case



Typical Case



Typical Case



Limitations of Thresheld Convergence

Random search leads to improved relative fitness

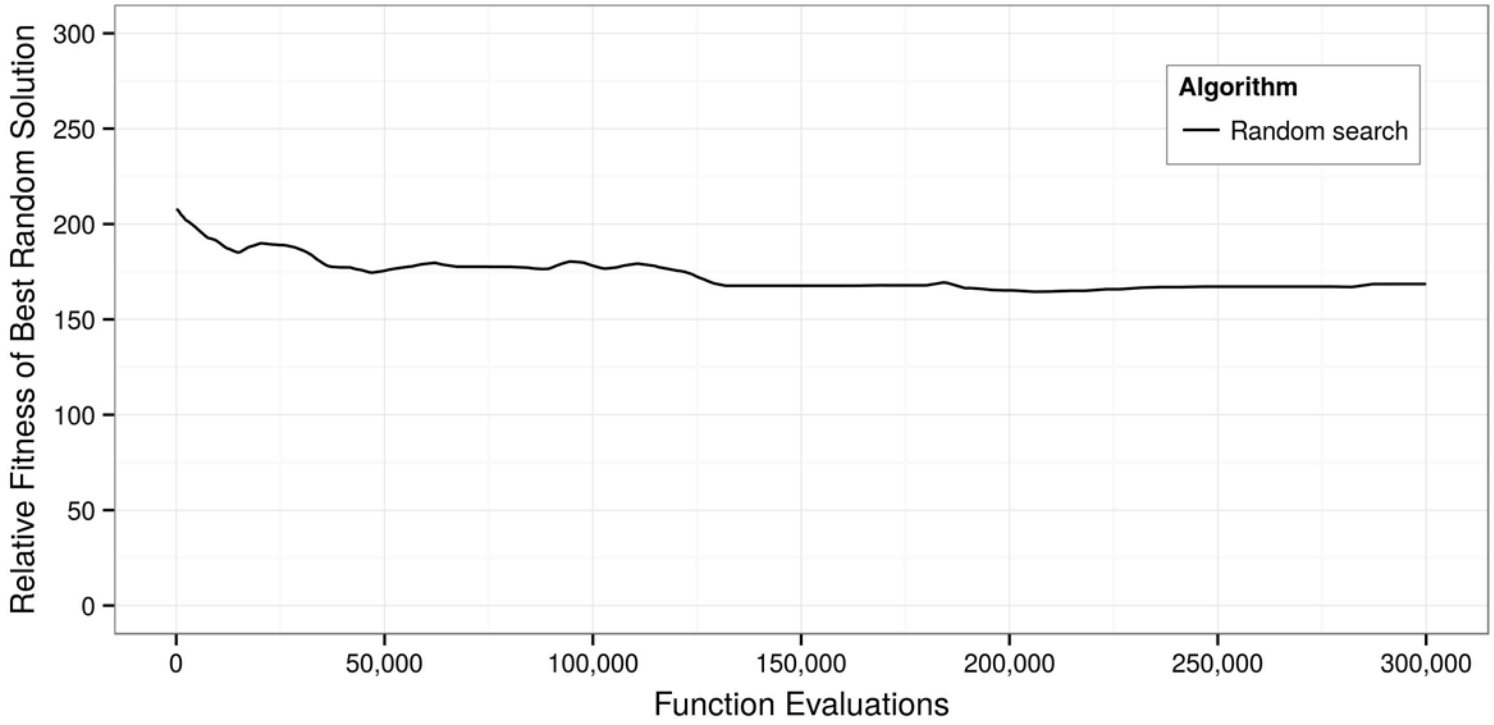


Image from Y. Gonzalez-Fernandez and S. Chen, "Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information", IEEE CEC, 2015.

Limitations of Thresheld Convergence

Better relative fitness means less effective exploration

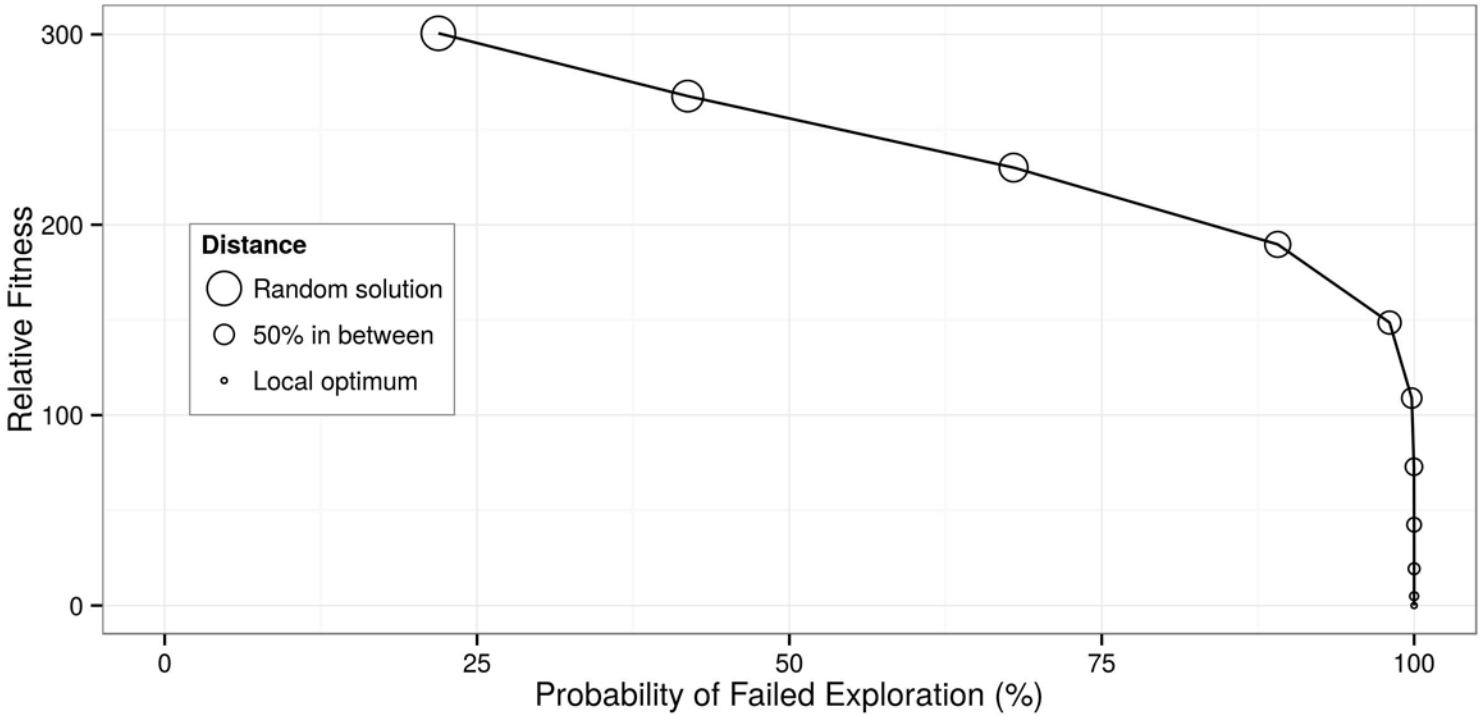
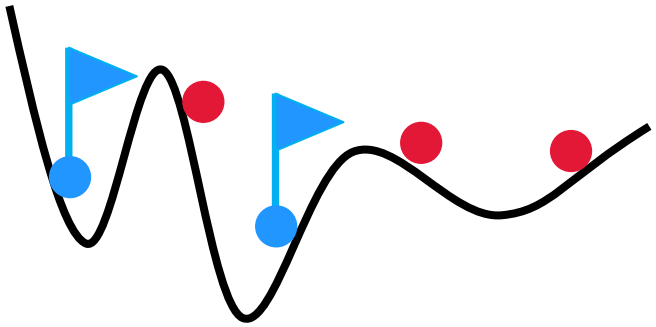


Image from Y. Gonzalez-Fernandez and S. Chen, "Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information", IEEE CEC, 2015.

Towards a New Metaheuristic

Leaders and Followers



UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

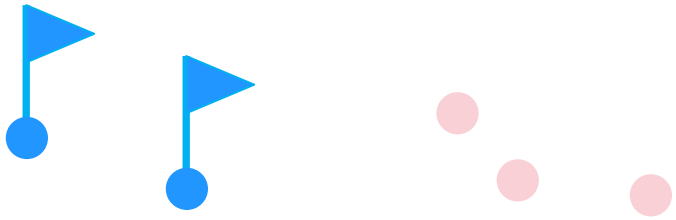
Towards a New Metaheuristic

- Exploration is only effective for a very few number of iterations
- Need rapid restarts
  - Each restart renews opportunity for effective exploration
- Goal of optimization is to find better solutions
  - Better solutions reduce effectiveness of exploration
- Solution: two populations
  - Leaders: best solutions which guide the search
  - Followers: new solutions which can explore effectively

Leaders and Followers

- A new metaheuristic explicitly developed for multi-modal search spaces
- Population control scheme designed to promote accurate comparison of attraction basins

Leaders



The best solutions found to date

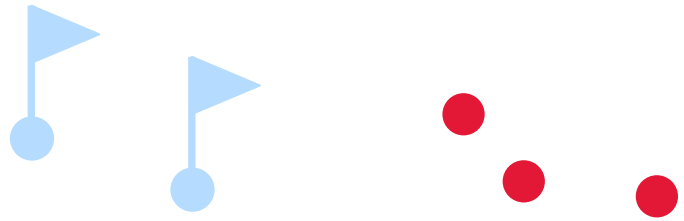
Guide the search

- Like pbests in PSO which guide/attract the current solutions

Do not evaluate the search

- Unlike PSO, new solutions are not compared to leaders/pbests

Followers



The current solutions looking for better attraction basins

Cannot be compared with leaders

- New search points are likely to have poor relative fitness
- Direct comparison with leaders/solutions with high relative fitness will lead to inaccurate comparisons

Compare with other followers

- After each restart, followers can have similar relative fitness

Leaders and Followers

- Leaders lead
- Followers follow
- How can search progress unless followers can become leaders?
- When can leaders and followers be compared fairly?
- Ideally, when both populations have similar relative fitness
- Practically, when both populations have the same median fitness

Leaders and Followers

Initialize Leaders (to random solutions)

Loop allowed iterations/function evaluations

    Initialize Followers to random solutions

    Loop until Followers have similar relative fitness

        Each follower moves towards a leader

        Each follower is best of current or previous position

    End Loop

    Select new Leaders

End Loop

Return best leader

Leaders and Followers

Relative fitness of *pbests* in PSO improves rapidly

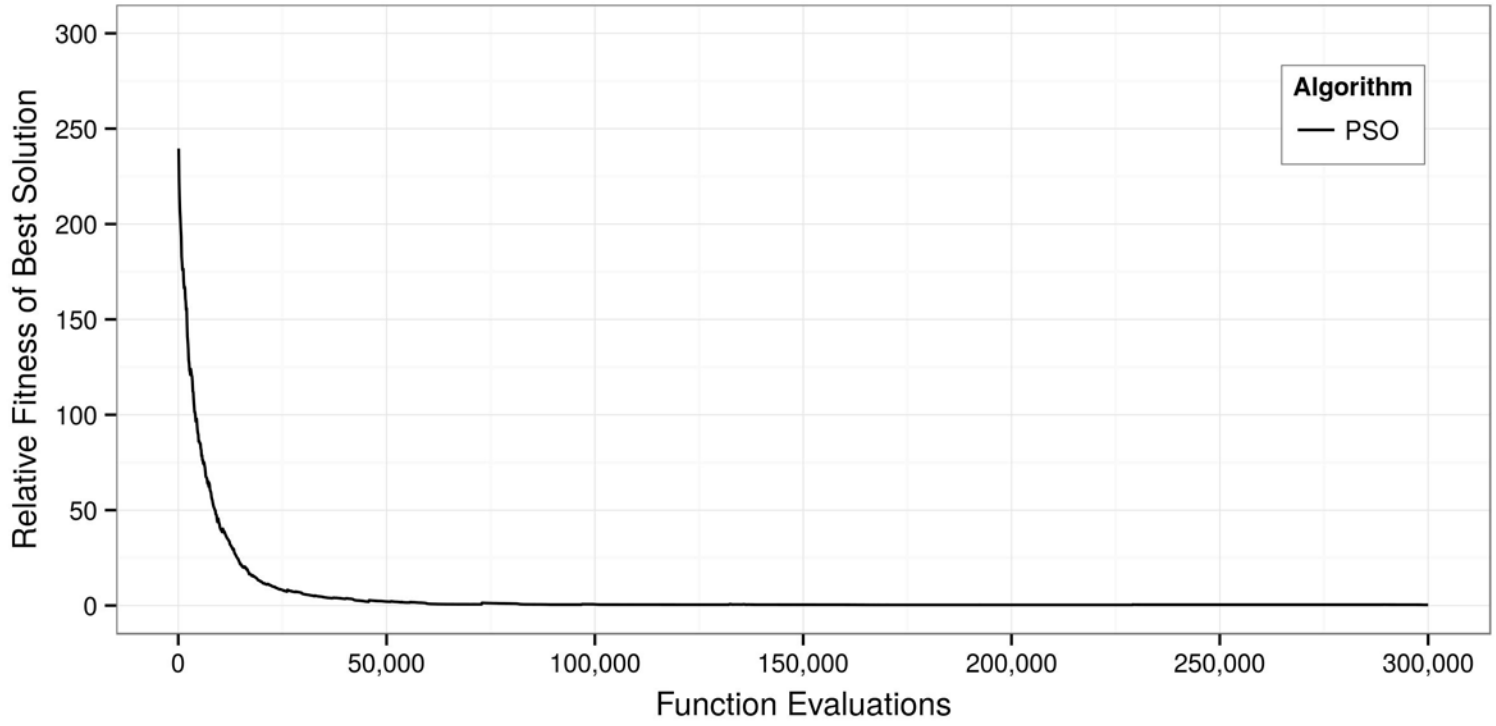


Image from Y. Gonzalez-Fernandez and S. Chen, "Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information", IEEE CEC, 2015.

UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

Leaders and Followers

Relative fitness of leaders also improves rapidly

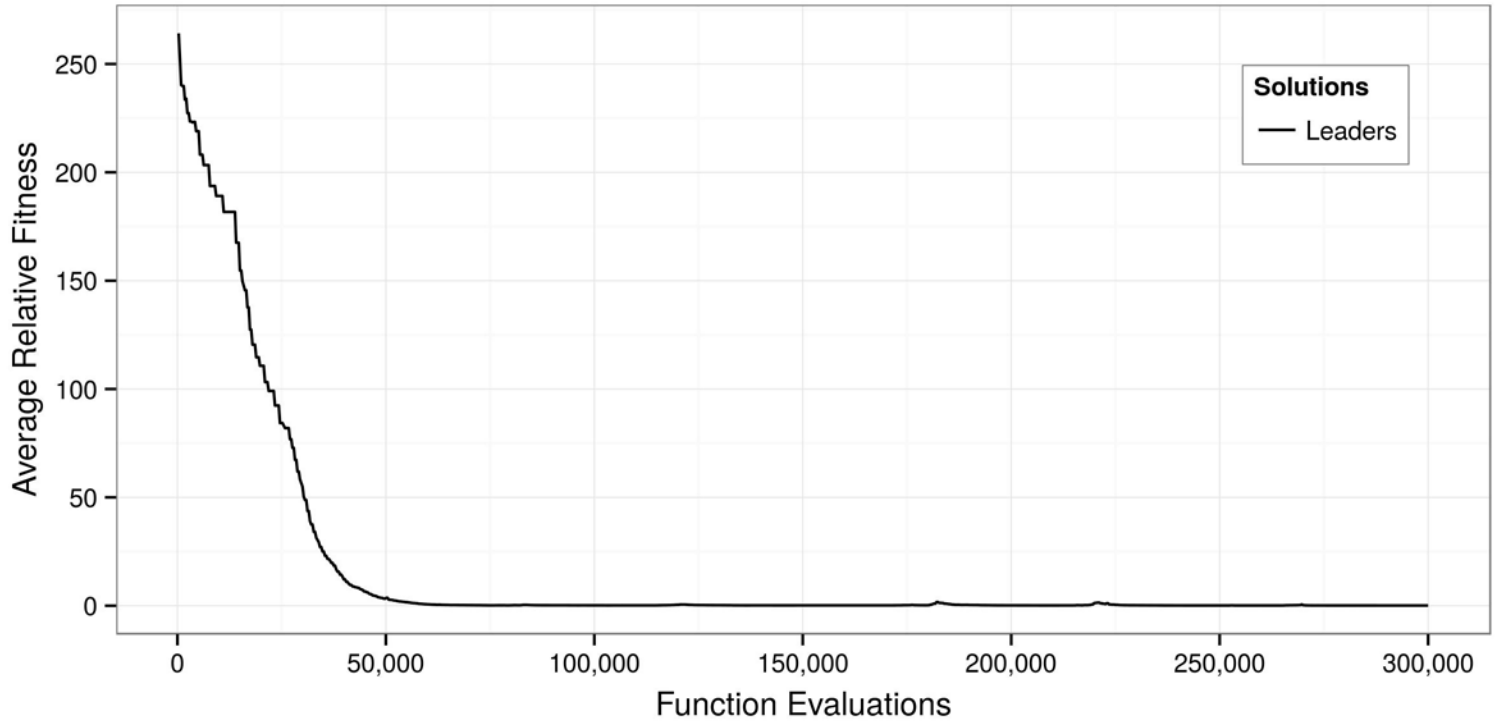


Image from Y. Gonzalez-Fernandez and S. Chen, "Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information", IEEE CEC, 2015.

Leaders and Followers

Search stalls in PSO, does not stall in Leaders and Followers – followers are not compared against leaders

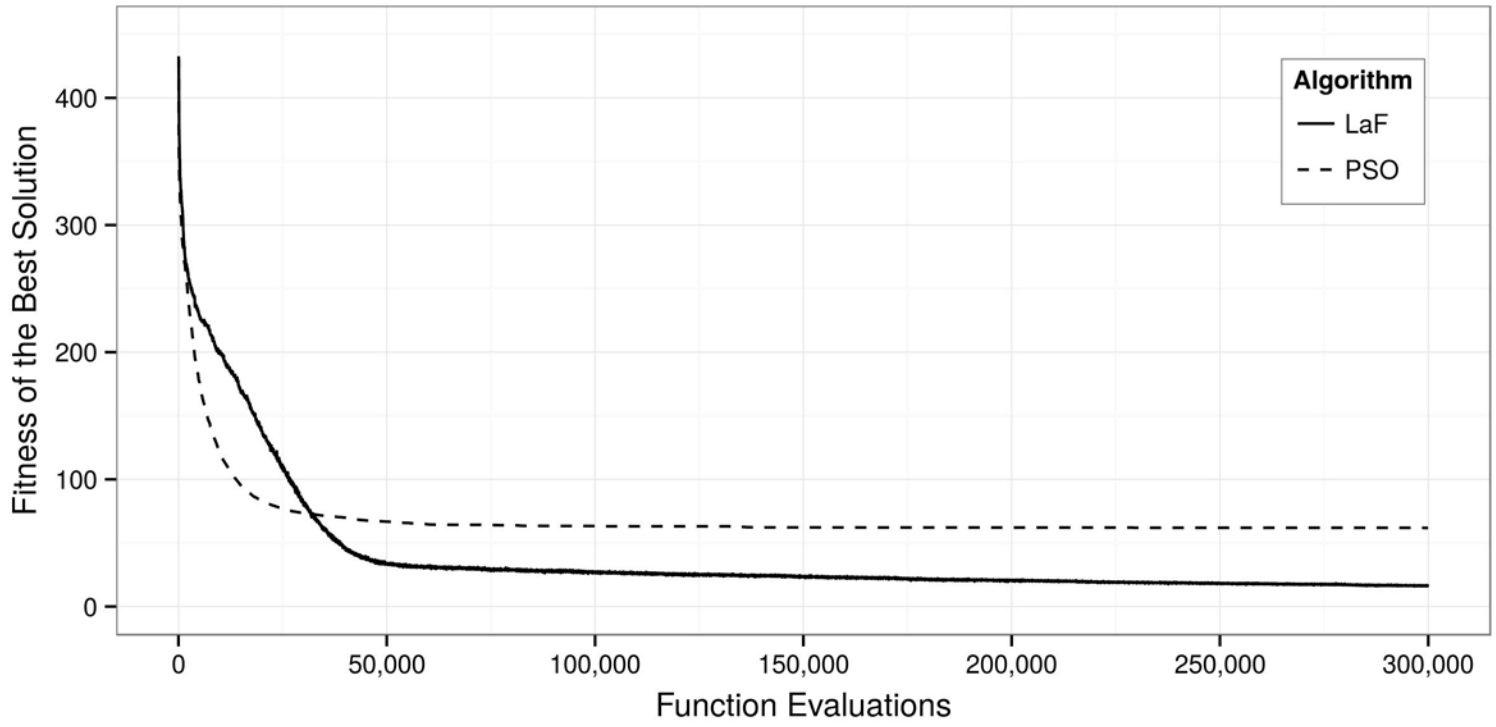
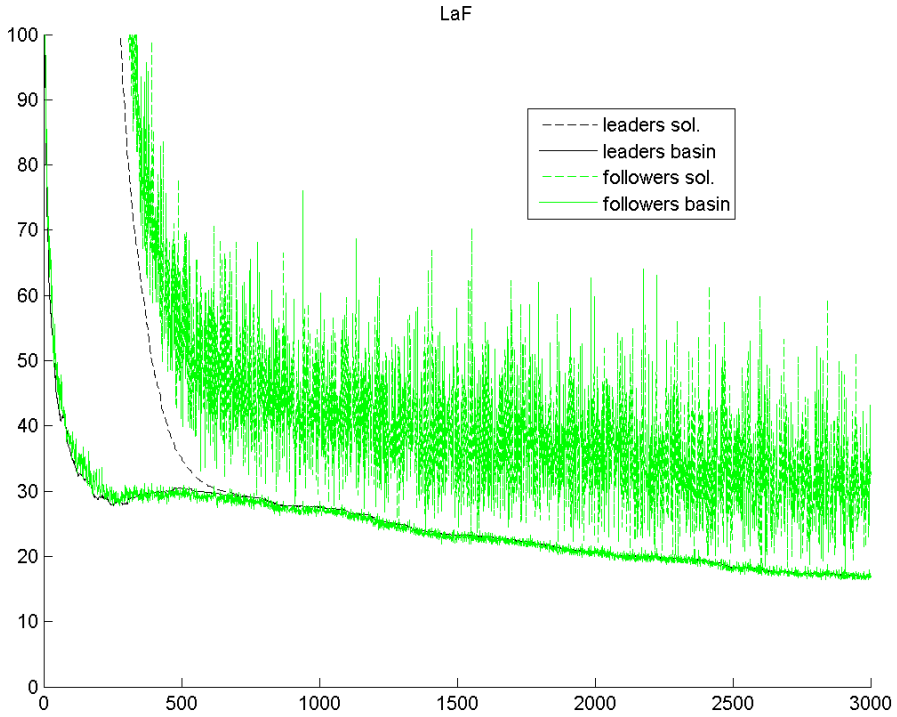


Image from Y. Gonzalez-Fernandez and S. Chen, "Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information", IEEE CEC, 2015.

Leaders and Followers

Each wave of Followers can still find better attraction basins than the Leaders



Leaders and Followers - Operation

- Followers reset every time their median fitness reaches the median fitness of the Leaders
- Each reset allows unbiased comparison of attraction basins
- Improved comparisons lead to improved performance in multi-modal search space
- Rapid restarts are implicit thresheld convergence
  - No threshold function used to encourage diversity

Leaders and Followers - Performance

- LaF performs 20% better than PSO on multi-modal functions in CEC2013 benchmark
- LaF performs 10% better than DE on multi-modal functions in CEC2013 benchmark
- LaF is simpler than both PSO and DE
  - Fewer parameters
  - Less-optimized parameters

Leaders and Followers - Future Work

- Add explicit thresheld convergence
  - Better exploration in Followers
    - Don’t create followers in the same attraction basins
    - Lots of wasted time in unnecessary exploitation
  - Better diversity in Leaders
    - Each leader in its own attraction basin
- Improve search distribution for Followers

Leaders and Followers – Why Bother?

- K. Sörensen, “Metaheuristics—the metaphor exposed,” International Transactions in Operational Research, 2013.
- Leaders and Followers is not metaphor based
  - Based on a detailed study of exploration in multi-modal search spaces (using the Rastrigin function)
- Leaders and Followers is distinctly different from other metaheuristics
  - Key feature is how solutions are compared
  - Explicitly designed for multi-modal search spaces

Thresheld Convergence

and what we have learned from its development

UNIVERSITY of TASMANIA

YORK UNIVERSITÉ UNIVERSITY

Optimization in multi-modal spaces should be two distinct phases

Thresheld convergence focuses on attraction basins in multi-modal search spaces

Optimization should be a two-phase process:

- Phase 1: Exploration → find the fittest attraction basins
- Phase 2: Exploitation → find its local optimum

Accuracy of comparing attraction basins is affected by the relative fitness of their sample solutions

Goal of thresheld convergence is to maintain similar relative fitness of all search solutions

Metaheuristic design should focus on how solutions are compared

Leaders and followers is a new metaheuristic designed explicitly for multi-modal search spaces

Key focus is on how solutions are **compared**

- Avoid comparison of solutions with dissimilar relative fitness

Most metaheuristics focus on how solutions are **created**

- PSO → attraction vectors
- DE → difference vectors
- UMDA → distributions

Improvements to how solutions are created should lead to large performance gains in leaders and followers

Key papers

A Review of Thresheld Convergence

S. Chen, J. Montgomery, A. Bolufé-Röhler, Y. Gonzalez-Fernandez

GECONTEC: Revista Internacional de Gestión del Conocimiento y la Tecnología, Vol. 3(1), pp 1-13, 2015.

Leaders and Followers – A New Metaheuristic to Avoid the Bias of Accumulated Information

Y. Gonzalez-Fernandez and S. Chen

CEC2015, pp 776-783.