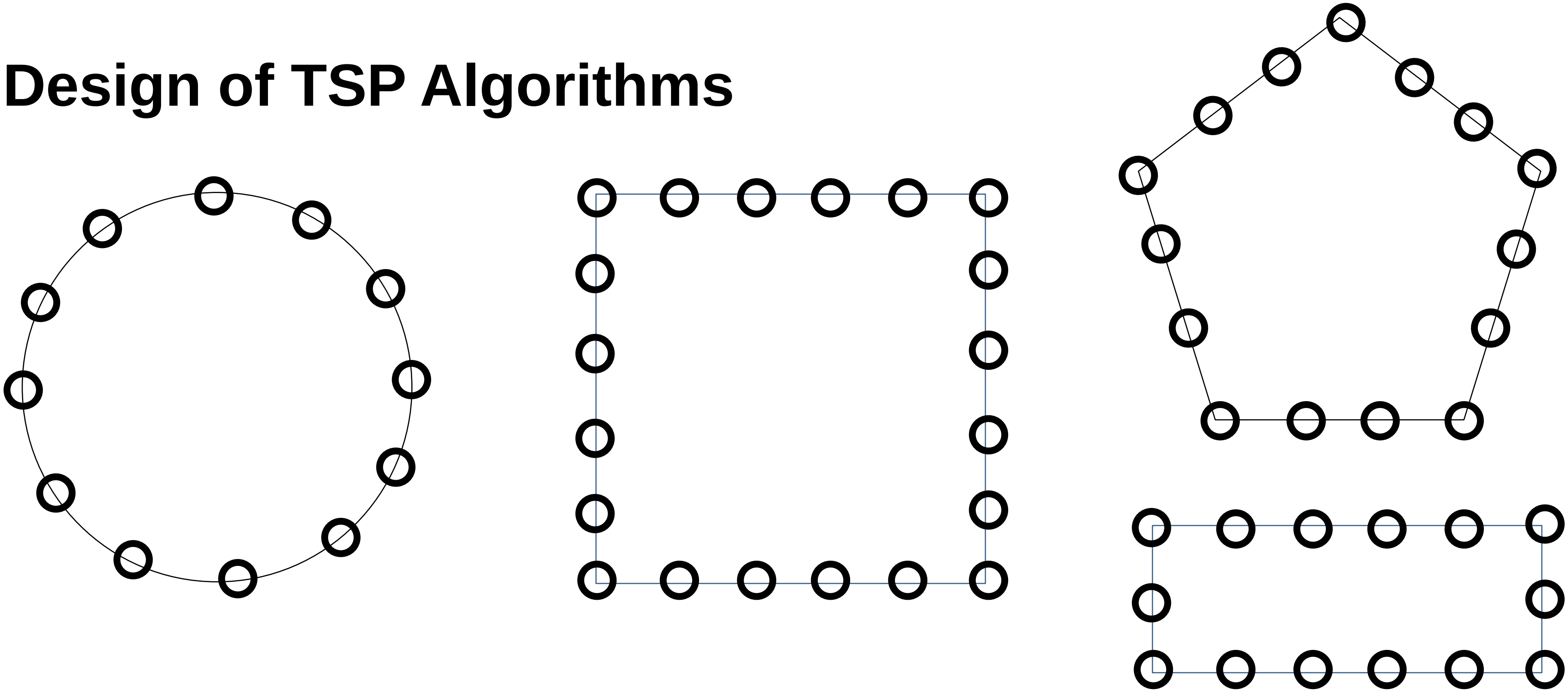
Optimization Methods

1. Introduction.
2. Greedy algorithms for combinatorial optimization.
3. LS and neighborhood structures for combinatorial optimization.
4. Variable neighborhood search, neighborhood descent, SA, TS, EC.
5. Branch and bound algorithms, and subset selection algorithms.
6. Linear programming problem formulations and applications.
7. Linear programming algorithms.
8. Integer linear programming algorithms.
9. Unconstrained nonlinear optimization and gradient descent.
10. Newton's methods and Levenberg-Marquardt modification.
11. Quasi-Newton methods and conjugate direction methods.
12. Nonlinear optimization with equality constraints.
13. Nonlinear optimization with inequality constraints.
14. Problem formulation and concepts in multi-objective optimization.
15. Search for single final solution in multi-objective optimization.
16: Search for multiple solutions in multi-objective optimization.

Algorithm Design

Algorithm design totally depends on test problems.

Design of TSP Algorithms

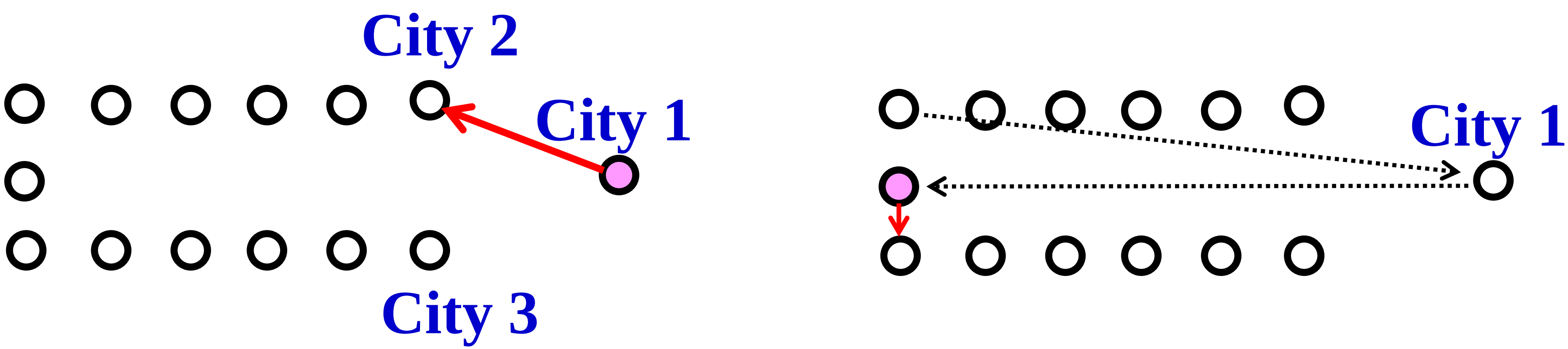


The nearest neighbor greedy algorithm from a randomly selected city always generates the optimal tour for all problems.

Algorithm Design

Algorithm design totally depends on test problems.

Design of TSP Algorithms



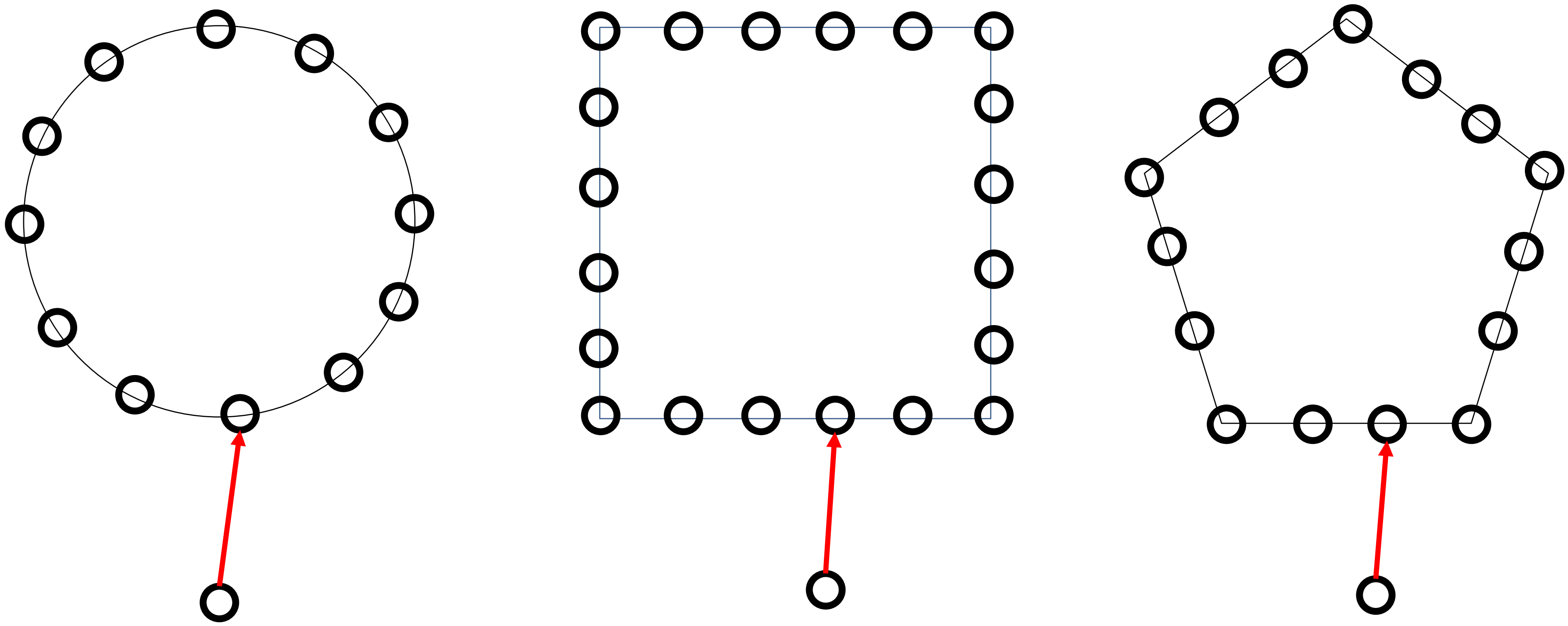
Only when City 1, 2 or 3 is used as the start city, the optimal tour is obtained by the nearest neighbor greedy algorithm.

==> The examination of all cities as the start city is a good idea.
The optimal tour is obtained.

Algorithm Design

Algorithm design totally depends on test problems.

Design of TSP Algorithms

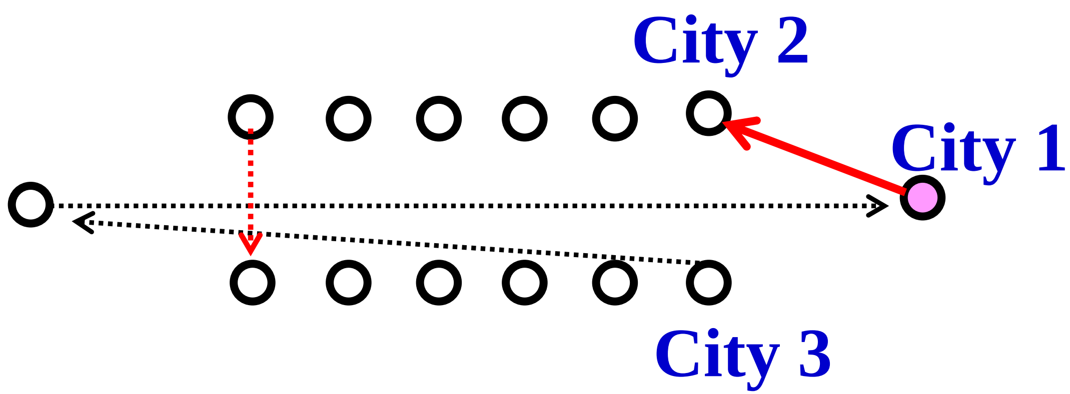


The examination of all cities as the start city ==> Always optimal.

Algorithm Design

Algorithm design totally depends on test problems.

Design of TSP Algorithms

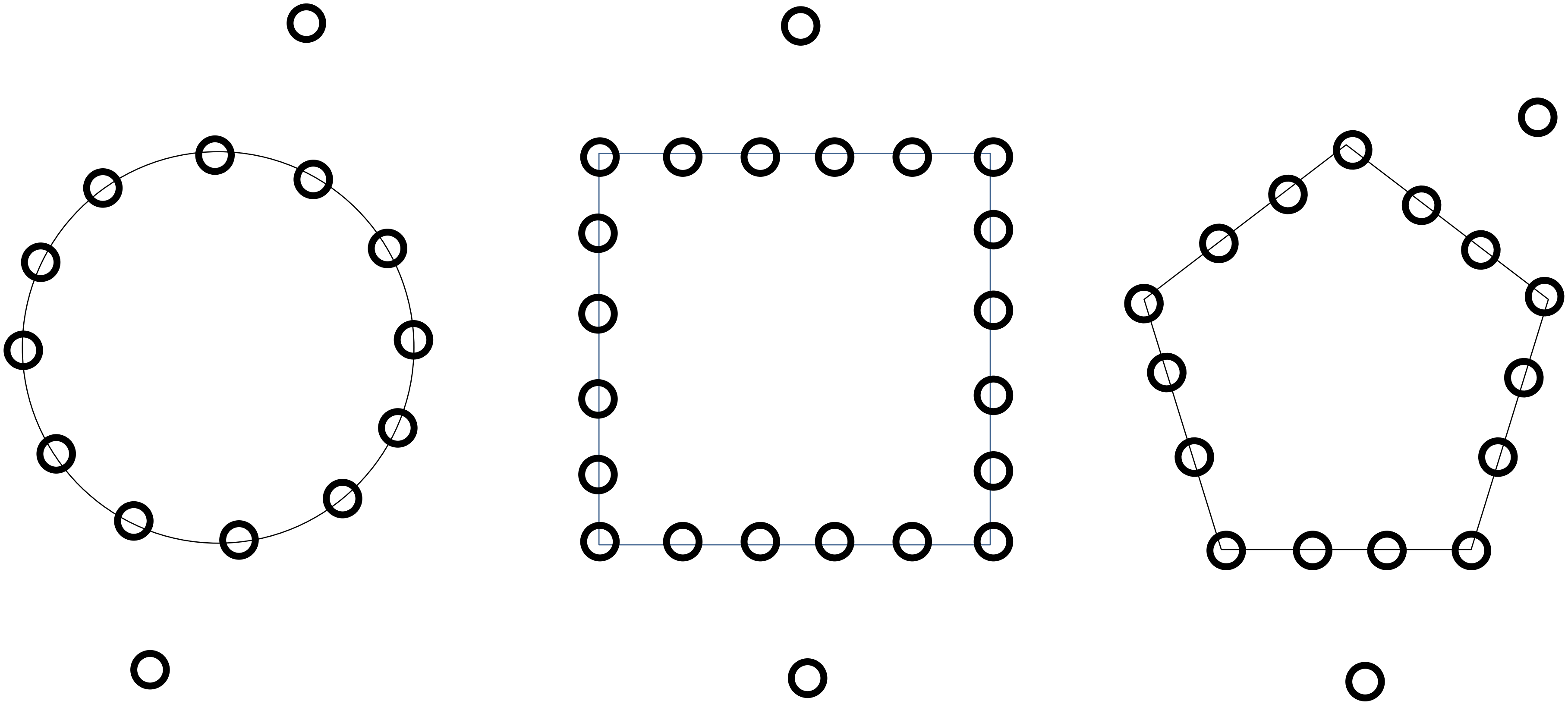


Independent of the choice of the start city, the optimal tour is never obtained by the nearest neighbor greedy algorithm.

==> Other heuristic or meta-heuristic algorithms may be needed.

Algorithm Design

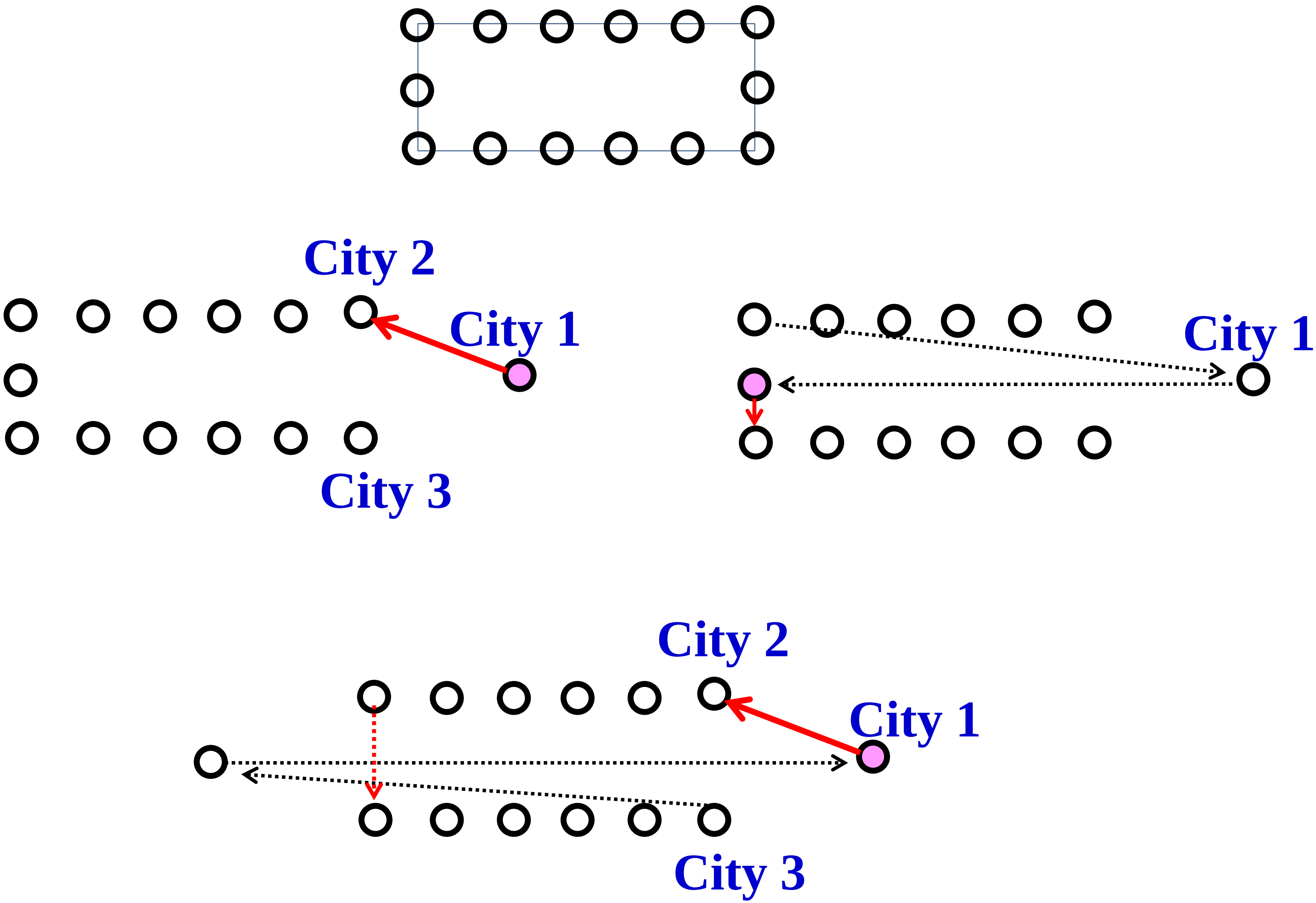
Algorithm design totally depends on test problems.



Independent of the choice of the start city, the optimal tour is never obtained by the nearest neighbor greedy algorithm.

==> Other heuristic or meta-heuristic algorithms may be needed.

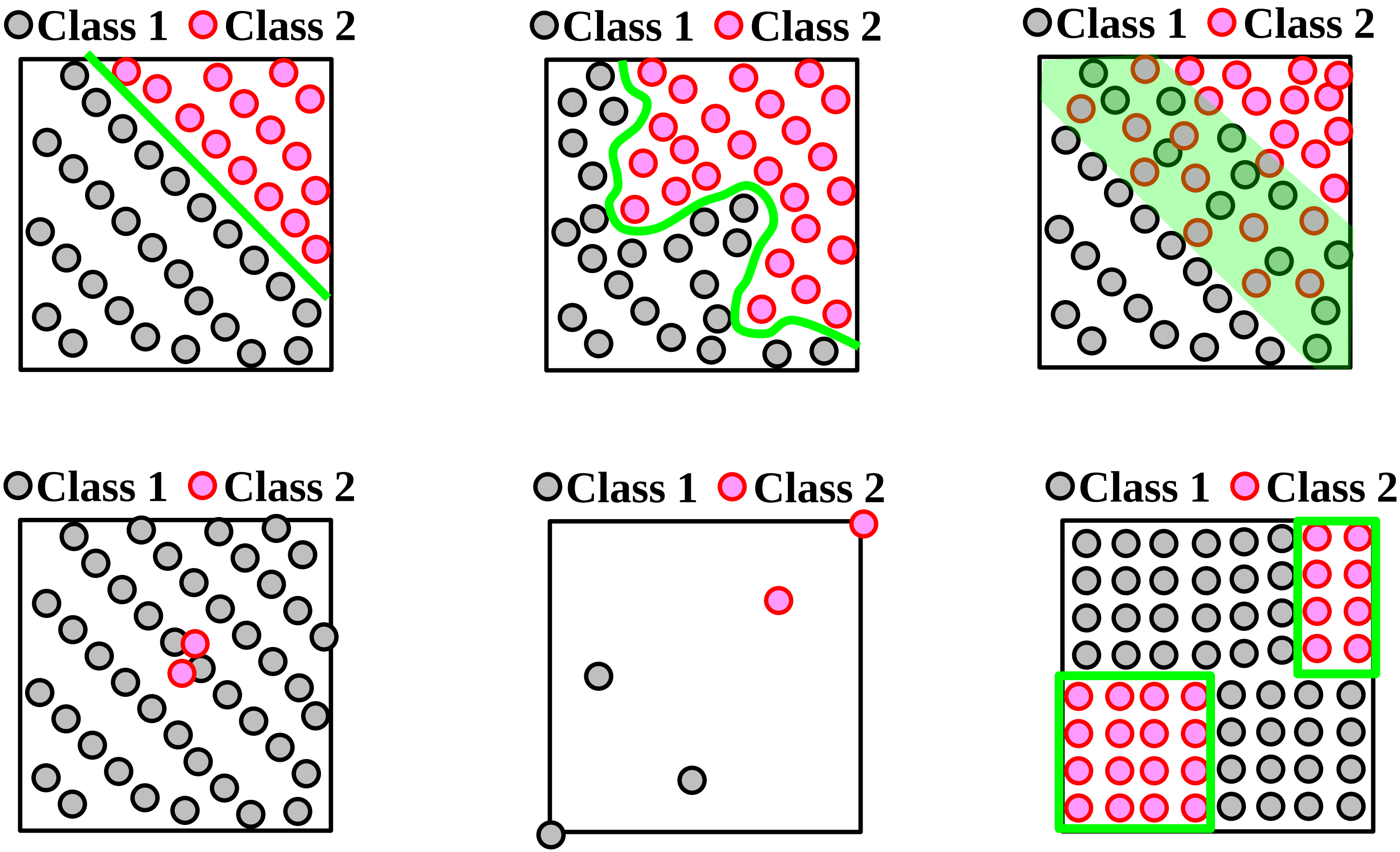
Choice of test problems is very important in algorithm design.



Choice of test problems is very important in algorithm design.

This can be applicable to almost all research fields.

e.g., Machine Learning



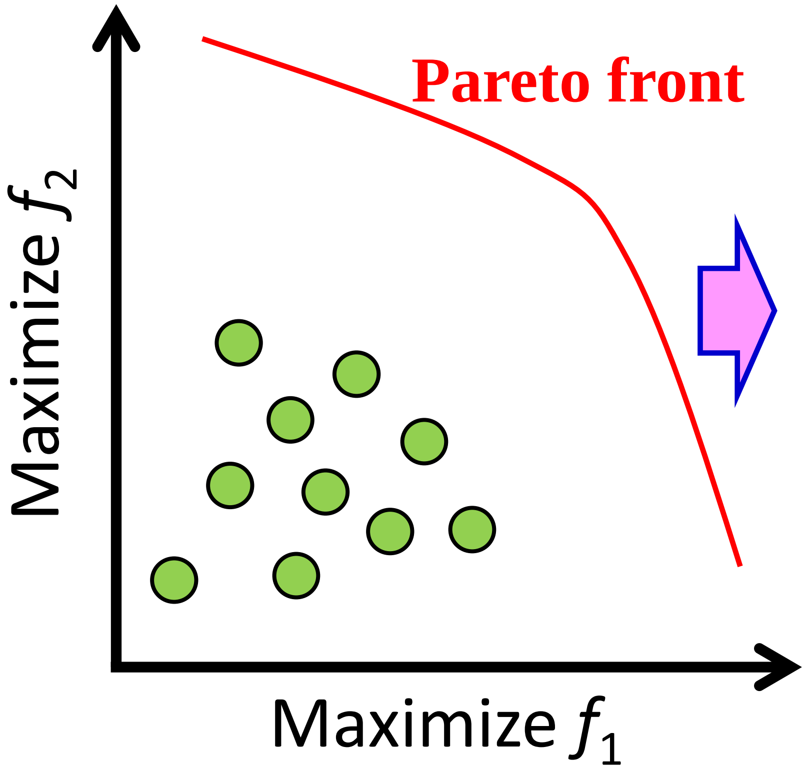
Another Viewpoint of Algorithm Design

Many researchers usually want to use
- their own proposed algorithms
- algorithms which they know very well

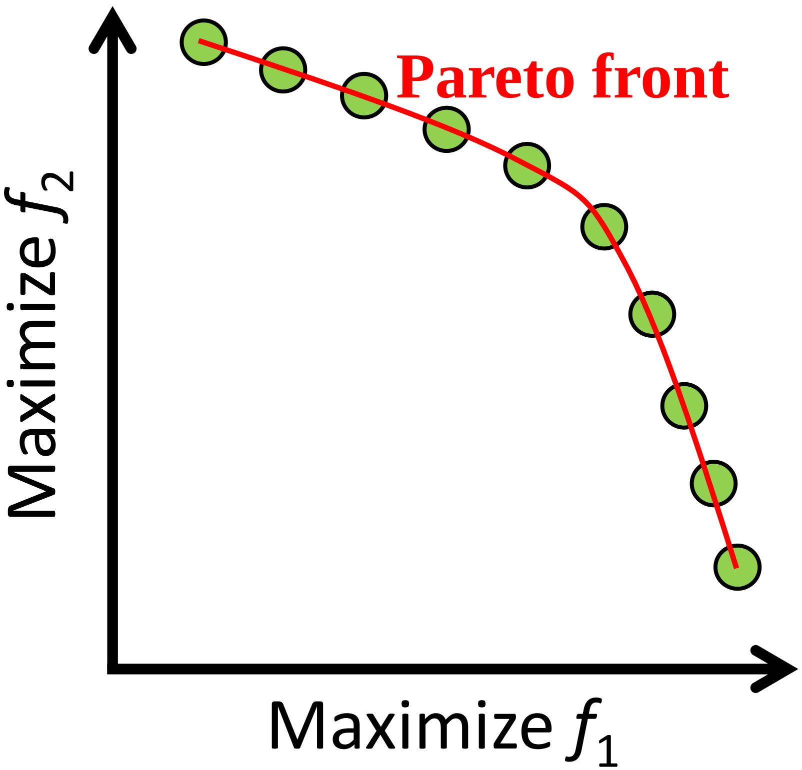
EMO Algorithm Design

Goal of Algorithm Design:

To find a set of well-distributed solutions over the entire Pareto front.



Initial Population

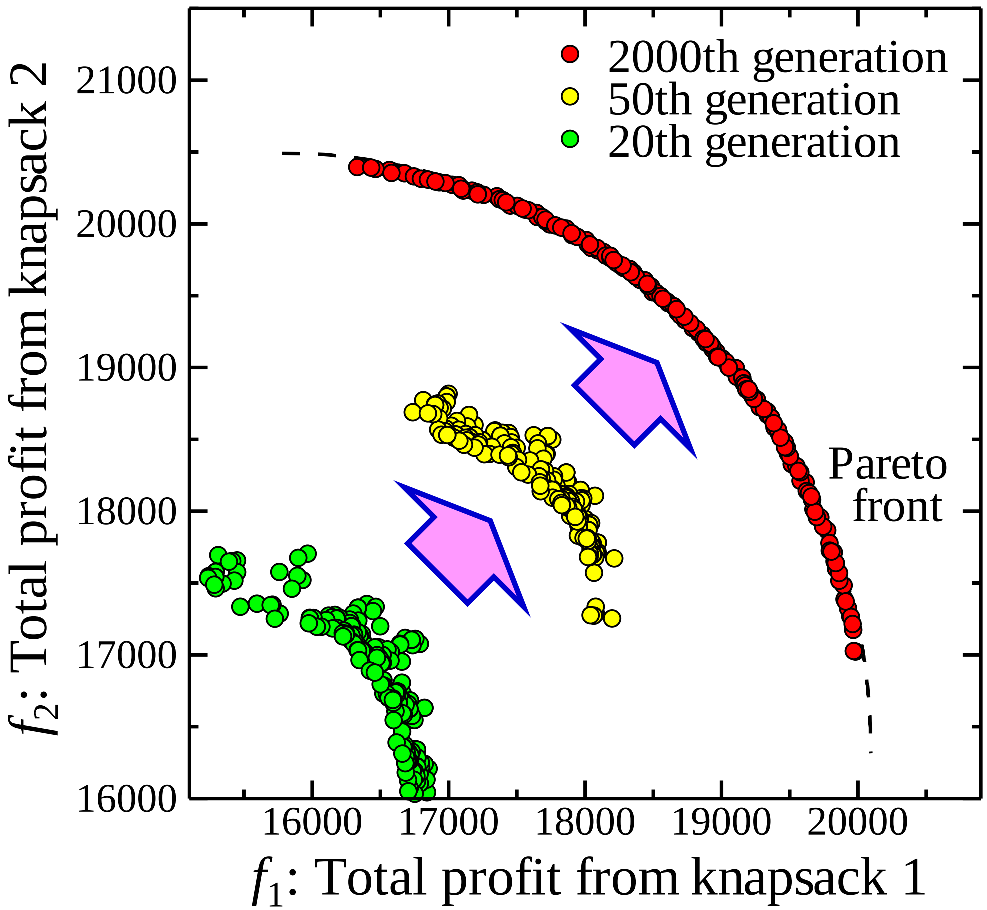


Final Population

EMO Algorithm Design

Goal of Algorithm Design:

To find a set of well-distributed solutions over the entire Pareto front.



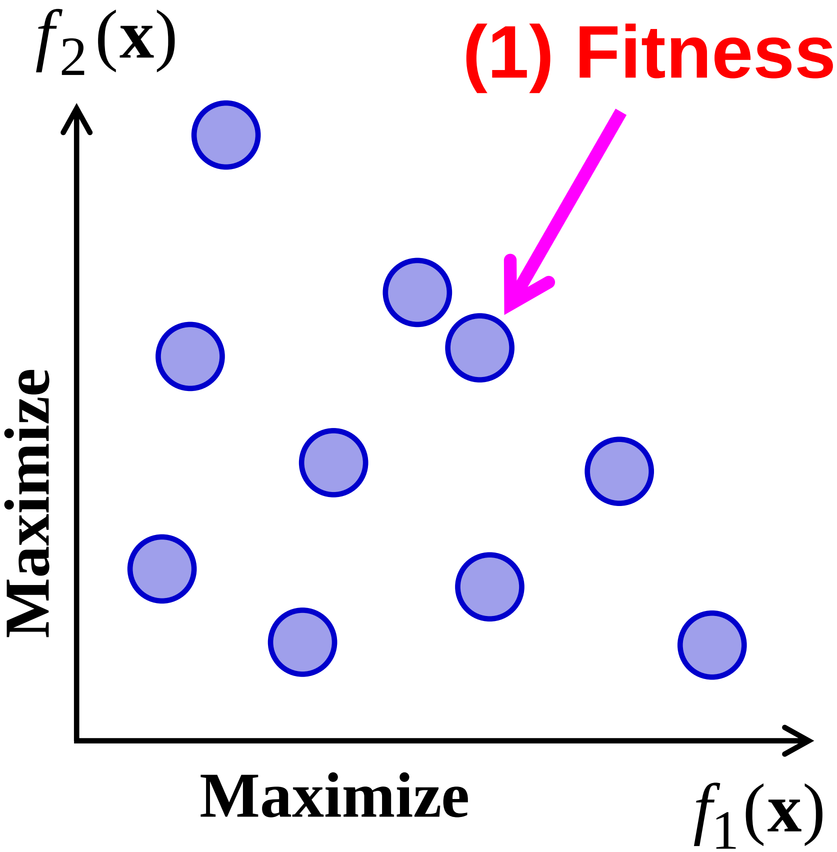
Two Levels of Evaluation

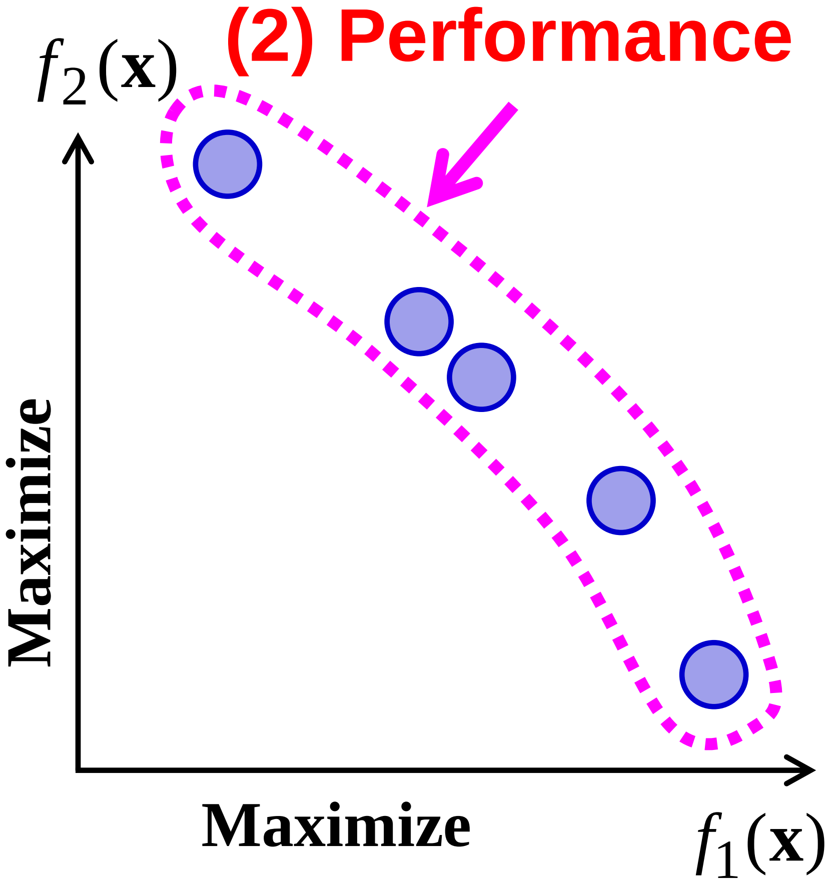
(1) Solution level:

for solution evaluation within an algorithm

(2) Solution set level:

for algorithm evaluation (solution set evaluation)

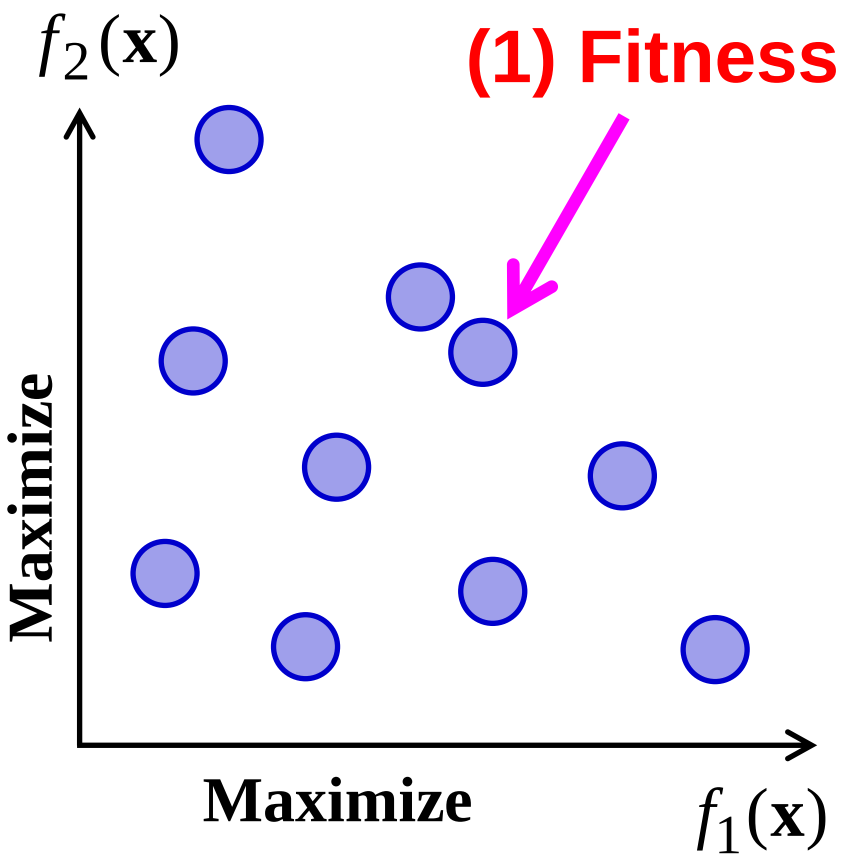


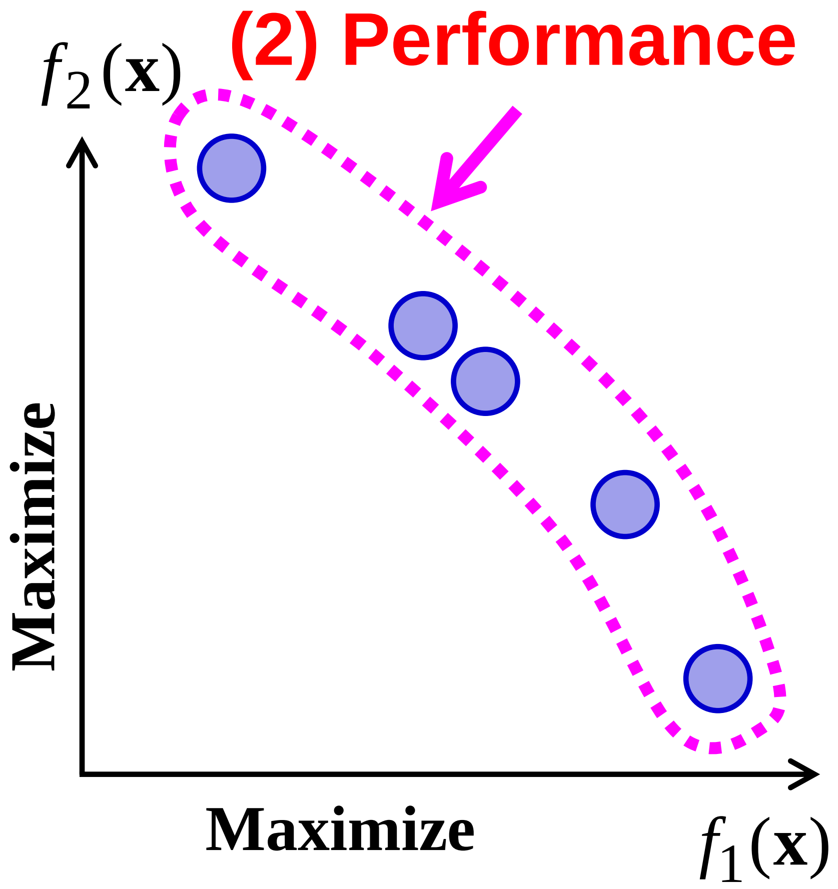


Two Levels of Evaluation

(1) Solution level: **Algorithm Design**
for solution evaluation within an algorithm

(2) Solution set level: **Algorithm Comparison**
for algorithm evaluation (solution set evaluation)

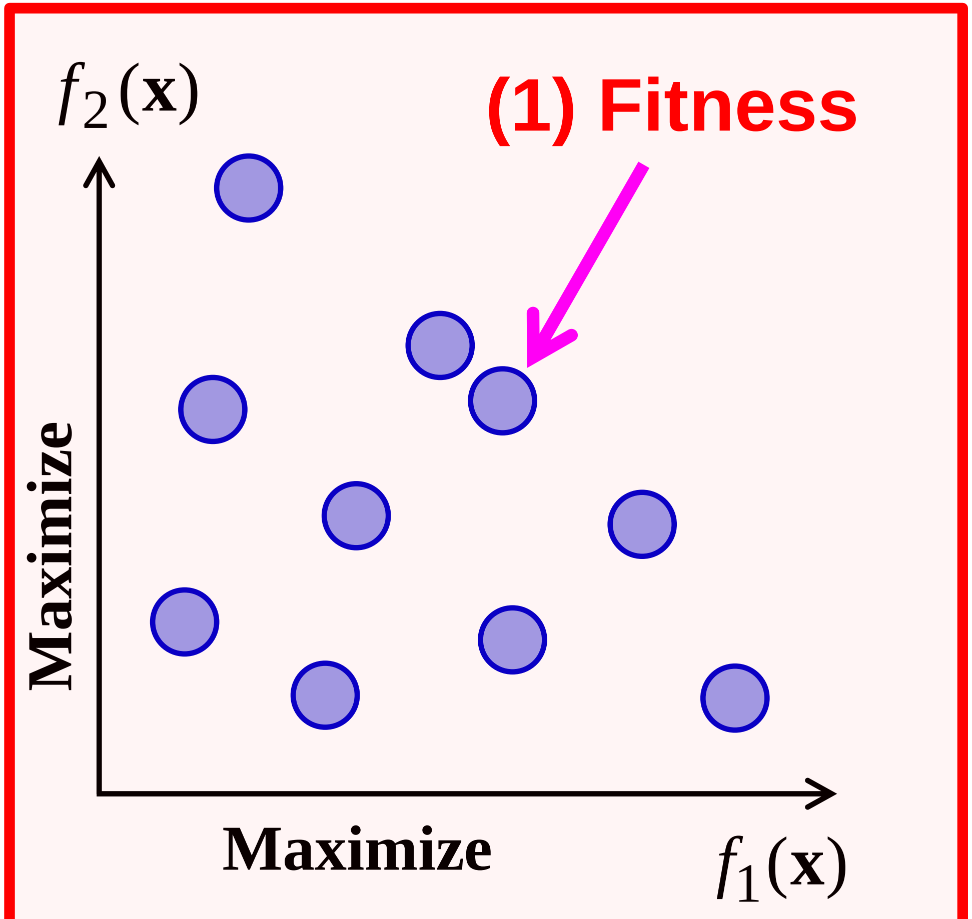


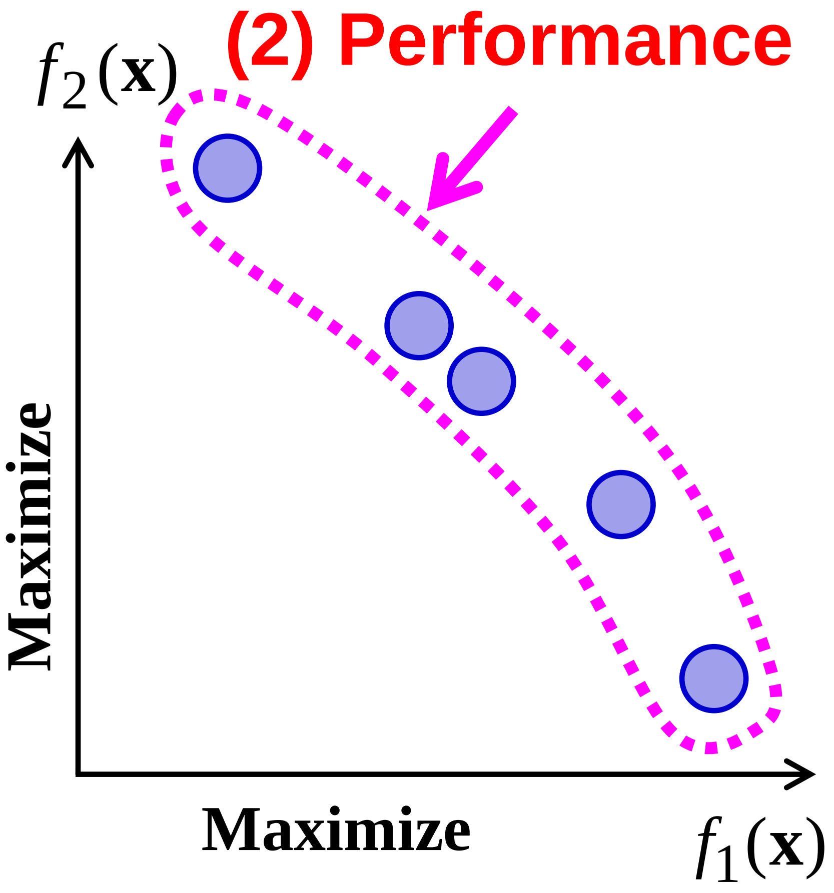


Two Levels of Evaluation

(1) Solution level: **Algorithm Design**
for solution evaluation within an algorithm

(2) Solution set level: **Algorithm Comparison**
for algorithm evaluation (solution set evaluation)





Your Task: To remove one solution (the worst solution) from the given six solutions to create the next population.

Which solution should be removed ?

Your Answer: ? .

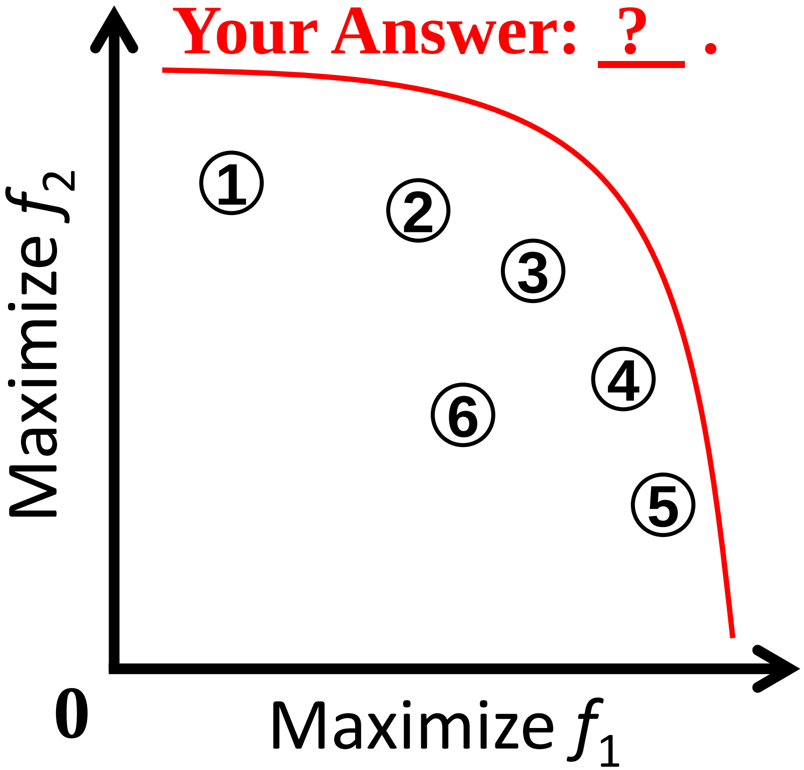
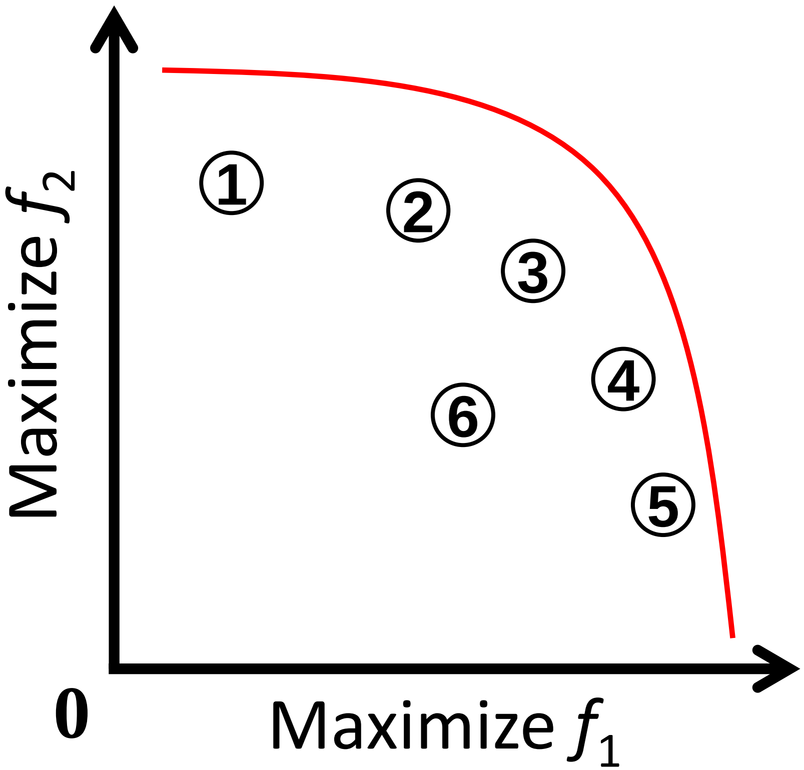


Fig. 5. Solution Set A.

(Maximization Problem)

Your Task: To remove one solution (the worst solution) from the given six solutions to create the next population.

Which solution should be removed ?



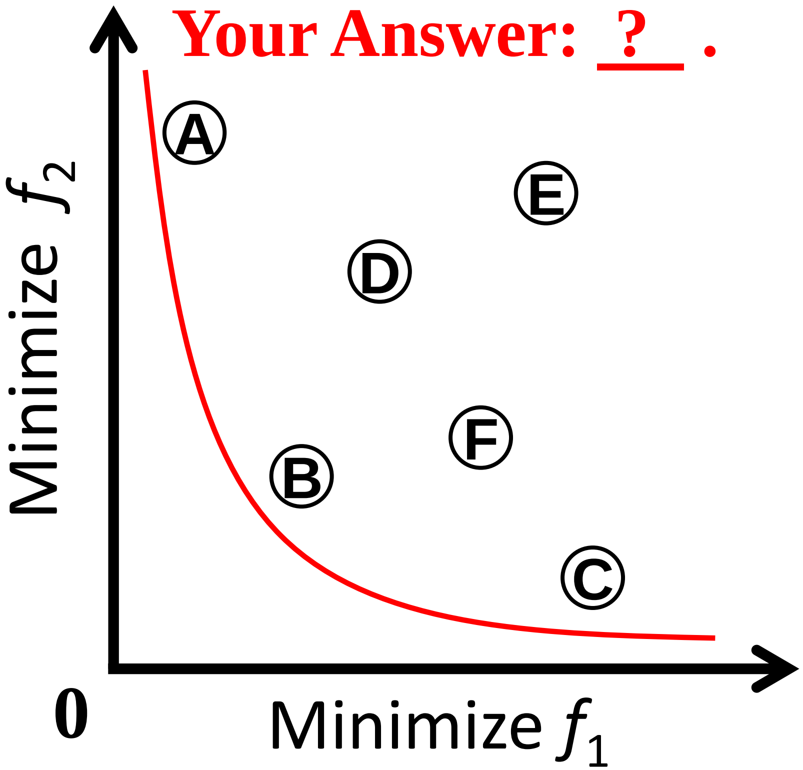


Fig. 5. Solution Set A.
(Maximization Problem)

Fig. 6. Solution Set B.
(Minimization Problem)

Your Task: To remove one solution (the worst solution) from the given six solutions to create the next population.

Which solution should be removed ?

Your Answer: __?__ .

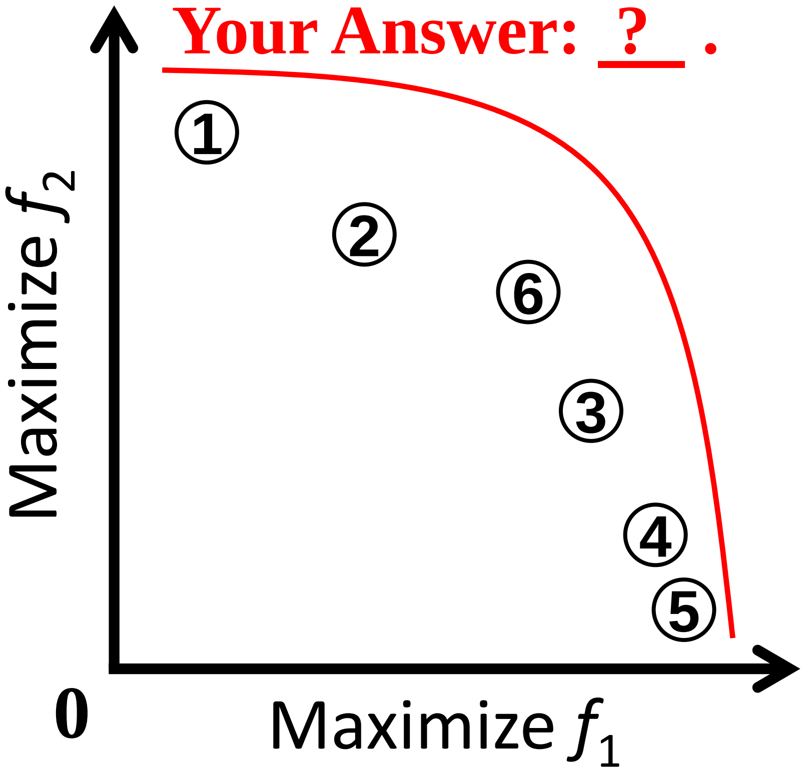


Fig. 7. Solution Set A.

(Maximization Problem)

Your Task: To remove one solution (the worst solution) from the given six solutions to create the next population.

Which solution should be removed ?

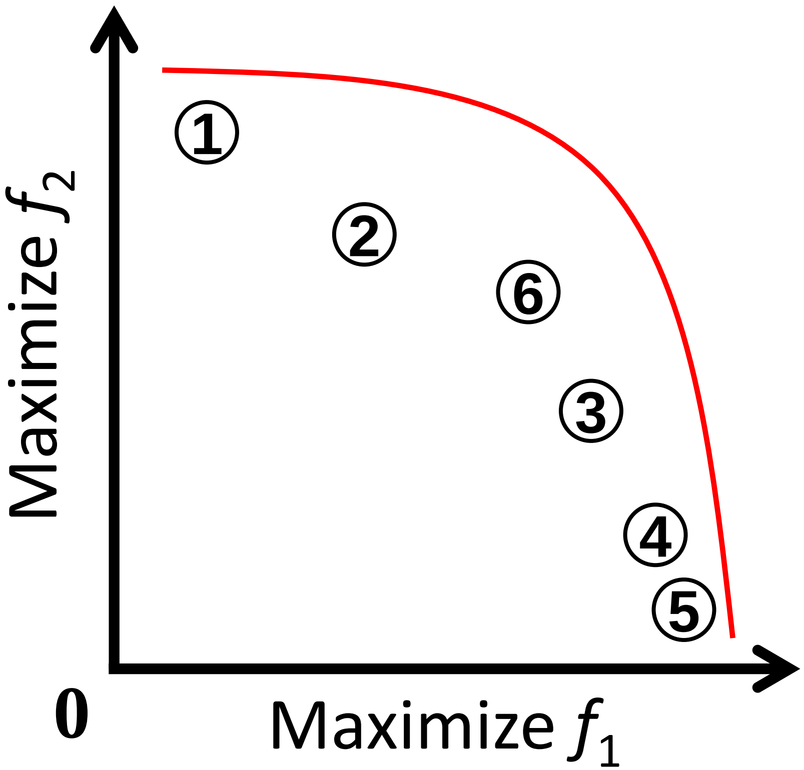


Fig. 7. Solution Set A.
(Maximization Problem)

Your Answer: __?__ .

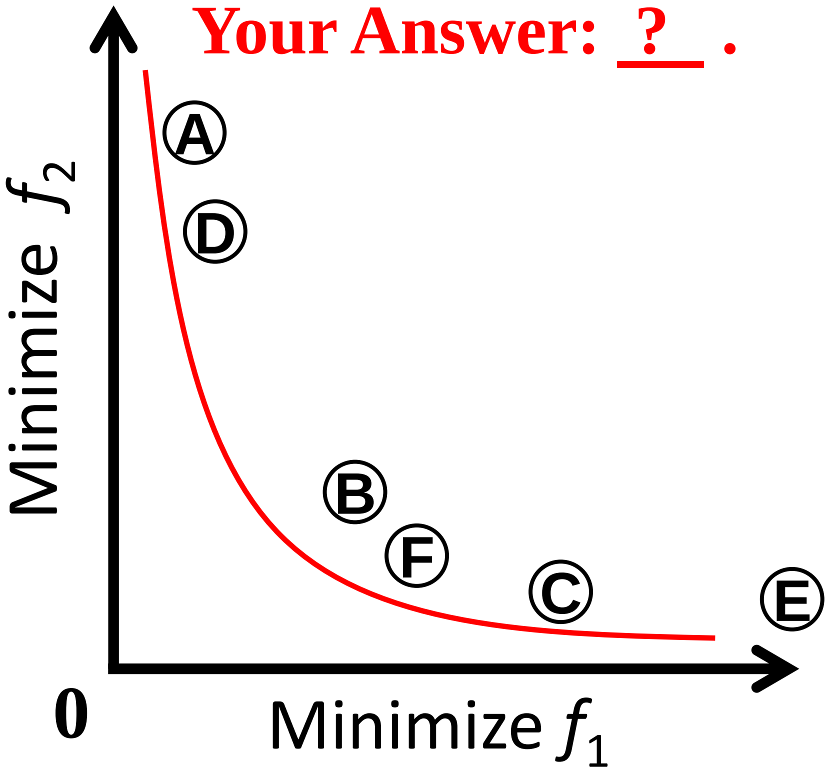
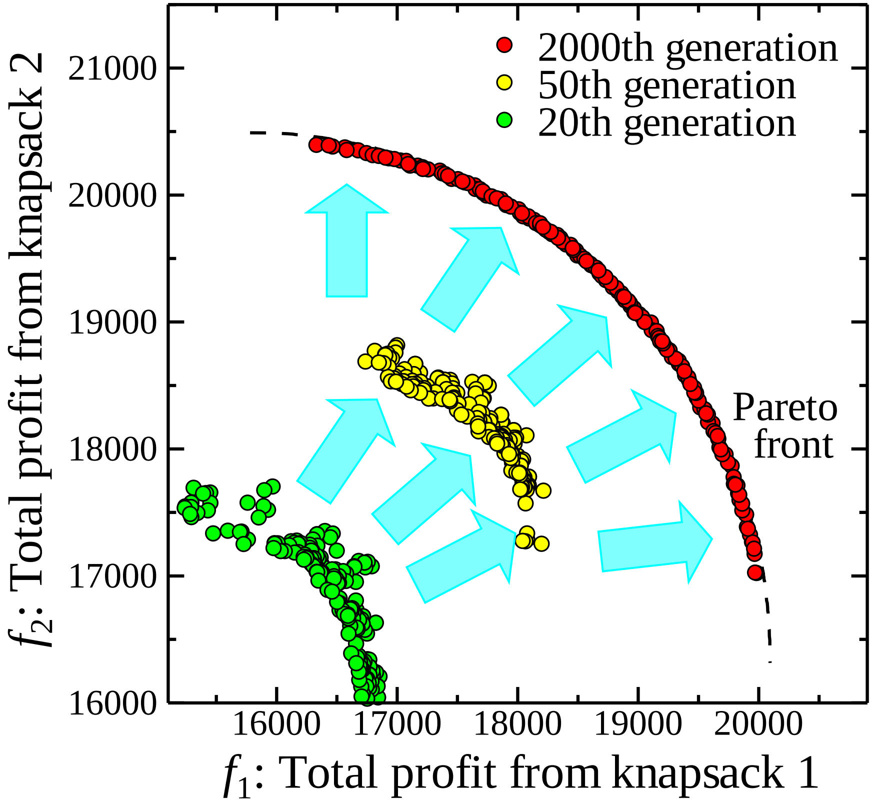


Fig. 8. Solution Set B.
(Minimization Problem)

Two Sub-goals:

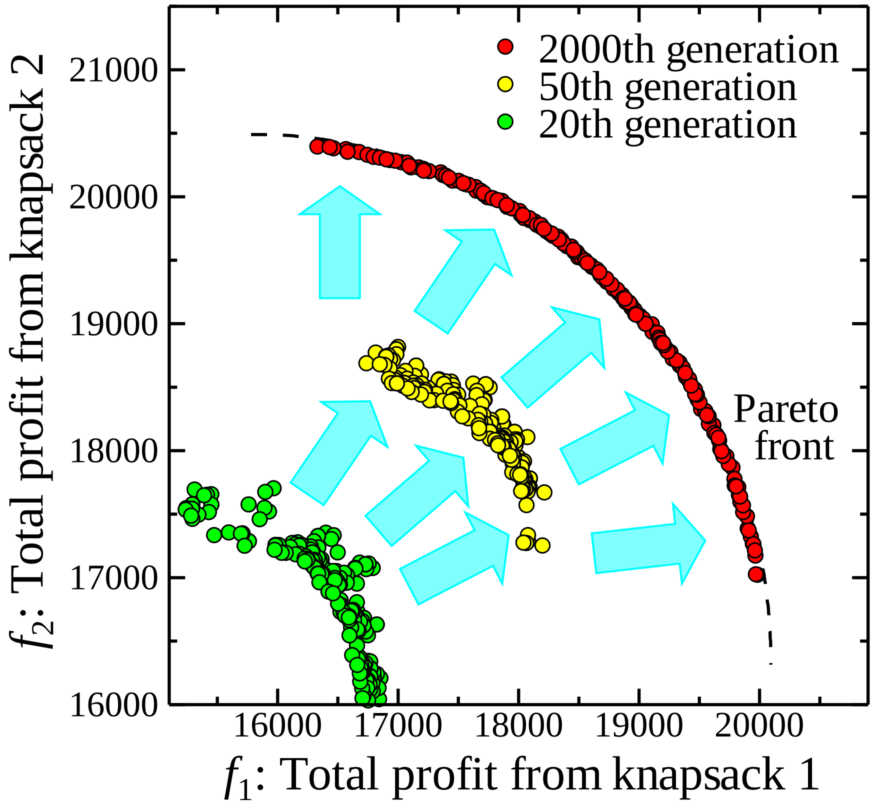
(i) Convergence: To push the population to the Pareto front as close as possible.

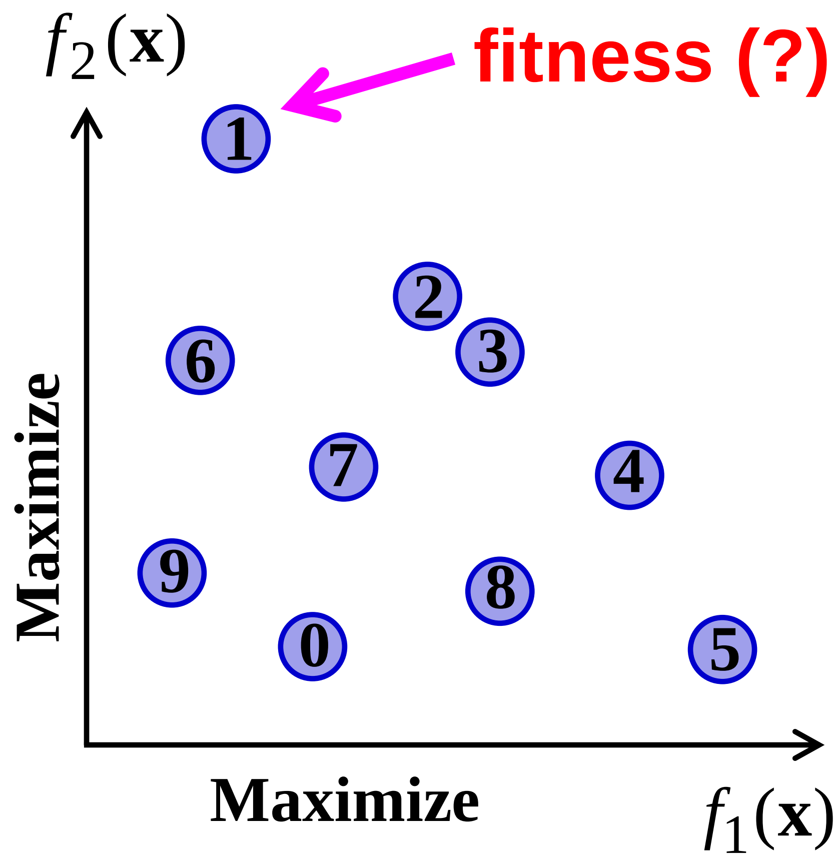
(ii) Diversity: To uniformly distribute the population over the entire Pareto front.



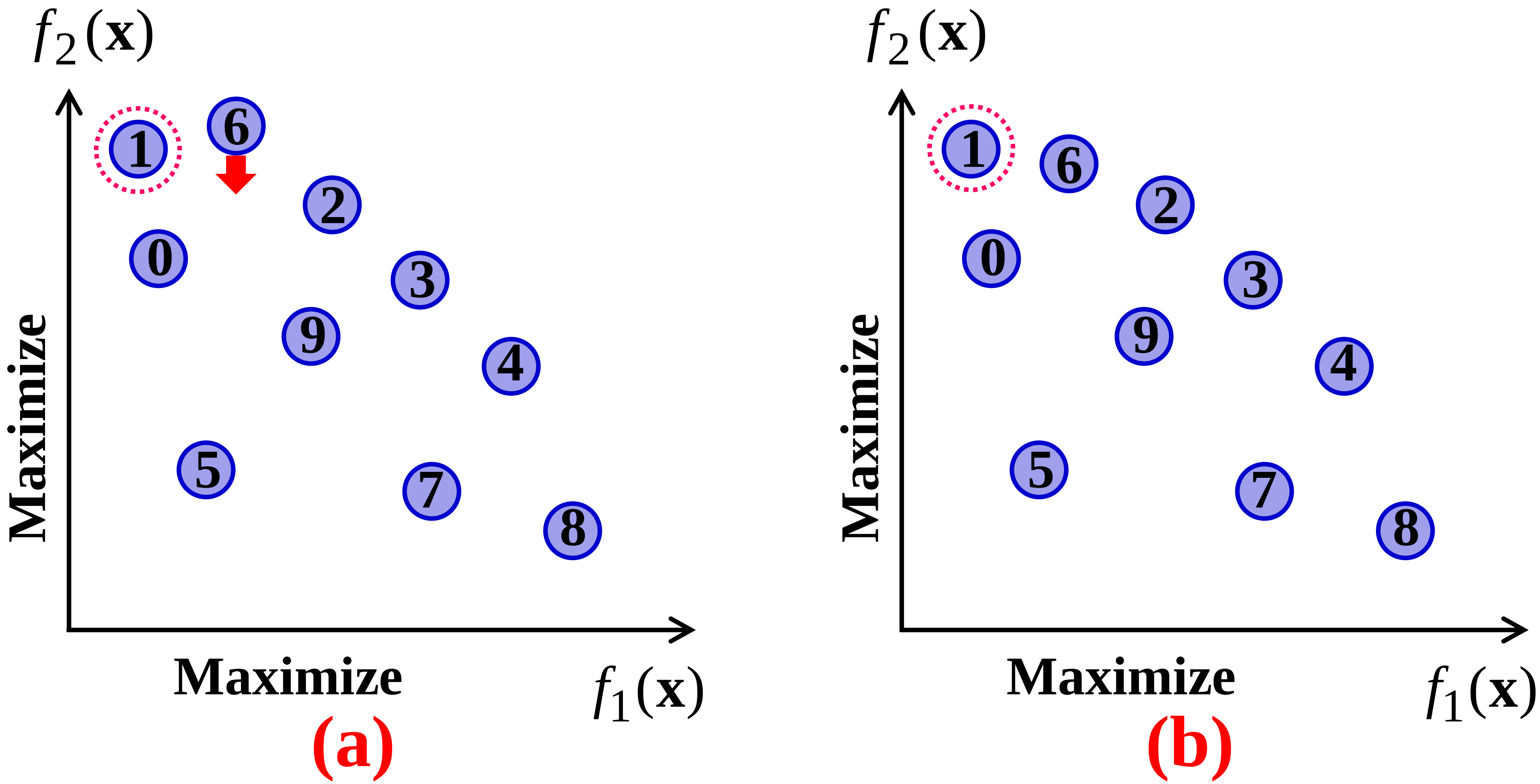
An important issue in the design of EMO algorithms:

How to evaluate a solution using multiple objectives
(Solution-level evaluation)

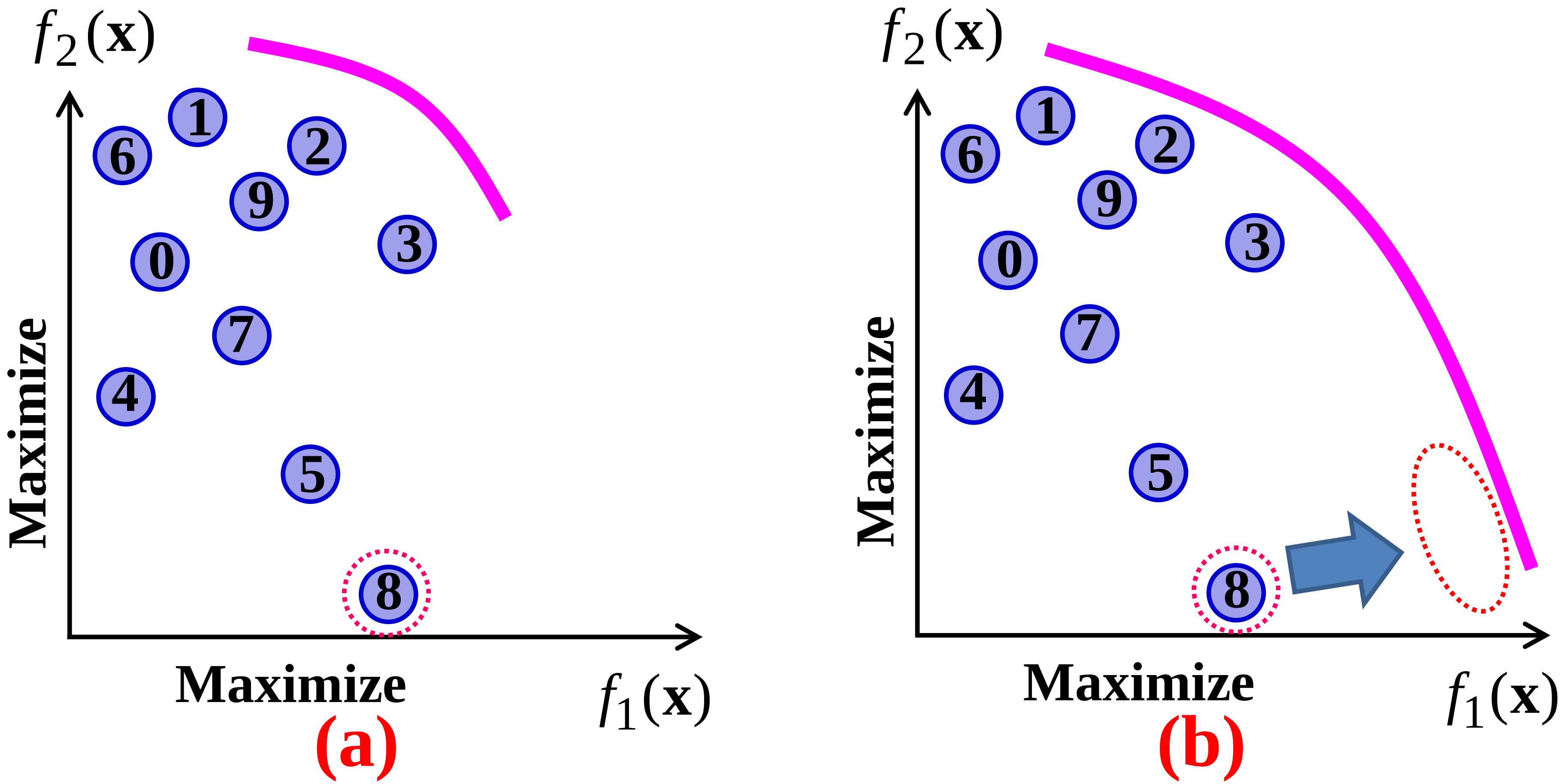




Q. Which solutions should have the highest fitness?
Q. Which solutions should have the lowest fitness?



The fitness of each solution totally depends on the other solutions. In (a), the fitness of Solution 1 is not high since it is a dominated solution. However, the fitness of the same solution in (b) is high since it is a non-dominated solution. The difference between (a) and (b) is a slight change of the location of Solution 6.



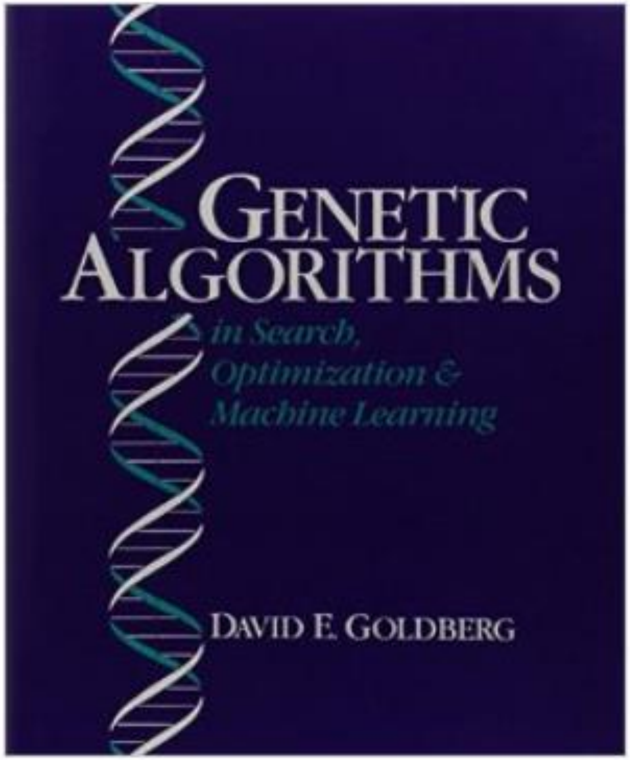
The importance of each solution totally depends on the Pareto front. In (a), Solution 8 is not important since it is far away from the Pareto front. However, in (b), the same solution may be very important in the search for the entire Pareto front. The difficulty is that we do not know the location of the true Pareto front.

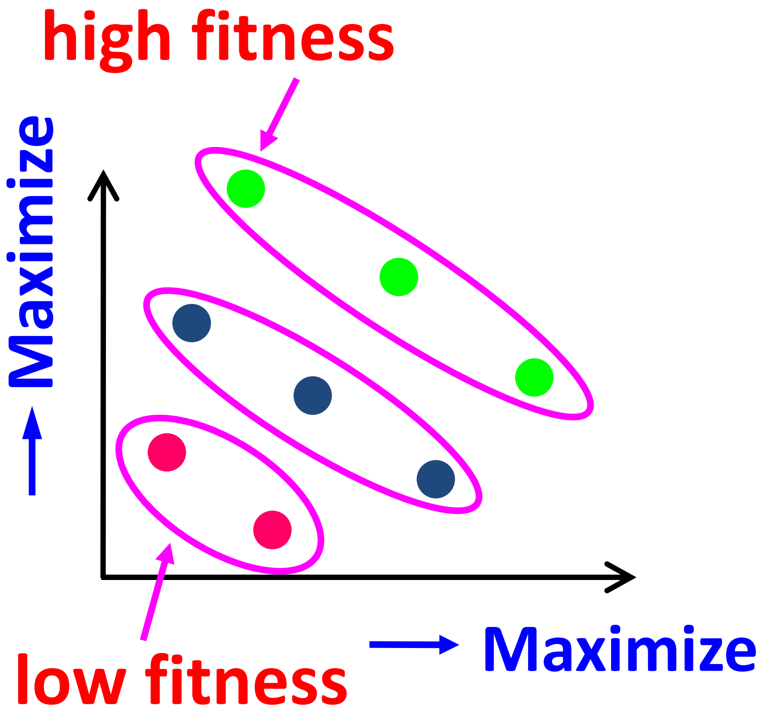
Basic Idea of EMO Algorithms

Pareto Dominance-based EMO Algorithms

(1) Pareto Dominance

This idea was suggested in Goldberg's GA book (1989).





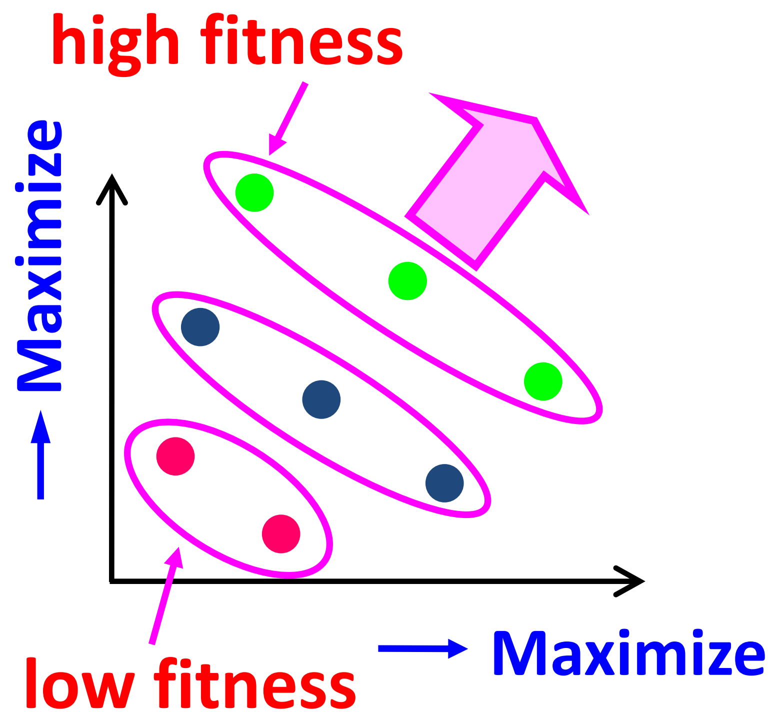
Basic Idea of EMO Algorithm

Pareto Dominance-based EMO Algorithms

(1) Pareto Dominance

This idea was suggested in Goldberg's GA book (1989).

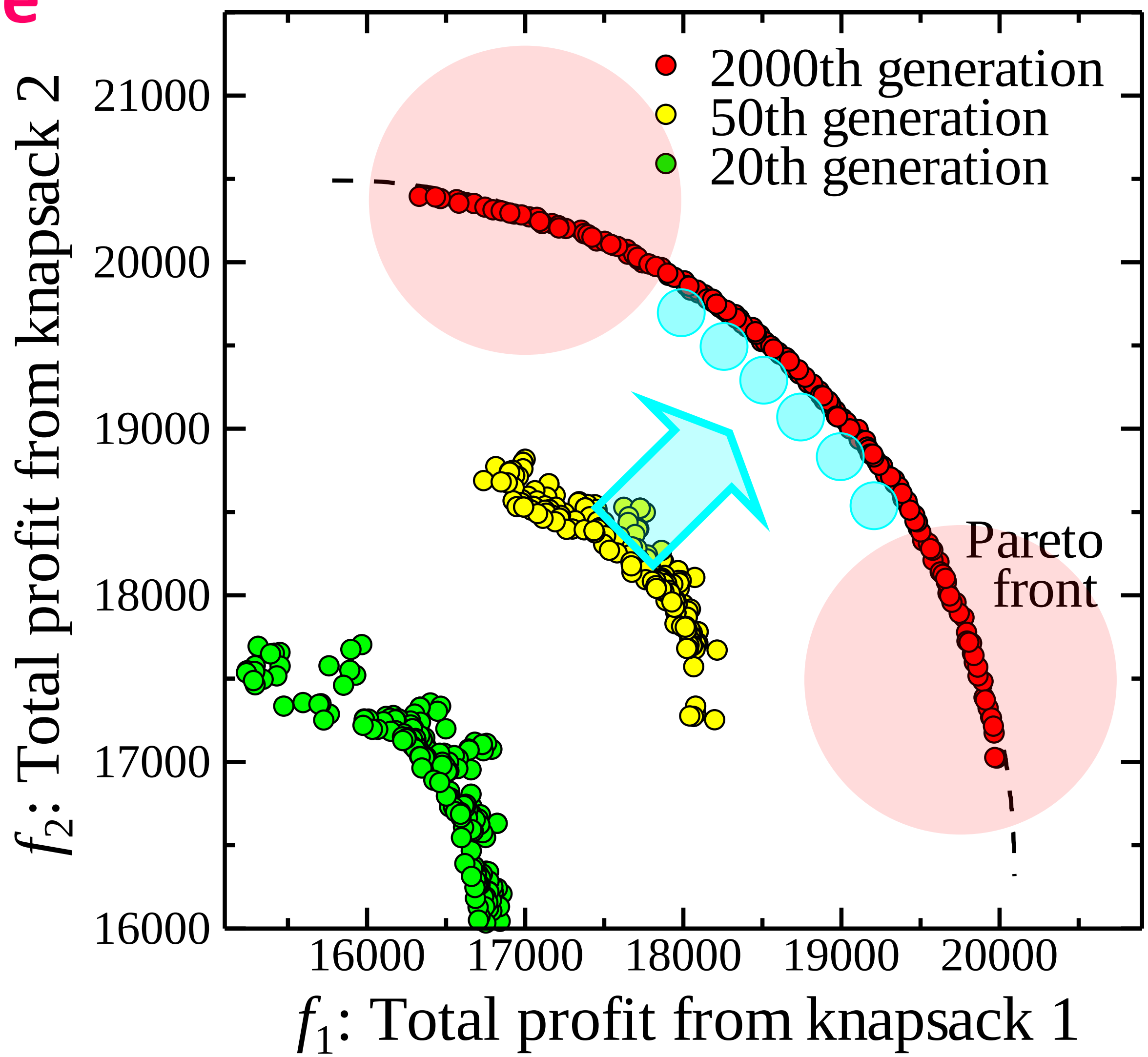
The point is to push the population towards the Pareto front.



Basic Idea of EMO Algorithm

Pareto Dominance-based EMO Algorithms

(1) Pareto Dominance

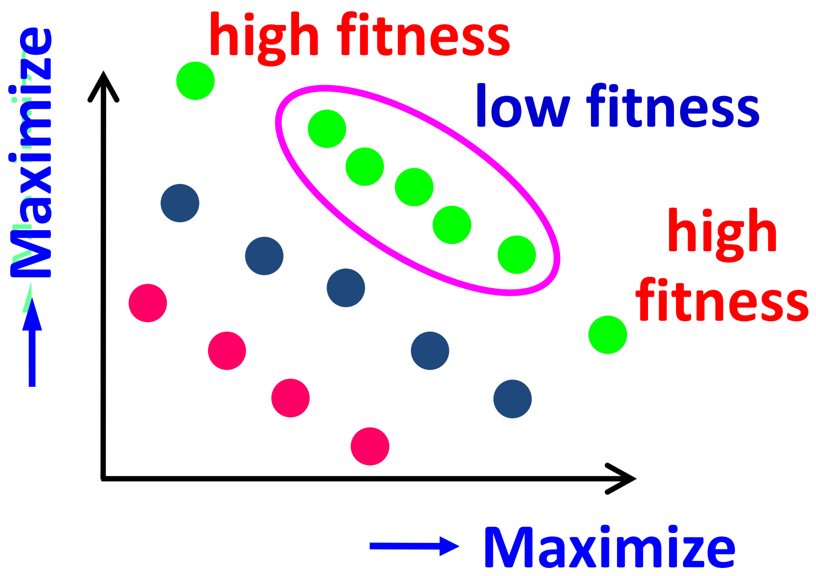


Basic Idea of EMO Algorithm

Pareto Dominance-based EMO Algorithms

(1) Pareto Dominance

This idea was suggested in Goldberg's GA book (1989).



(2) Crowding

This idea has been used since the early 1990s.

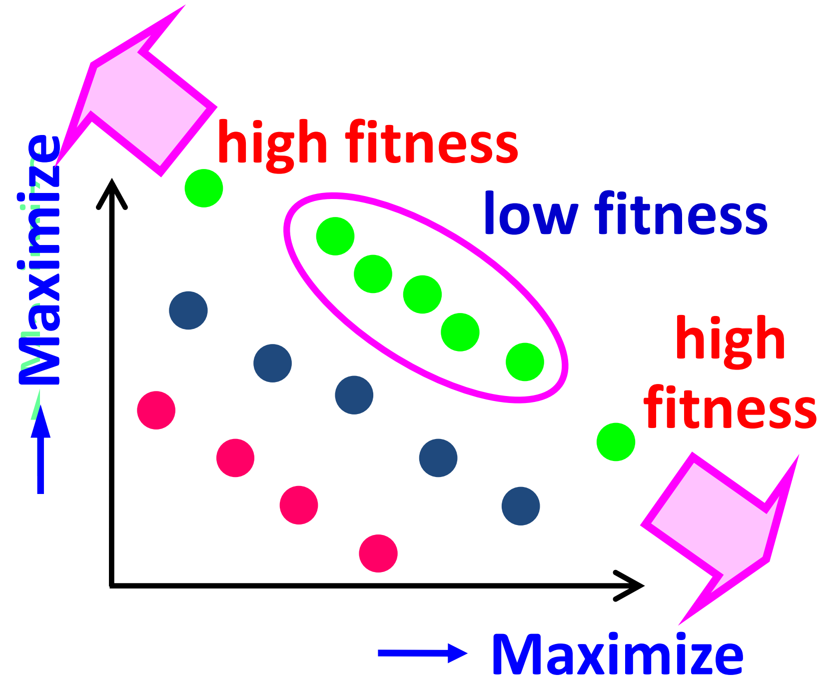
Example: The number of solutions in the neighborhood of each solution (the smaller number, the better).

Basic Idea of EMO Algorithm

Pareto Dominance-based EMO Algorithms

(1) Pareto Dominance

This idea was suggested in Goldberg's GA book (1989).



(2) Crowding

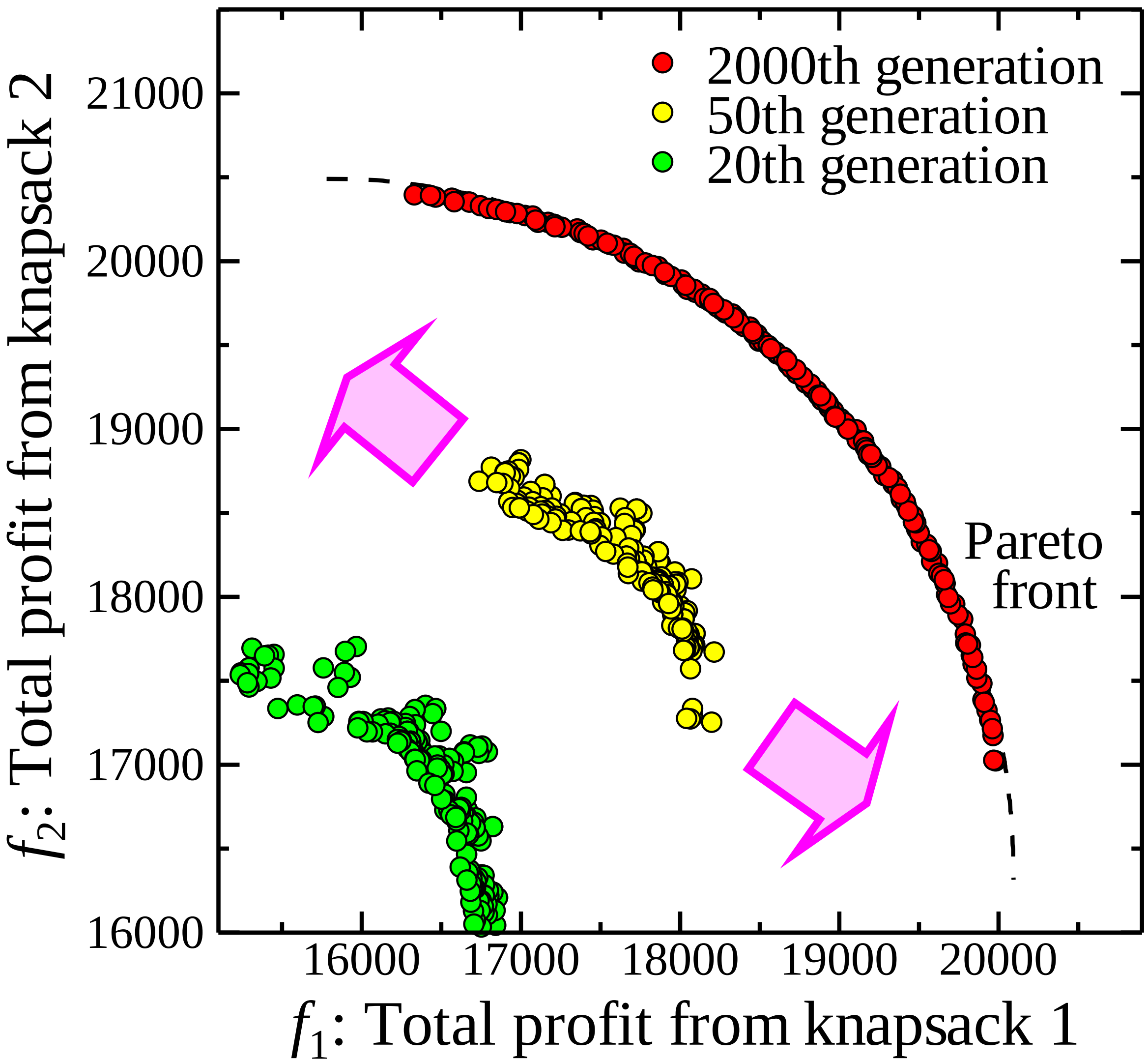
This idea has been used since the early 1990s.

The point is to maintain (or increase) the diversity of solutions in the population.

Basic Idea of EMO Algorithm

Pareto Dominance-based EMO Algorithms

(2) Crowding

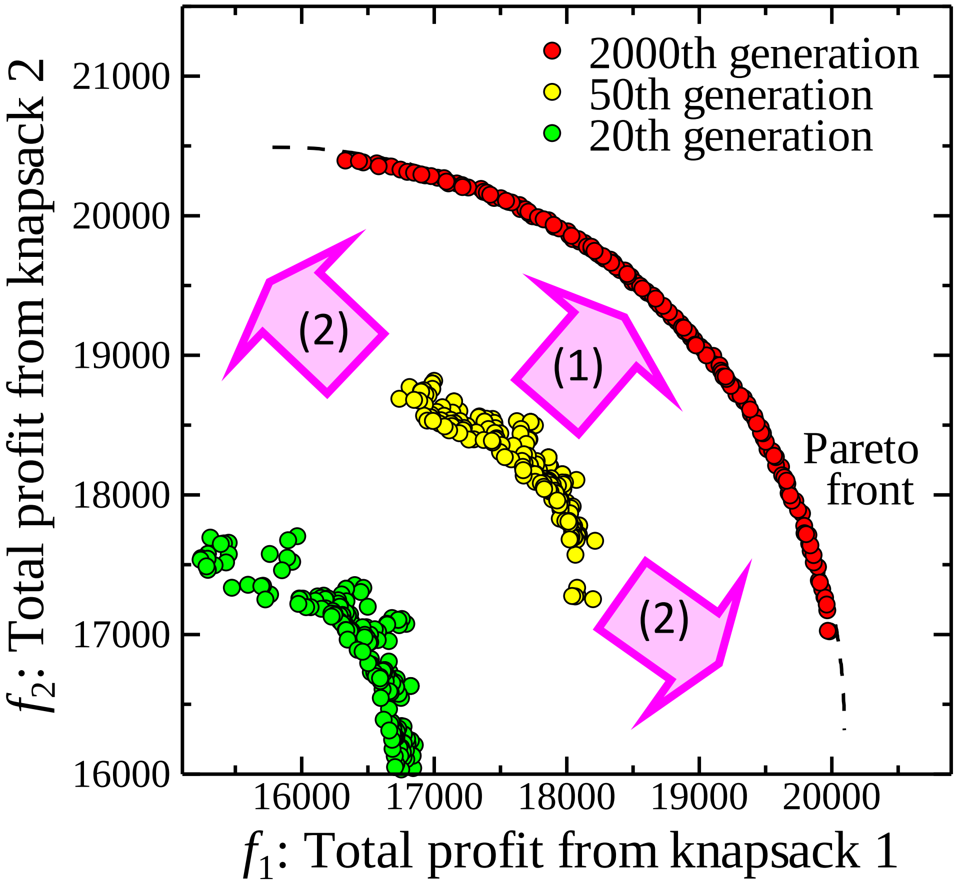


Basic Idea of EMO Algorithm

Pareto Dominance-based EMO Algorithms

(1) Pareto Dominance (Convergence to the Pareto front)

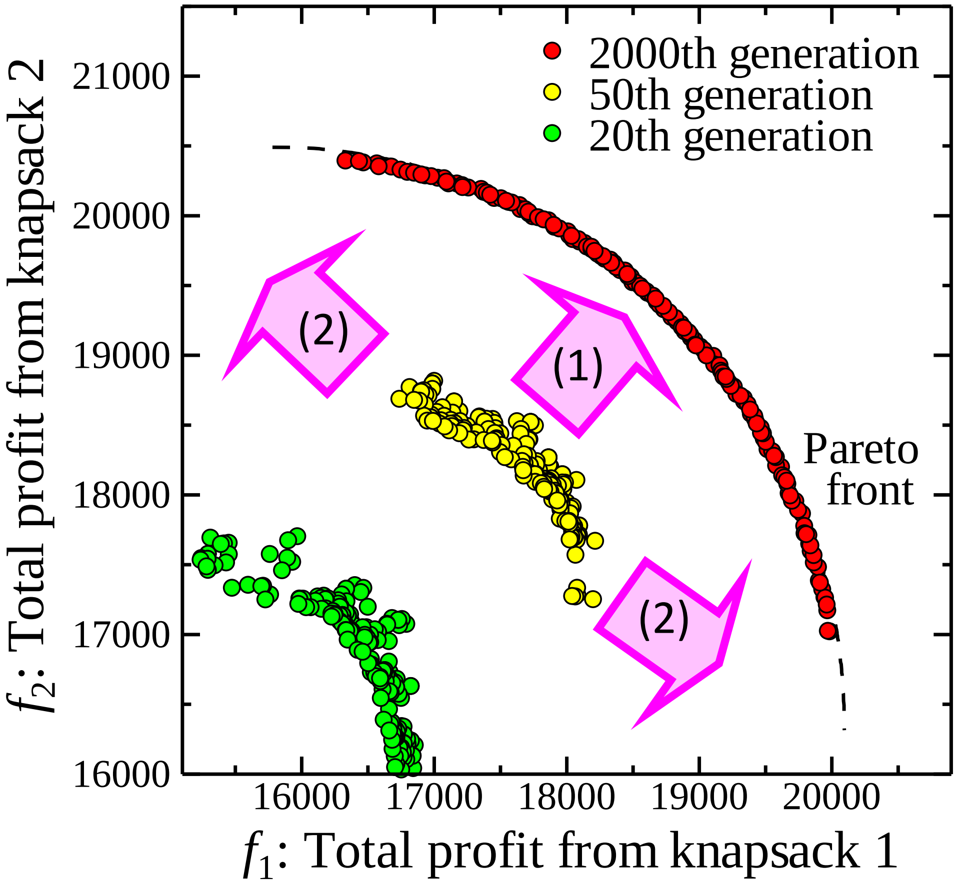
(2) Crowding (Diversity Maintenance)



Basic Idea of EMO Algorithm

Pareto Dominance-based EMO Algorithms

(1) Pareto Dominance (Convergence to the Pareto front)
(2) Crowding (Diversity Maintenance)



The most important issue in EMO algorithm design:
How to maintain a good balance between
(1) the convergence, and
(2) the diversity.

Highly-Cited Papers published in Mid-1990s

Non-Elitist Pareto Dominance-Based Algorithms

[1] C. M. Fonseca and P. J. Fleming, “Genetic Algorithms for Multiobjective Optimization: Formulation, Discussion and Generalization,” 5th ICGA, pp. 416-423, 1993. MOGA

[2] J. Horn, N. Nafpliotis, and D. E. Goldberg, “A Niched Pareto Genetic Algorithm for Multiobjective Optimization”, 1st IEEE ICEC (in WCCI 1994), pp. 82–87, 1994. NPGA

[3] N. Srinivas and K. Deb, “Multiobjective Optimization Using Nondominated Sorting in Genetic Algorithms”, Evolutionary Computation, Vol. 2, No. 3, pp. 221-248 (1995). NSGA

Highly-Cited Papers published in Mid-1990s

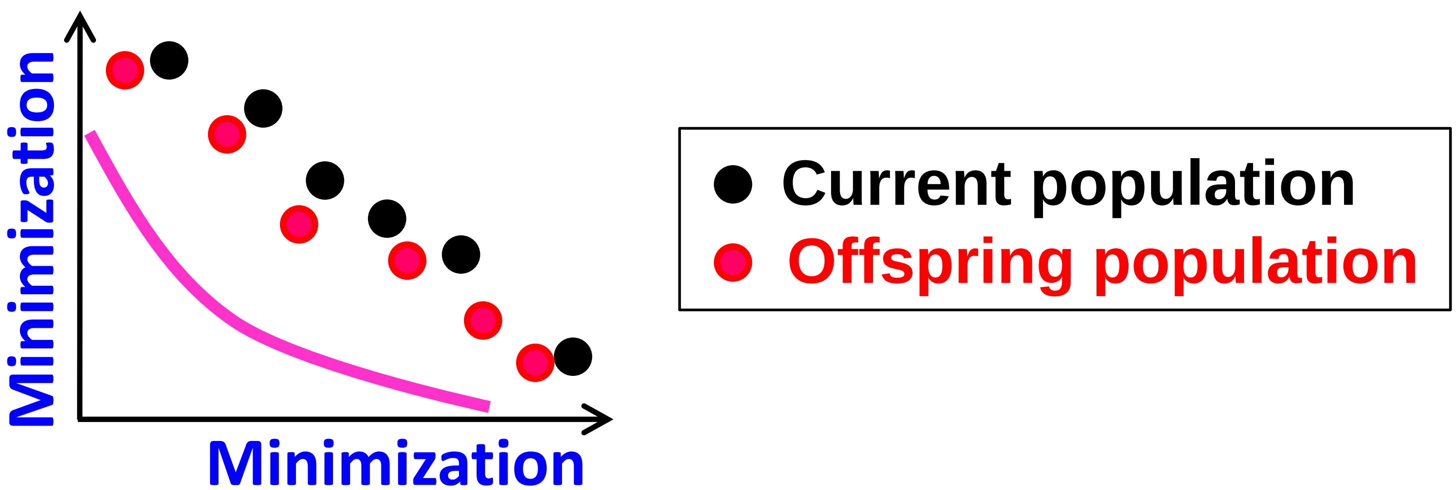
Main Feature: Weak Convergence Ability

(1) Non-Elitist

(2) Pareto dominance-based fitness evaluation

(3) Diversification mechanism

The point: To generate a good offspring population



Highly-Cited Papers published in Mid-1990s

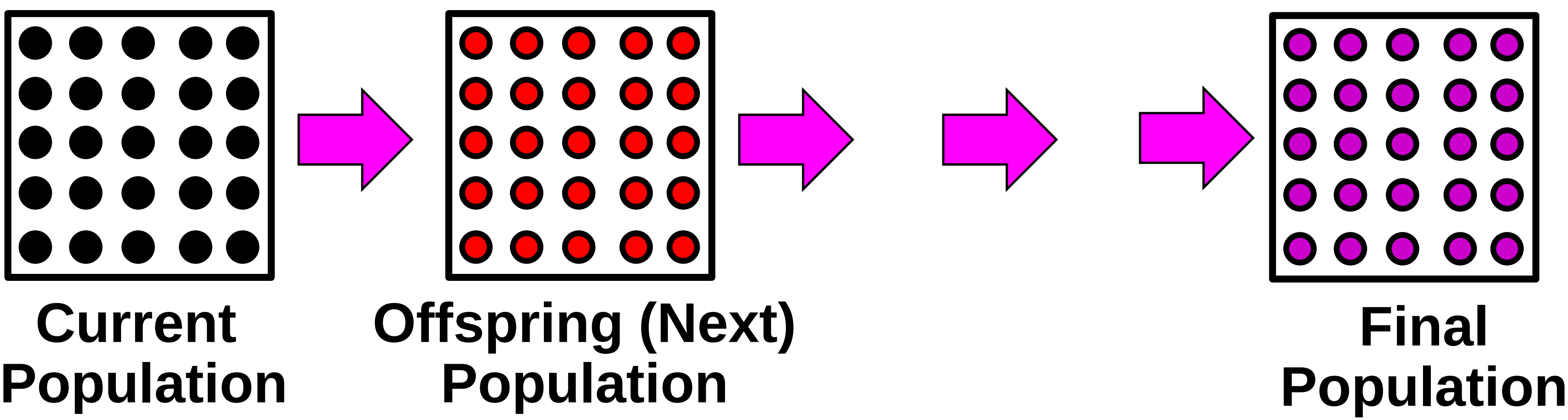
Main Feature: Weak Convergence Ability

(1) Non-Elitist

(2) Pareto dominance-based fitness evaluation

(3) Diversification mechanism

The point: To generate a good offspring population



Highly-Cited Papers published in Mid-1990s

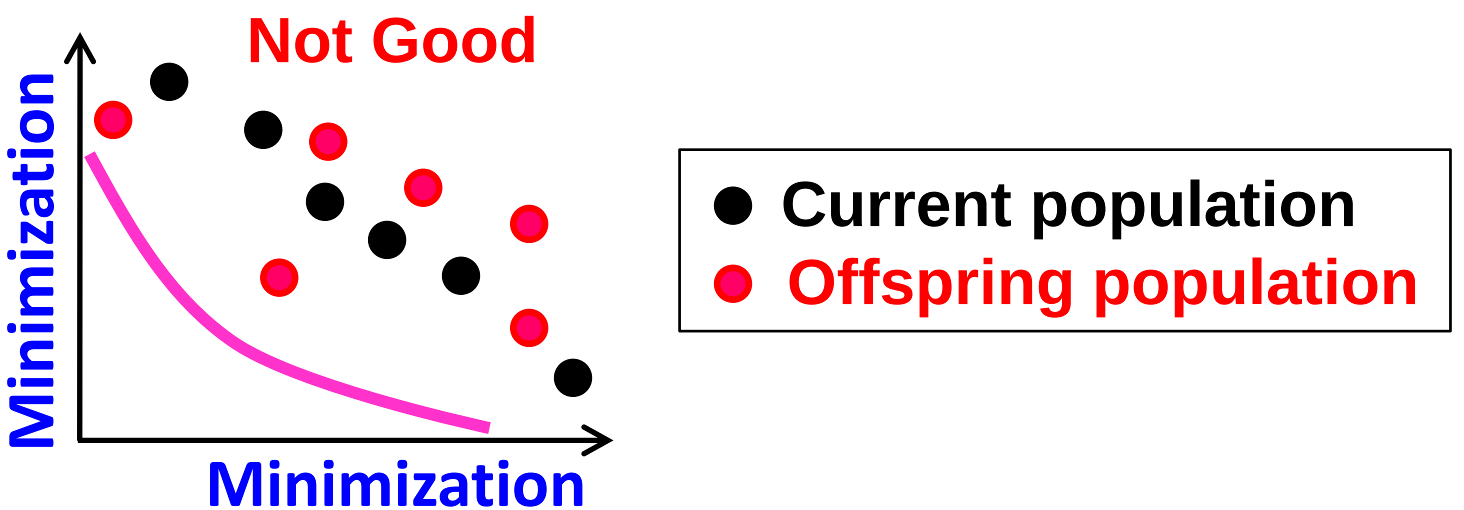
Main Feature: Weak Convergence Ability

(1) Non-Elitist

(2) Pareto dominance-based fitness evaluation

(3) Diversification mechanism

The point: To generate a good offspring population



Highly-Cited Papers published in Mid-1990s

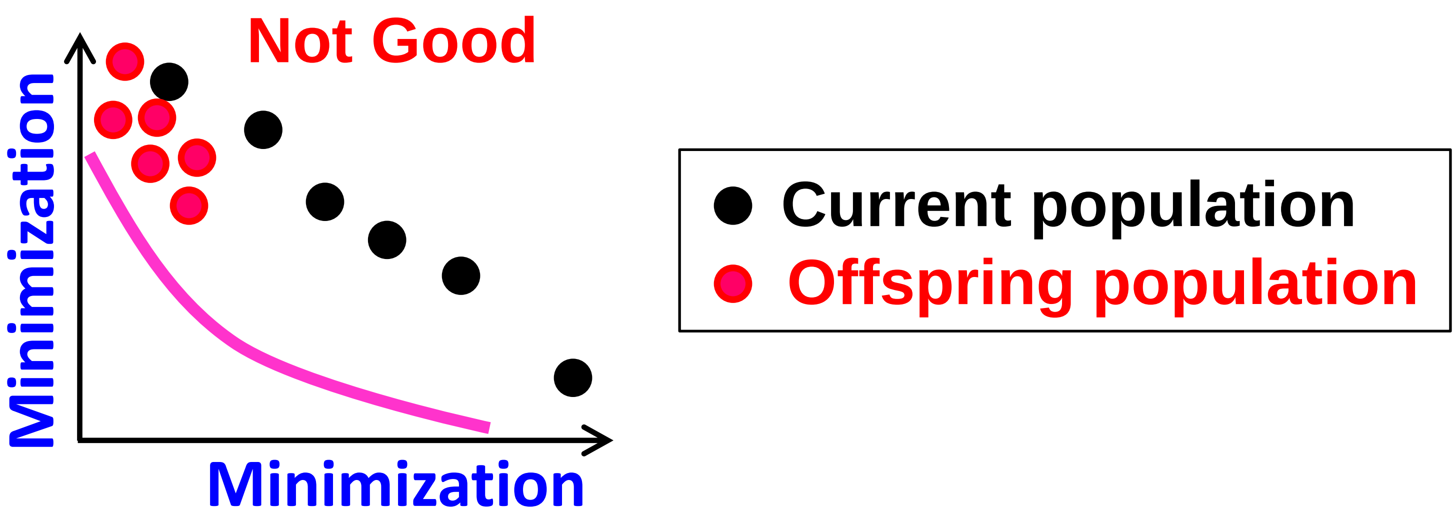
Main Feature: Weak Convergence Ability

(1) Non-Elitist

(2) Pareto dominance-based fitness evaluation

(3) Diversification mechanism

The point: To generate a good offspring population



Highly-Cited Papers published in Mid-1990s

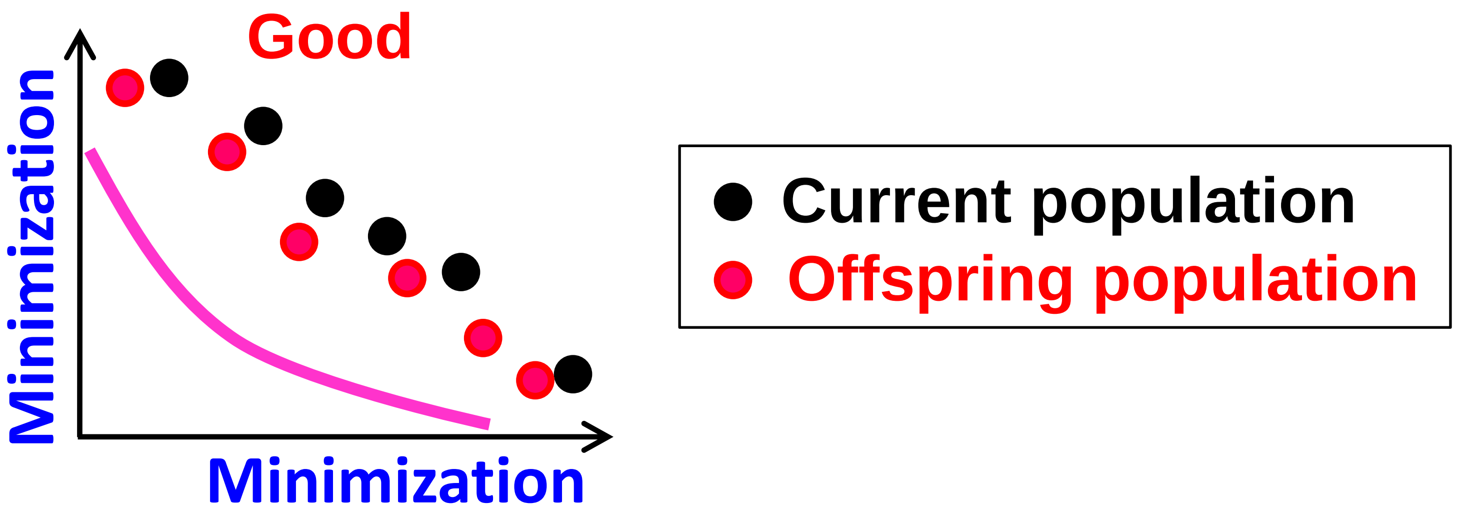
Main Feature: Weak Convergence Ability

(1) Non-Elitist

(2) Pareto dominance-based fitness evaluation

(3) Diversification mechanism

The point: To generate a good offspring population



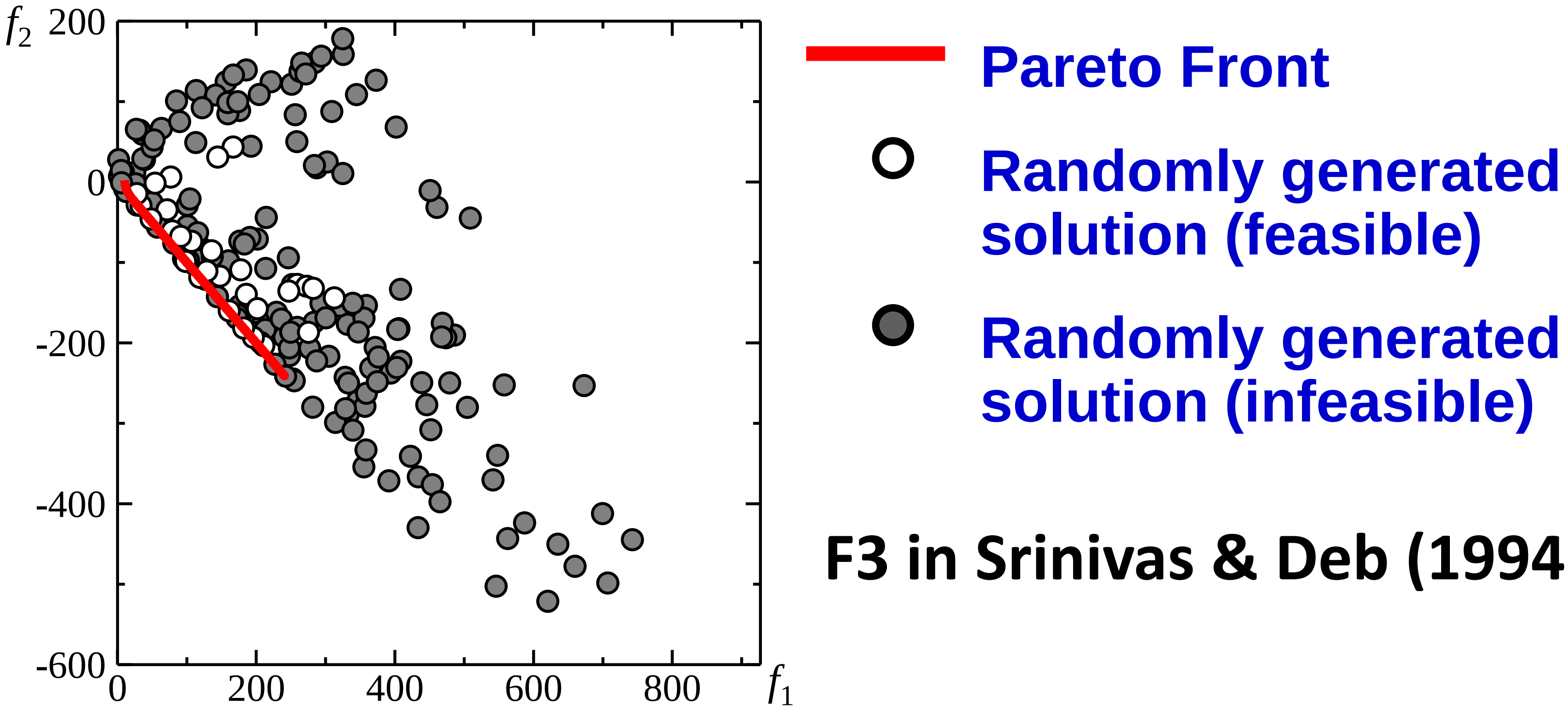
Why Non-Elitist EMO Algorithms ?

Answer:

Strong convergence ability is not needed.

Why Non-Elitist EMO Algorithms ?

Strong convergence ability is not needed.



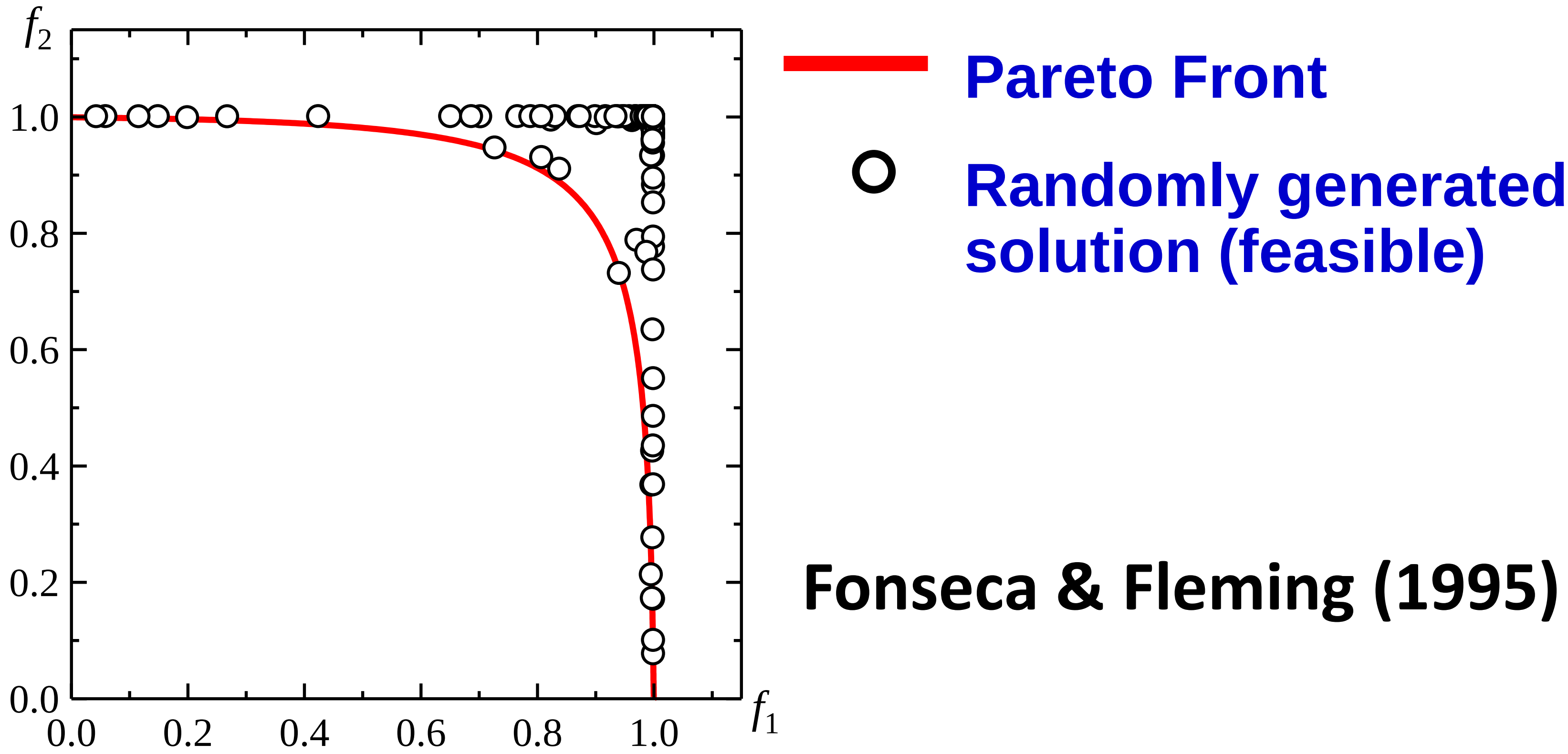
F3 in Srinivas & Deb (1994)

Pareto Front and Randomly Generated 200 Initial Solutions

Some randomly generated solutions are on the Pareto front.

Why Non-Elitist EMO Algorithms ?

Strong convergence ability is not needed.

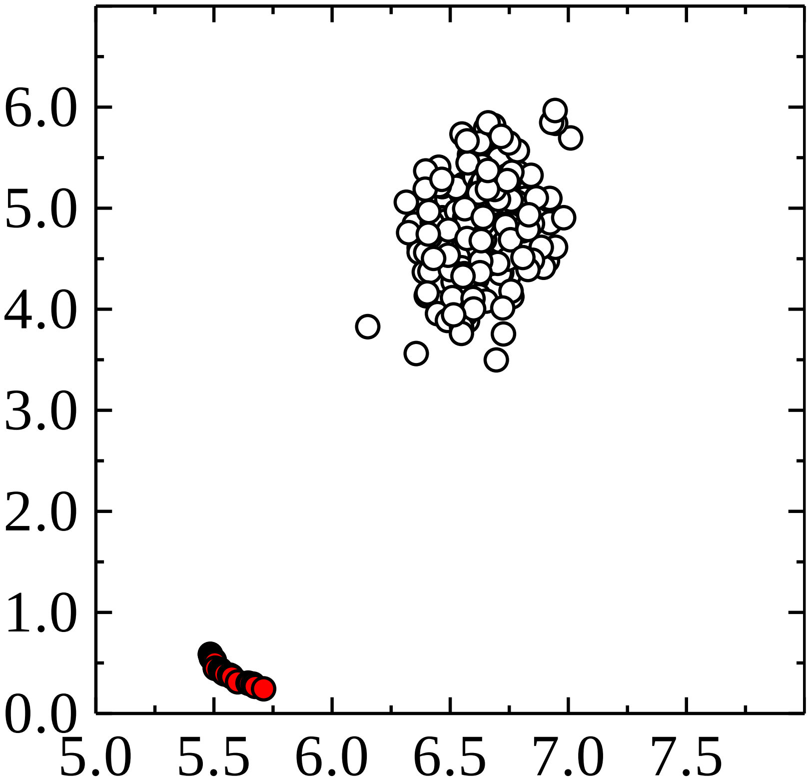


Fonseca & Fleming (1995)

Pareto Front and Randomly Generated 200 Initial Solutions

Some randomly generated solutions are on the Pareto front.

When we have different problems, different algorithms are developed.



Flowshop Scheduling Problem with 80 Jobs

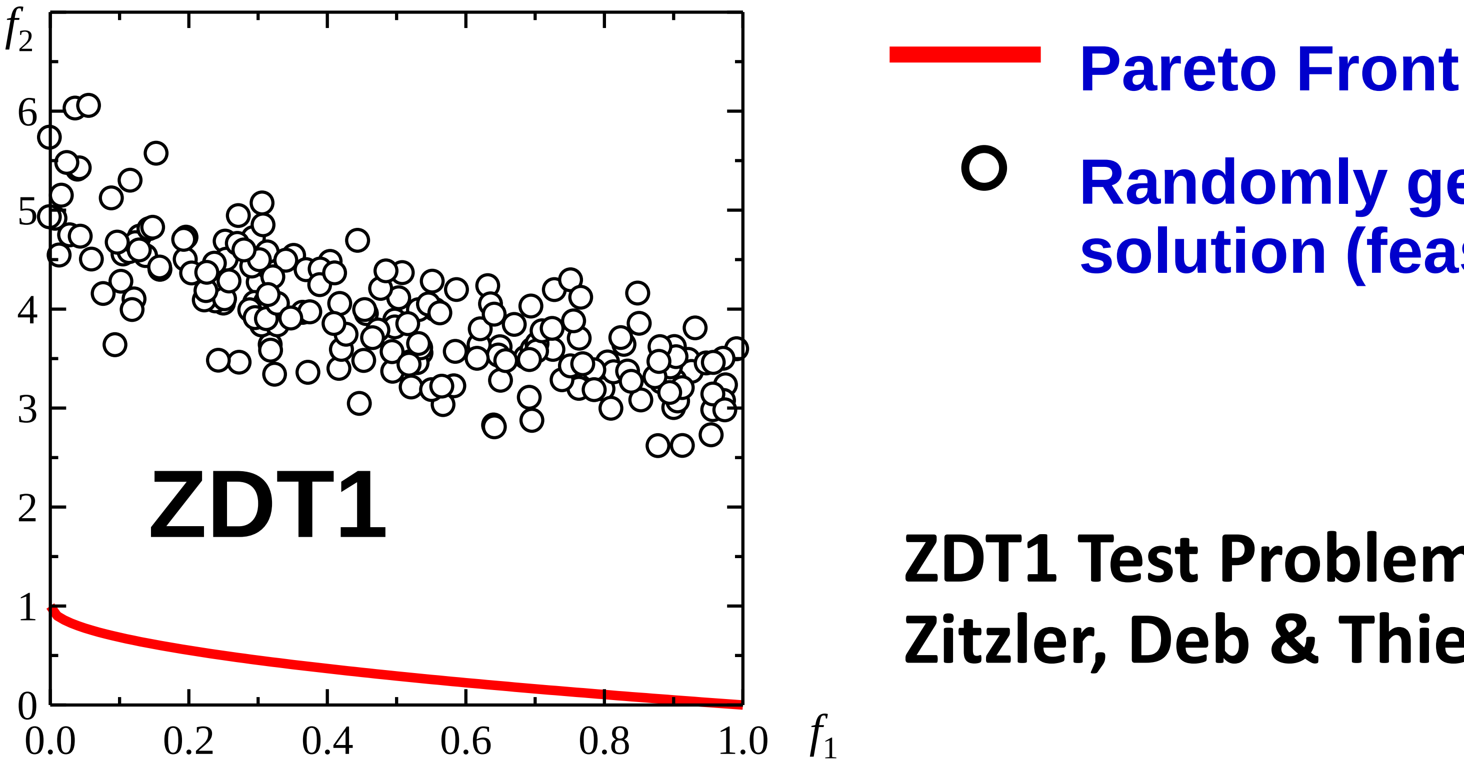
- Randomly generated initial solution
- Final solution

Elitist Hybrid EMO Algorithm with Local Search

Ishibuchi & Murata (CEC 1996, T-SMC Part C 1998)

This is not in the main stream of the EMO research.

ZDT Test Problems (2000)

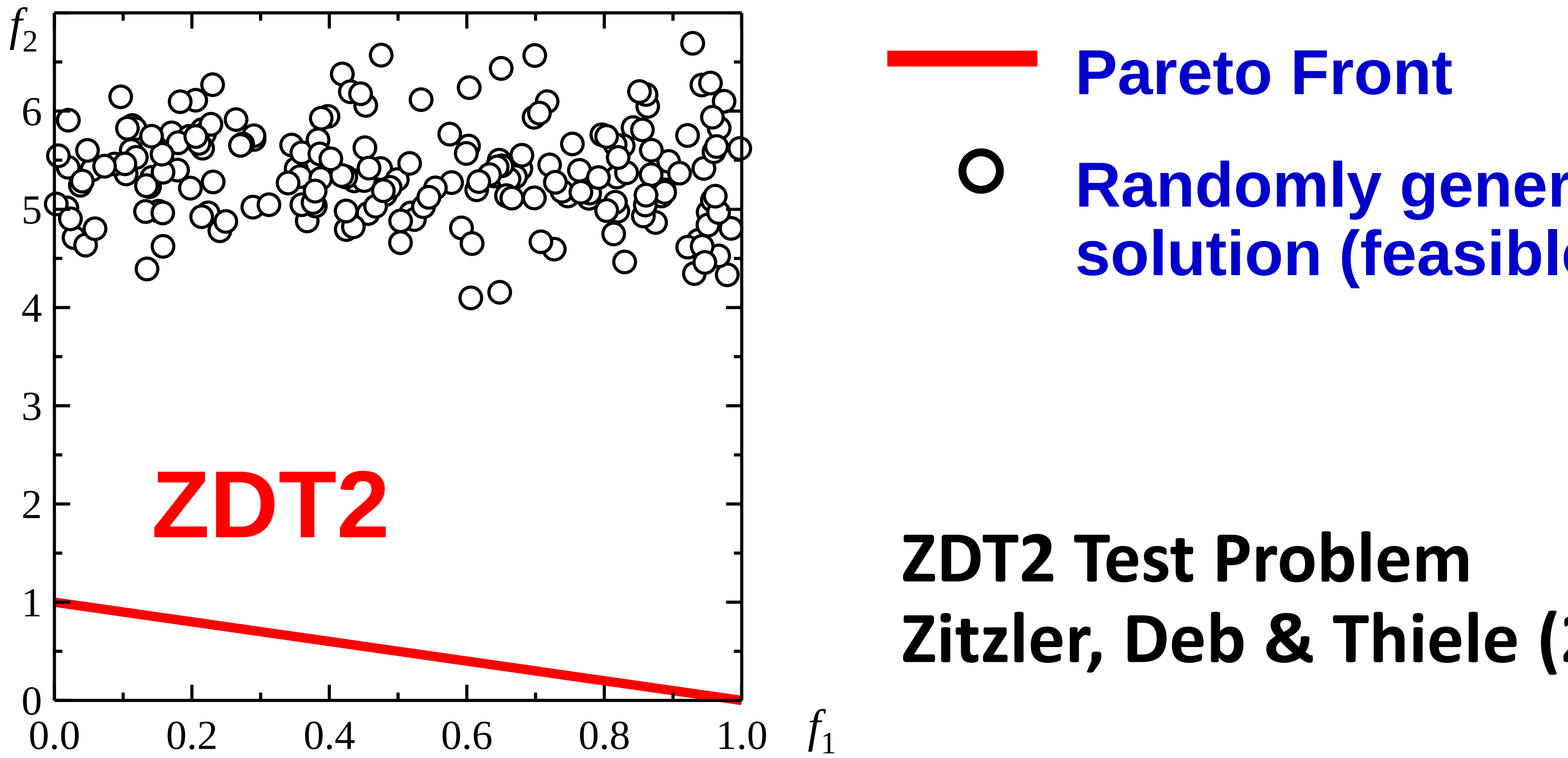


ZDT1 Test Problem
Zitzler, Deb & Thiele (2000)

Pareto Front and Randomly Generated 200 Initial Solutions

No randomly generated solutions are close to the Pareto front ==> Strong convergence ability is needed.

ZDT Test Problems (2000)

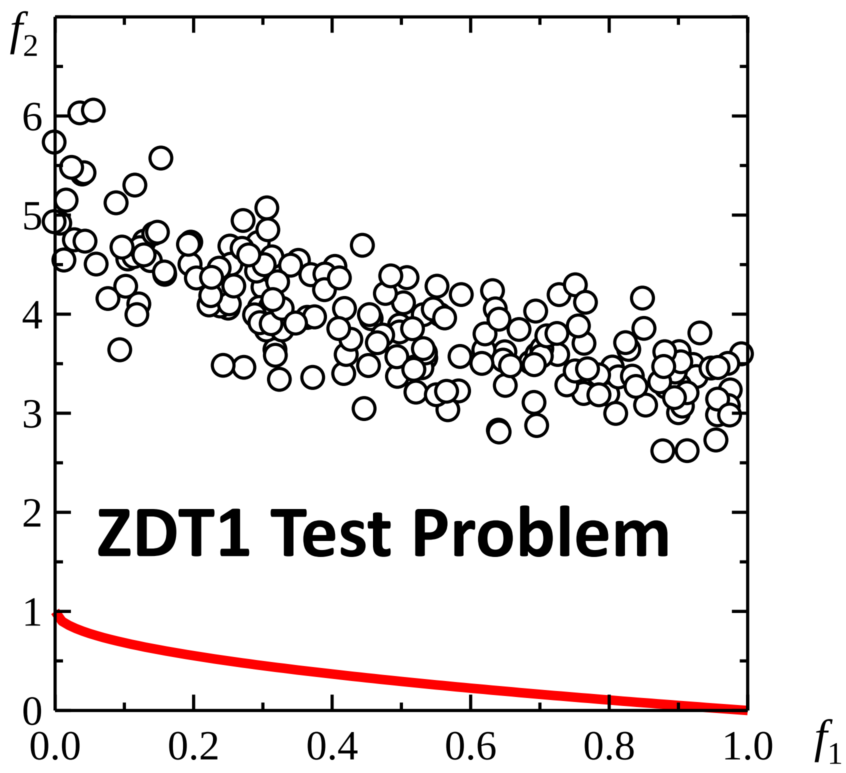


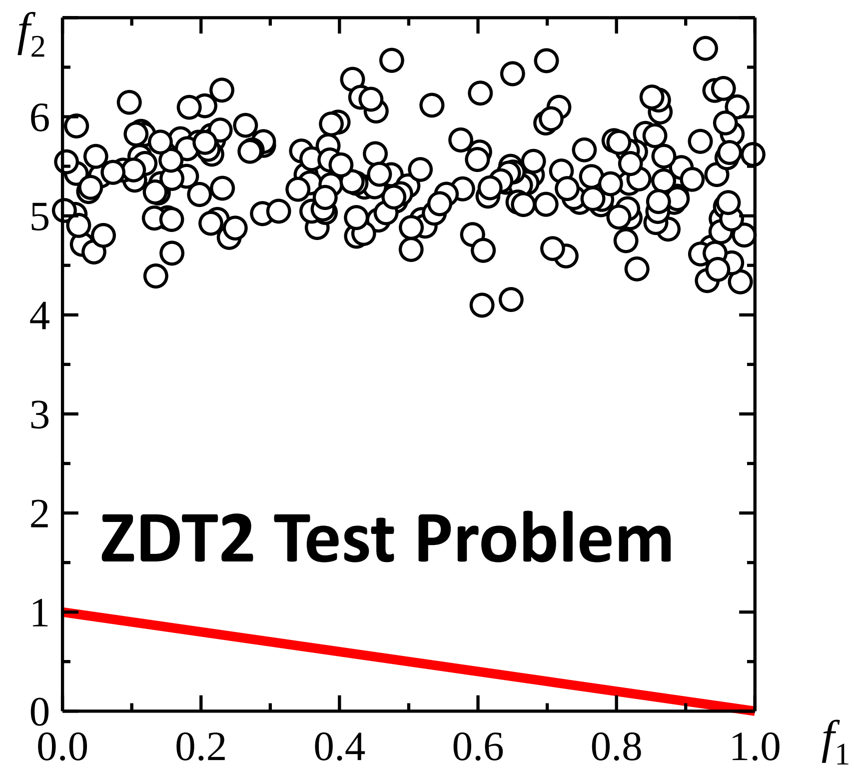
ZDT2 Test Problem
Zitzler, Deb & Thiele (2000)

Pareto Front and Randomly Generated 200 Initial Solutions

No randomly generated solutions are close to the Pareto front ==> Strong convergence ability is needed.

ZDT Test Problems (2000)





Randomly generated initial solutions are not close to the Pareto front.

==> Strong convergence ability is needed.

Highly-Cited Papers published in Late 1990s and 2000s

Elitist EMO Algorithms

[1] **SPEA** **"Multiobjective evolutionary algorithms: A comparative case study and the strength Pareto approach" (E Zitzler, L. Thiele) (1999)**

[2] **NSGA-II** **"A fast and elitist multiobjective genetic algorithm: NSGA-II" (K Deb et al.) (2002)**

[3] **MOEA/D** **"MOEA/D: A multiobjective evolutionary algorithm based on decomposition" (Q Zhang, H Li) (2007)**

Google Scholar

FOLLOW

Kalyanmoy Deb

Koenig Endowed Chair Professor, Electrical and Computer Engineering, Michigan State University
Verified email at egr.msu.edu - Homepage

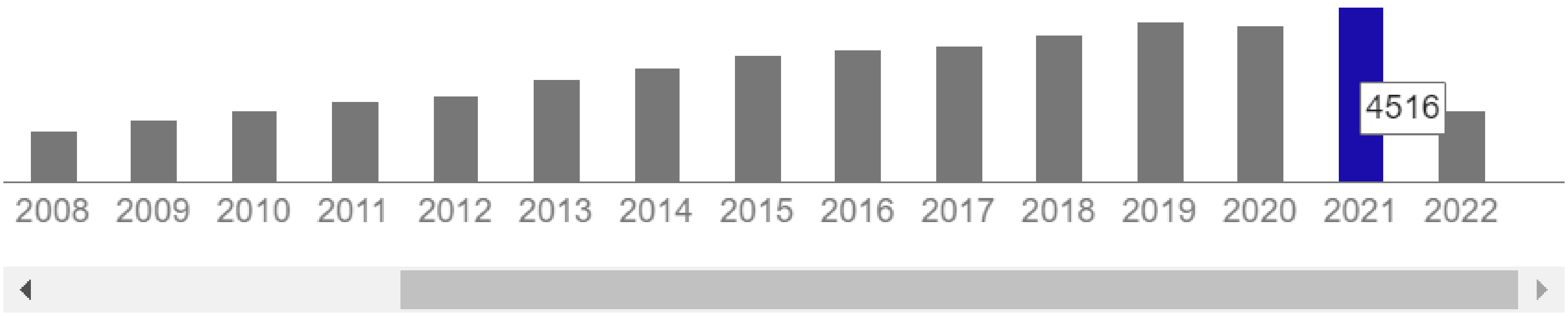
Optimization   Evolutionary Computation   Multiobjective optimization   Machine Learning   Artificial Intelligence

TITLE		CITED BY	YEAR
A fast and elitist multiobjective genetic algorithm: NSGA-II K Deb, A Pratap, S Agarwal, T Meyarivan IEEE transactions on evolutionary computation 6 (2), 182-197	NSGA-II	46091 *	2002
Multi-objective optimization using evolutionary algorithms K Deb Wiley	EMO Book	20166 *	2005
Muiltiobjective optimization using nondominated sorting in genetic algorithms N Srinivas, K Deb Evolutionary computation 2 (3), 221-248		9026	1994
Comparison of multiobjective evolutionary algorithms: Empirical results E Zitzler, K Deb, L Thiele Evolutionary computation 8 (2), 173-195	NSGA	6455	2000

NSGA-II Paper

IEEE Trans. on Evolutionary Computation, 2002.

Cited by 46091



A fast and elitist multiobjective genetic algorithm: NSGA-II
K Deb, A Pratap, S Agarwal, T Meyarivan - IEEE transactions on evolutionary computation, 2002
Cited by 40242   Related articles   All 37 versions

Highly-Cited Papers published in Late 1990s and 2000s

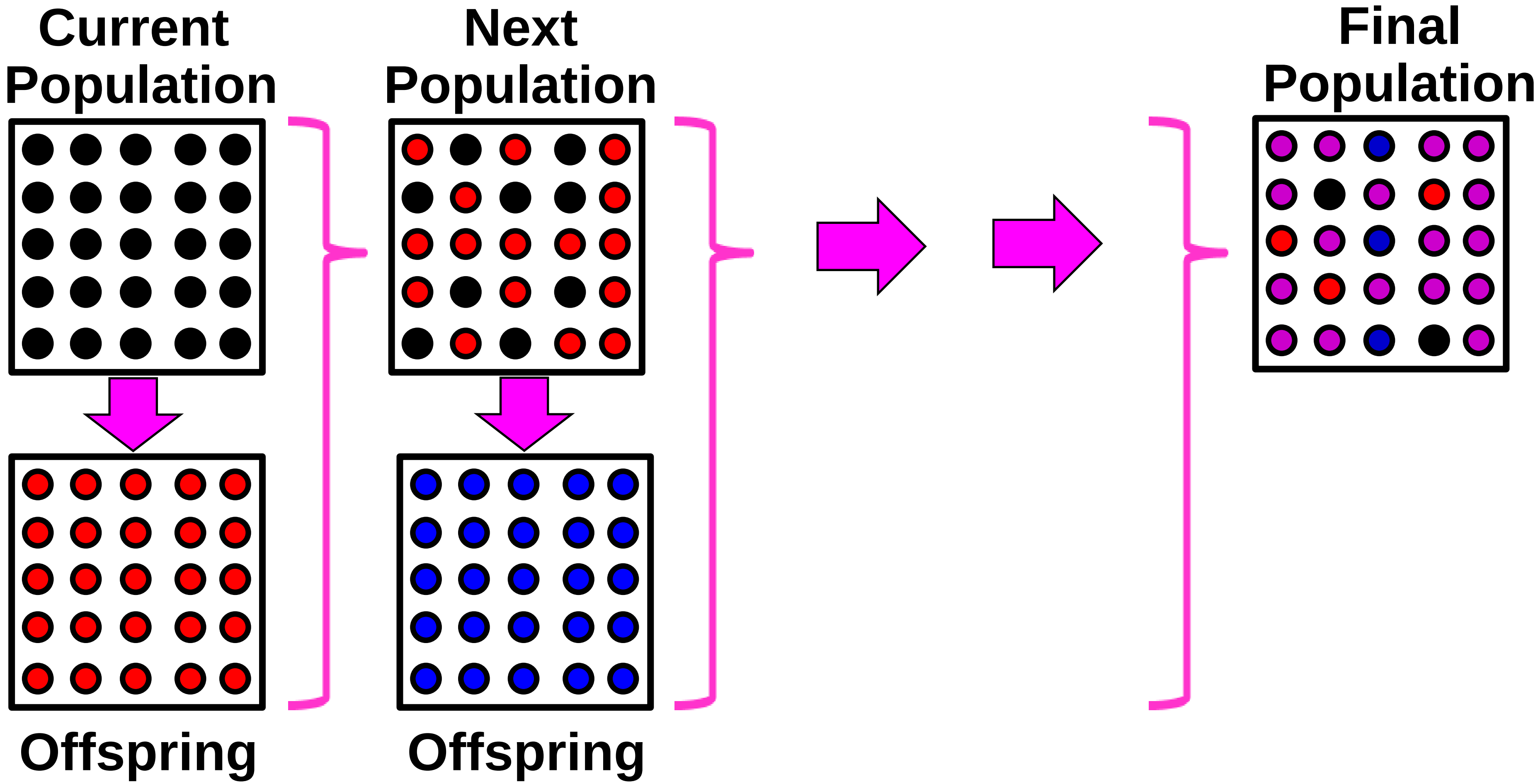
Main Feature: Strong Convergence Ability

(1) Elitist

The point: To generate some good offspring and to appropriately choose solutions for the next population

Highly-Cited Papers published in Late 1990s and 2000s

The point: To generate some good offspring and to appropriately choose solutions for the next population

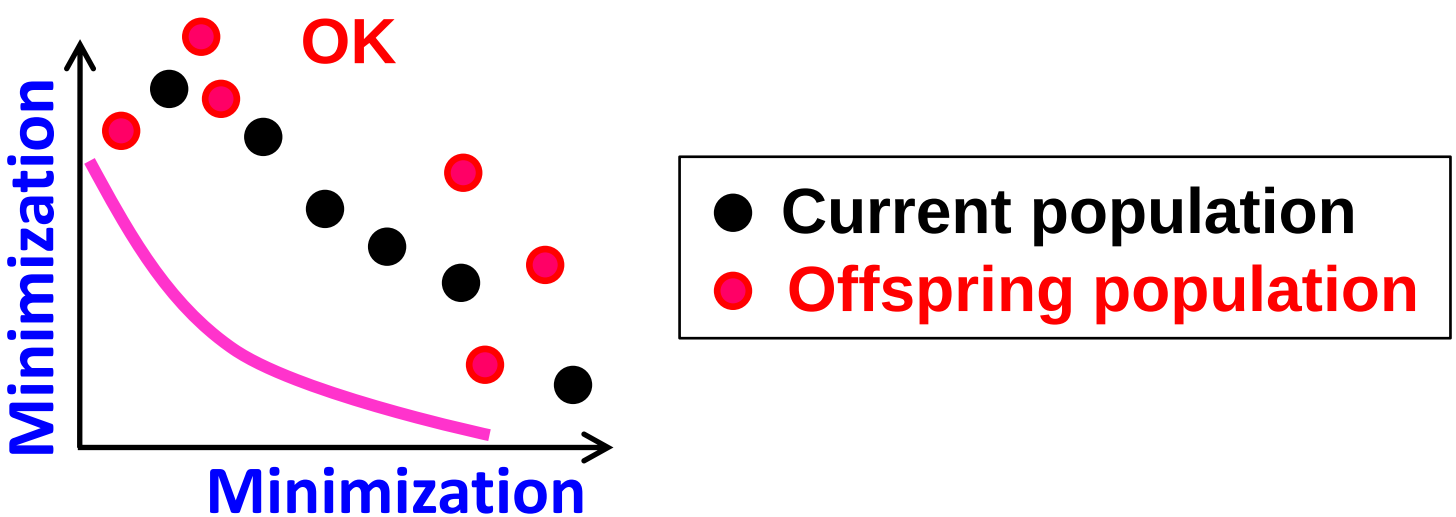


Highly-Cited Papers published in Late 1990s and 2000s

Main Feature: Strong Convergence Ability

(1) Elitist

The point: To generate some good offspring and to appropriately choose solutions for the next population

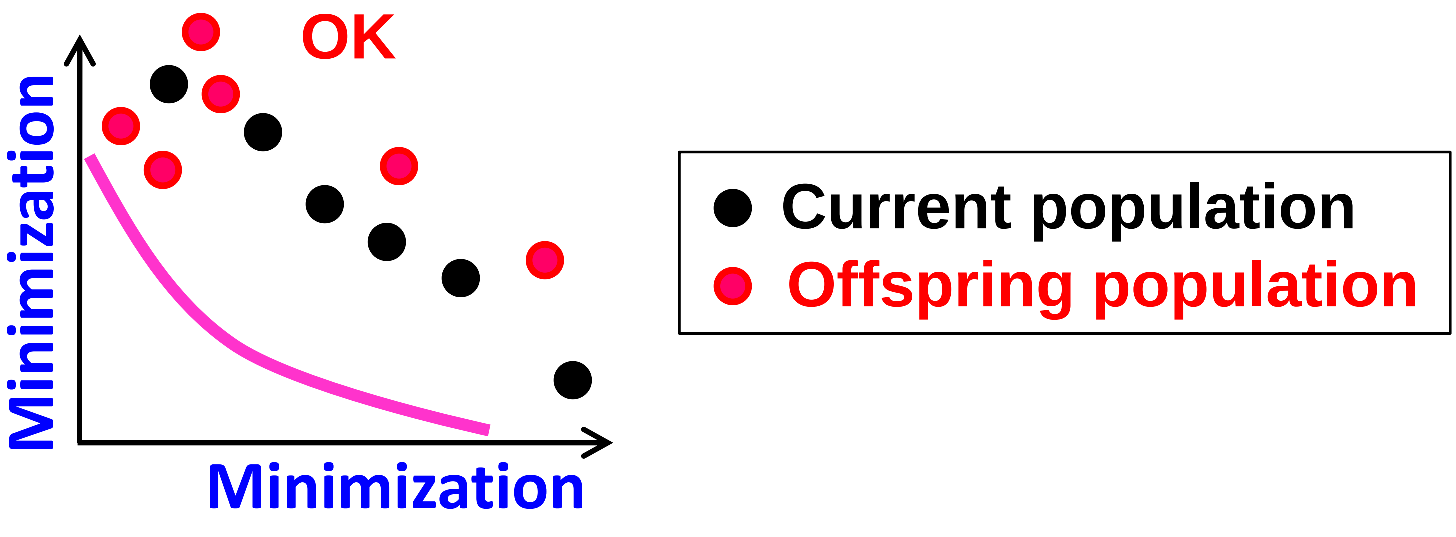


Highly-Cited Papers published in Late 1990s and 2000s

Main Feature: Strong Convergence Ability

(1) Elitist

The point: To generate some good offspring and to appropriately choose solutions for the next population

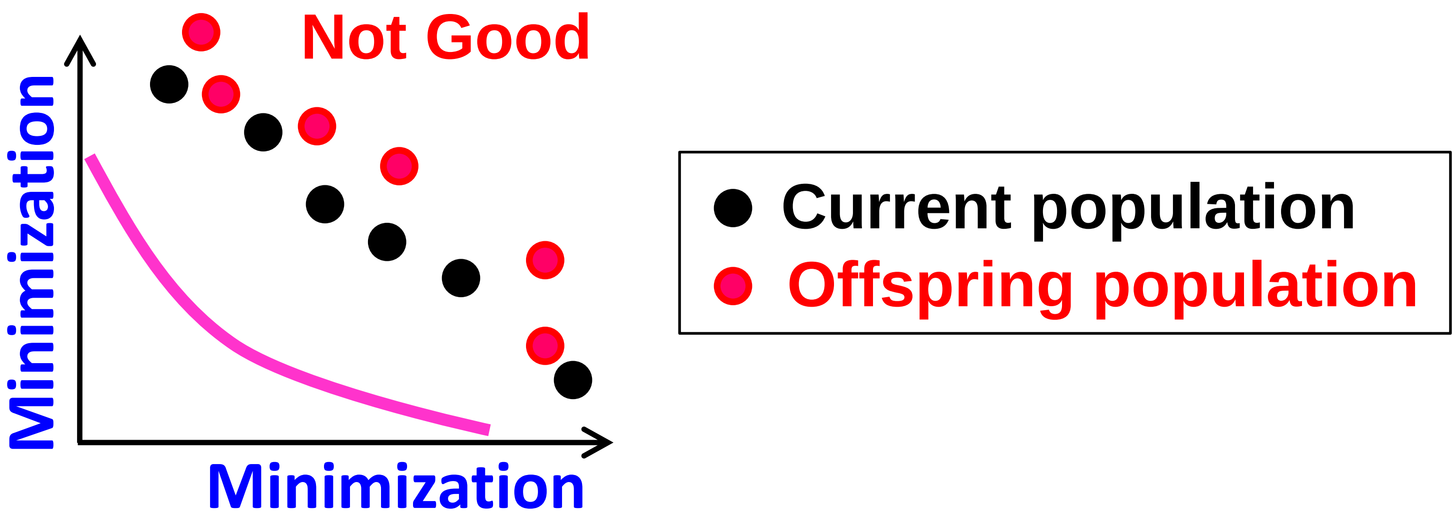


Highly-Cited Papers published in Late 1990s and 2000s

Main Feature: Strong Convergence Ability

(1) Elitist

The point: To generate some good offspring and to appropriately choose solutions for the next population



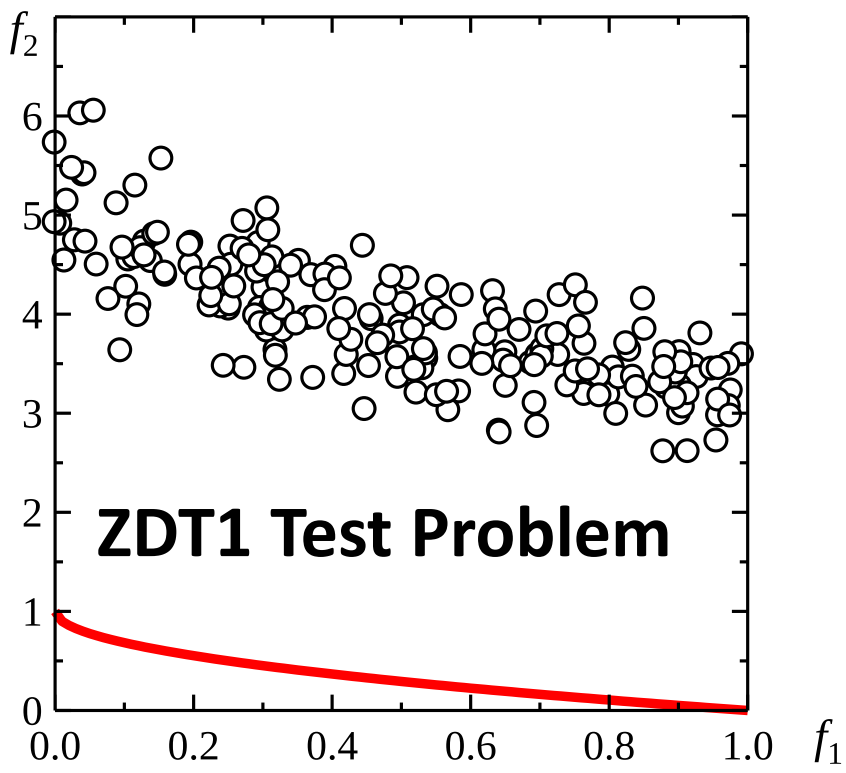
Even in this case, the population is not degraded since the current population will be used as the next population. This is a clear difference from non-elitist algorithms.

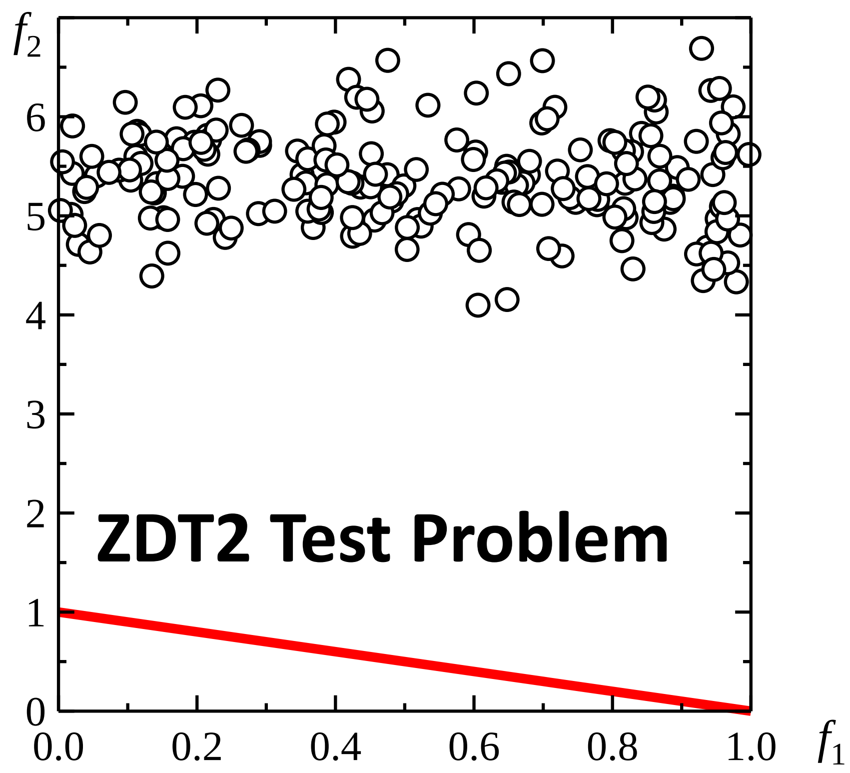
Why Elitist EMO Algorithms ?

Answer:

Strong convergence ability is needed.

ZDT Test Problems (2000)

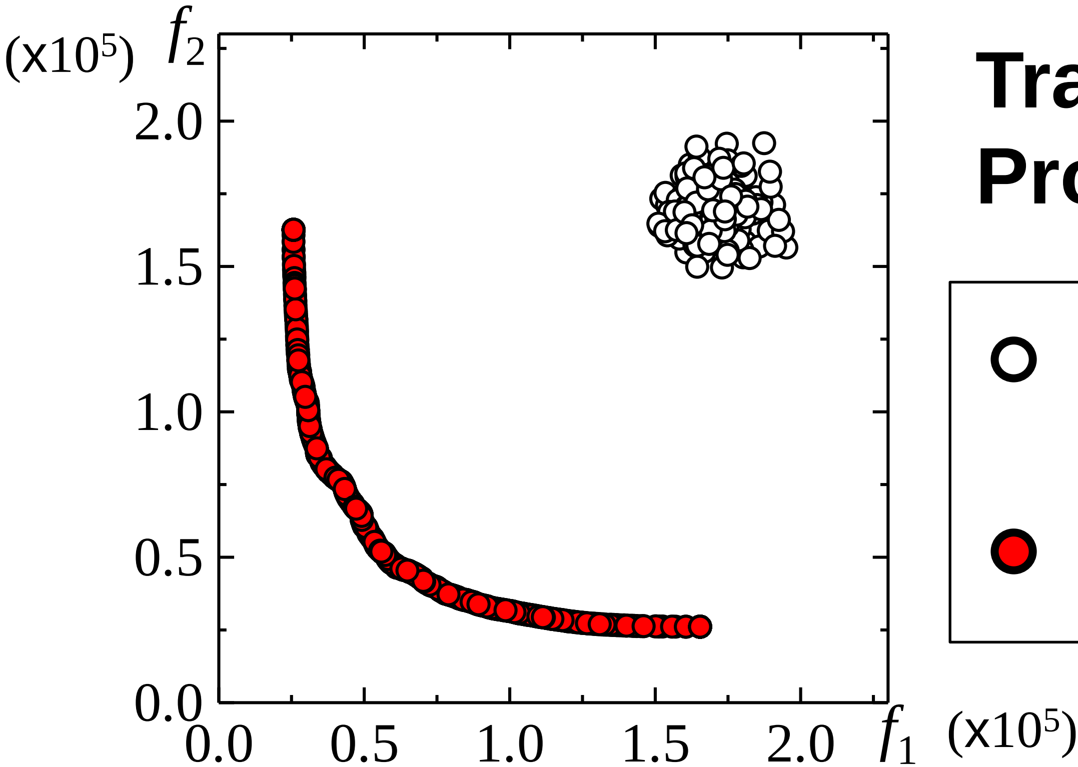




Randomly generated initial solutions are not close to the Pareto front.

==> Strong convergence ability is needed.

When we have different problems, different algorithms are developed.



Travelling Salesperson Problem (TSP)

Randomly generated solutions are far from the Pareto front.
The Pareto front has larger diversity than initial solutions.

Strong convergence ability and strong diversification ability are needed: Memetic EMO (Jaszkiewicz (2002))

NSGA-II Algorithm (IEEE TEVC, 2002)

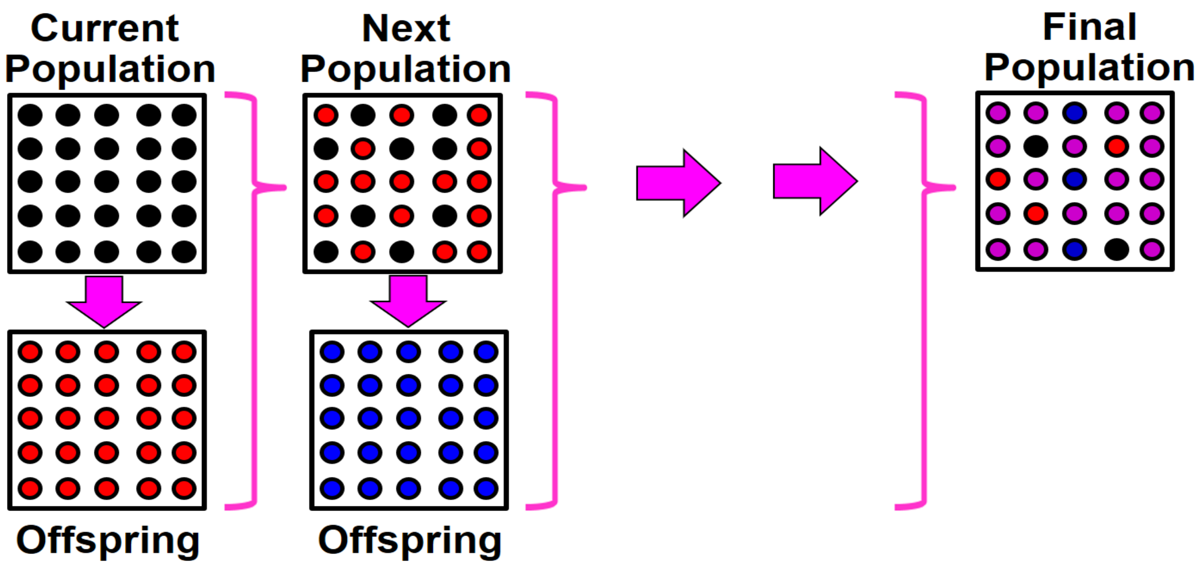
Algorithm Framework: (+)-Style

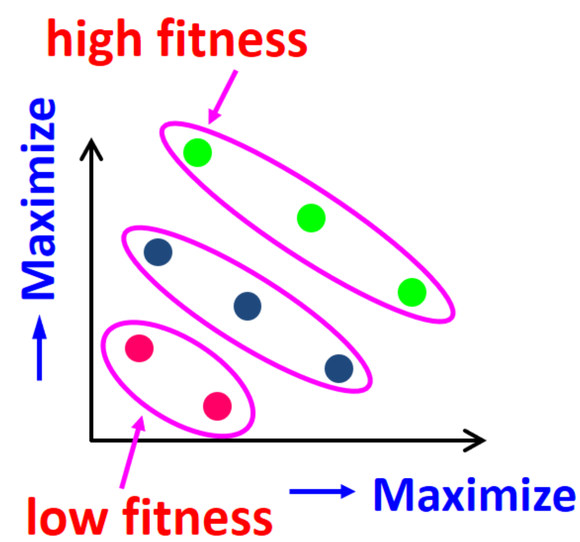
- Main population size:
- Offspring population size:
- Next population:

The best solutions in the current and offspring populations

Fitness Evaluation:

- Main Criterion: Non-dominated sorting
- Secondary Criterion: Crowding distance

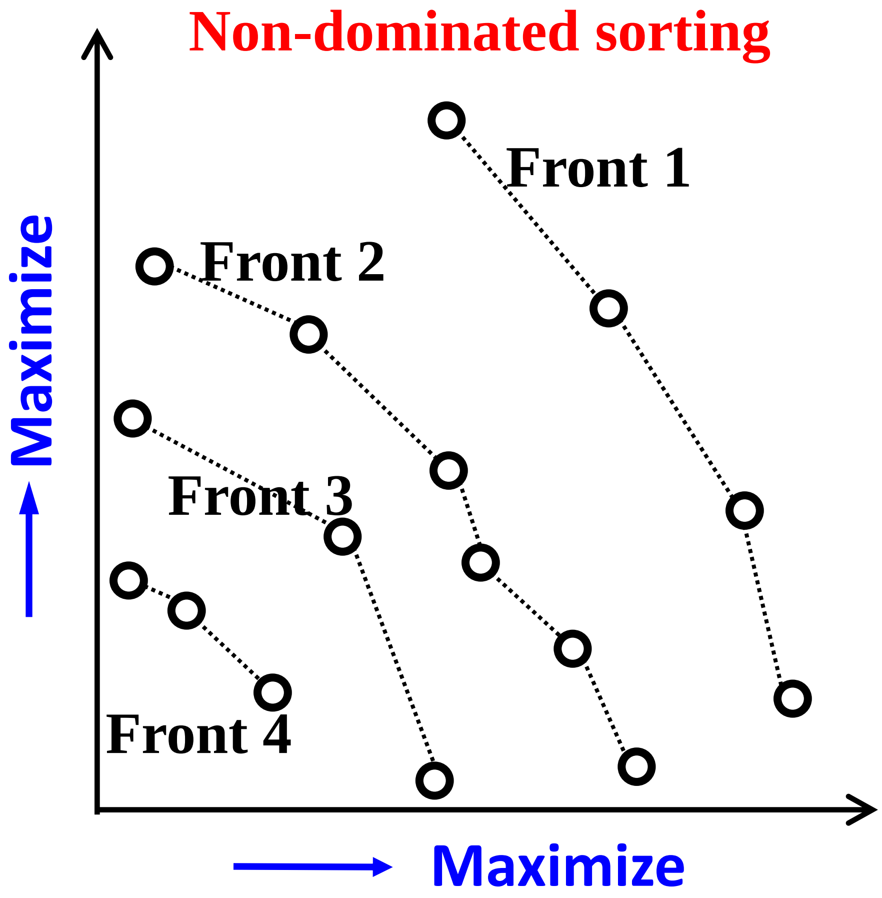


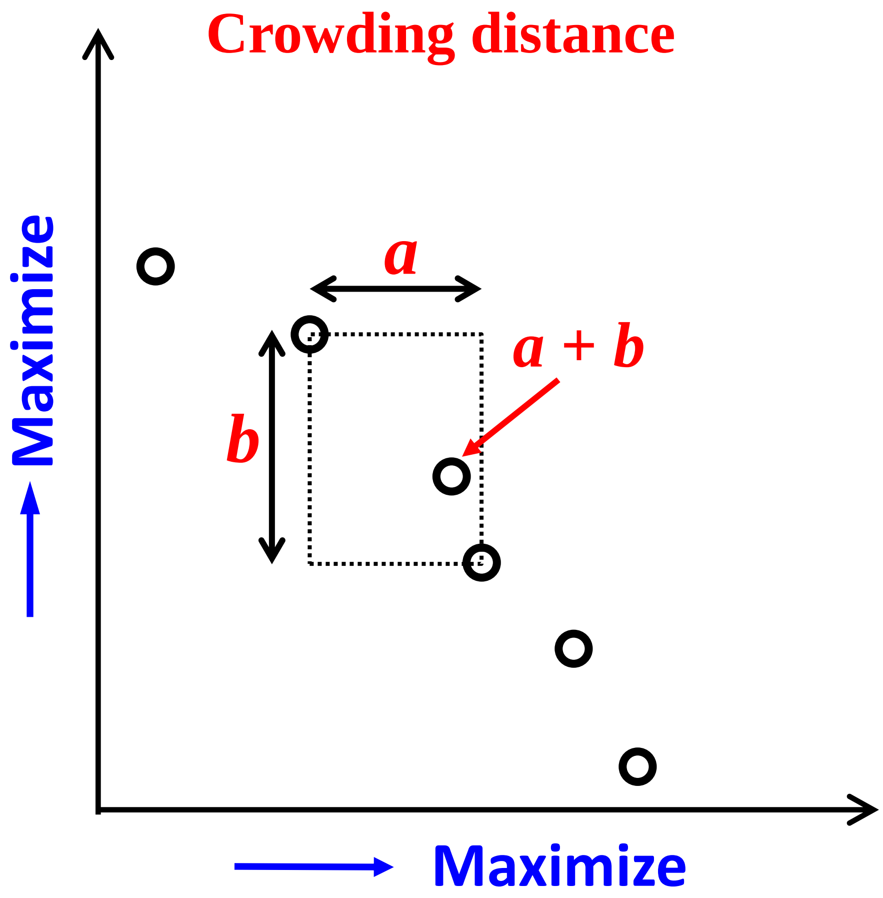


NSGA-II Algorithm (IEEE TEVC, 2002)

Fitness Evaluation:

- Main Criterion: Non-dominated sorting
- Secondary Criterion: Crowding distance



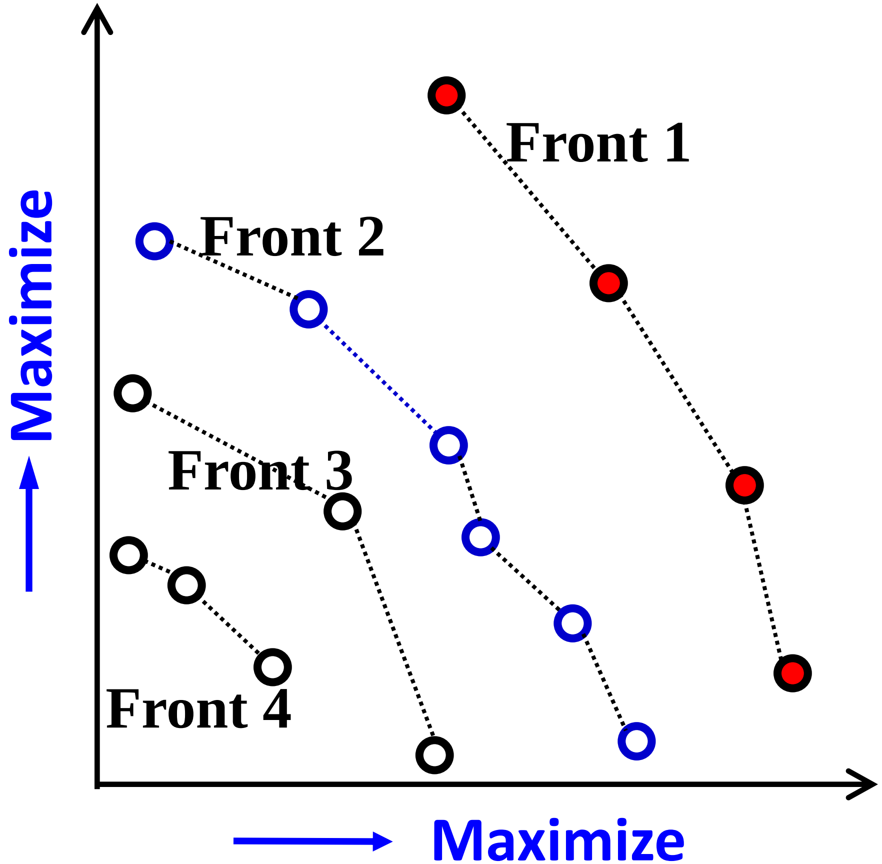


NSGA-II Algorithm (IEEE TEVC, 2002)

Fitness Evaluation:

- Main Criterion: Non-dominated sorting
- Secondary Criterion: Crowding distance

Generation Update: (8 + 8): Selection of the best 8 solutions



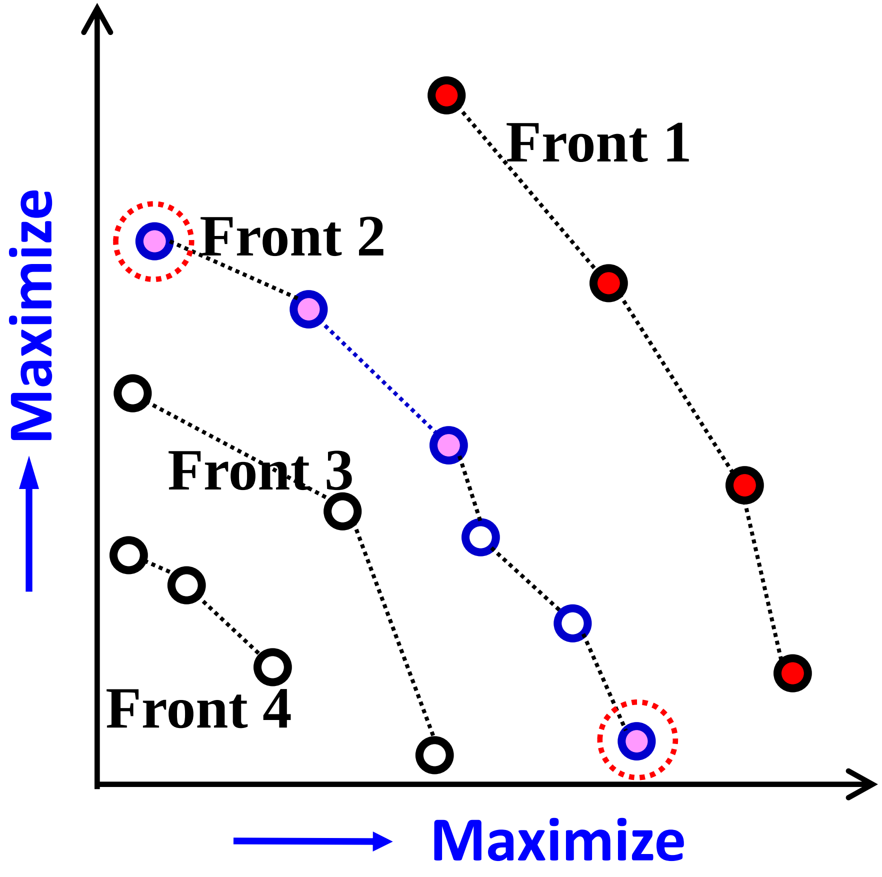
(1) Front 1:
All the four solutions.

NSGA-II Algorithm (IEEE TEVC, 2002)

Fitness Evaluation:

- Main Criterion: Non-dominated sorting
- Secondary Criterion: Crowding distance

Generation Update: (8 + 8): Selection of the best 8 solutions



(1) Front 1:
All the four solutions.

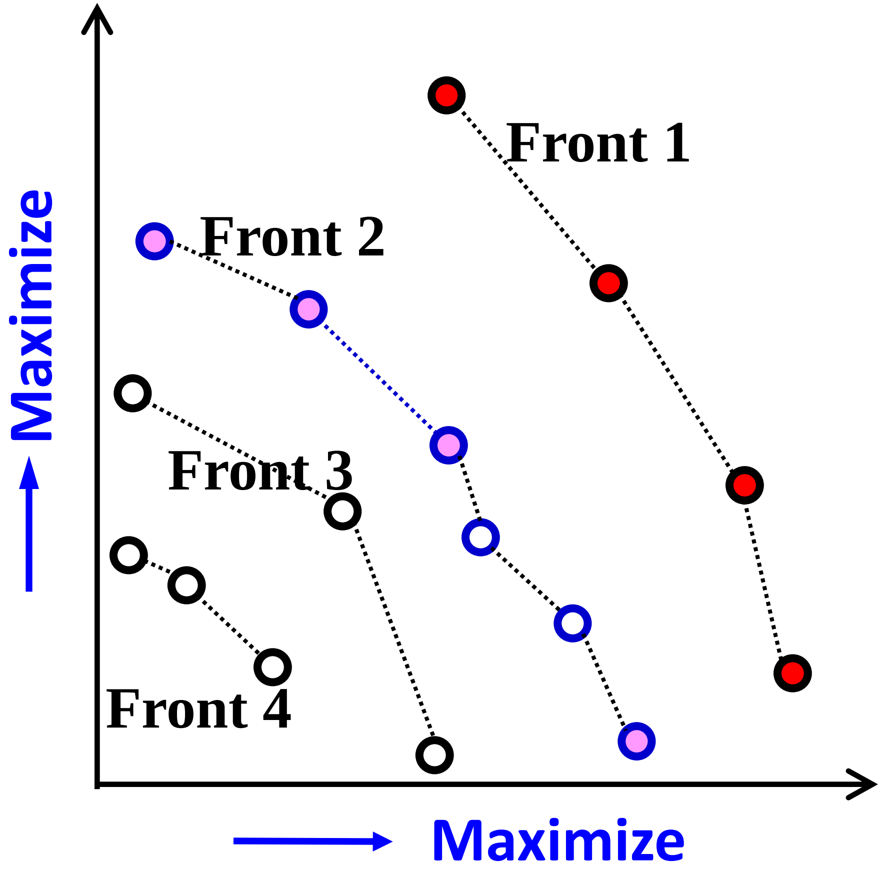
(2) Front 2:
Four out of the six solutions.
(based on the crowding distance).

NSGA-II Algorithm (IEEE TEVC, 2002)

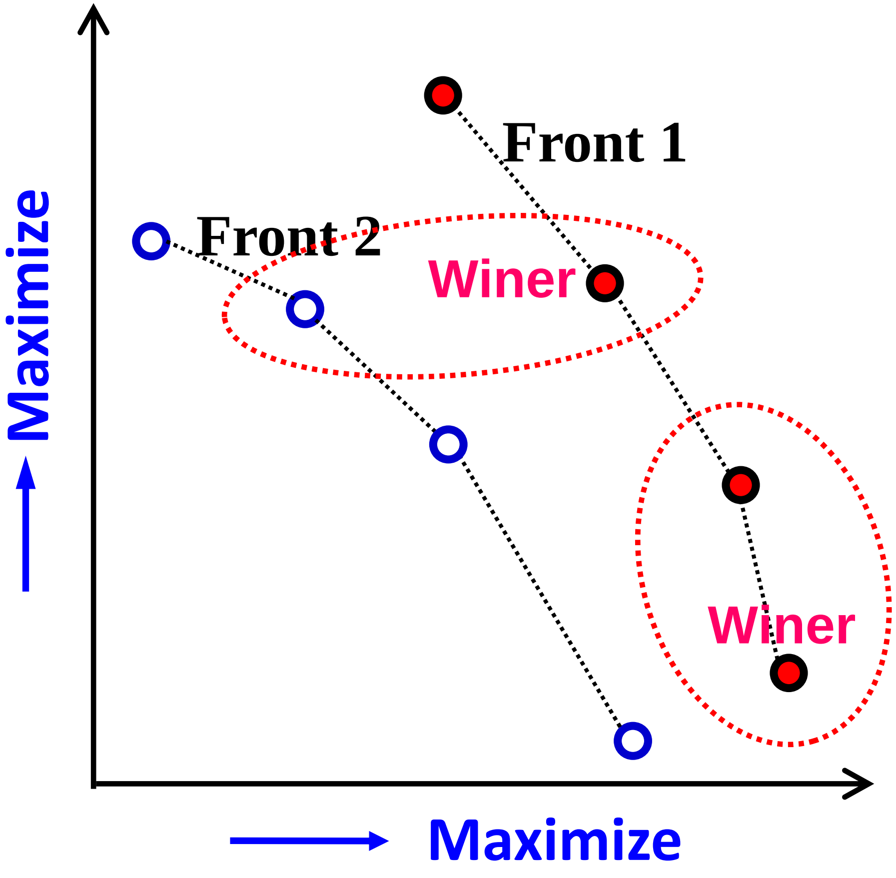
Fitness Evaluation:

- Main Criterion: Non-dominated sorting
- Secondary Criterion: Crowding distance

Generation Update: (8 + 8)



Parent Selection: Binary Tournament

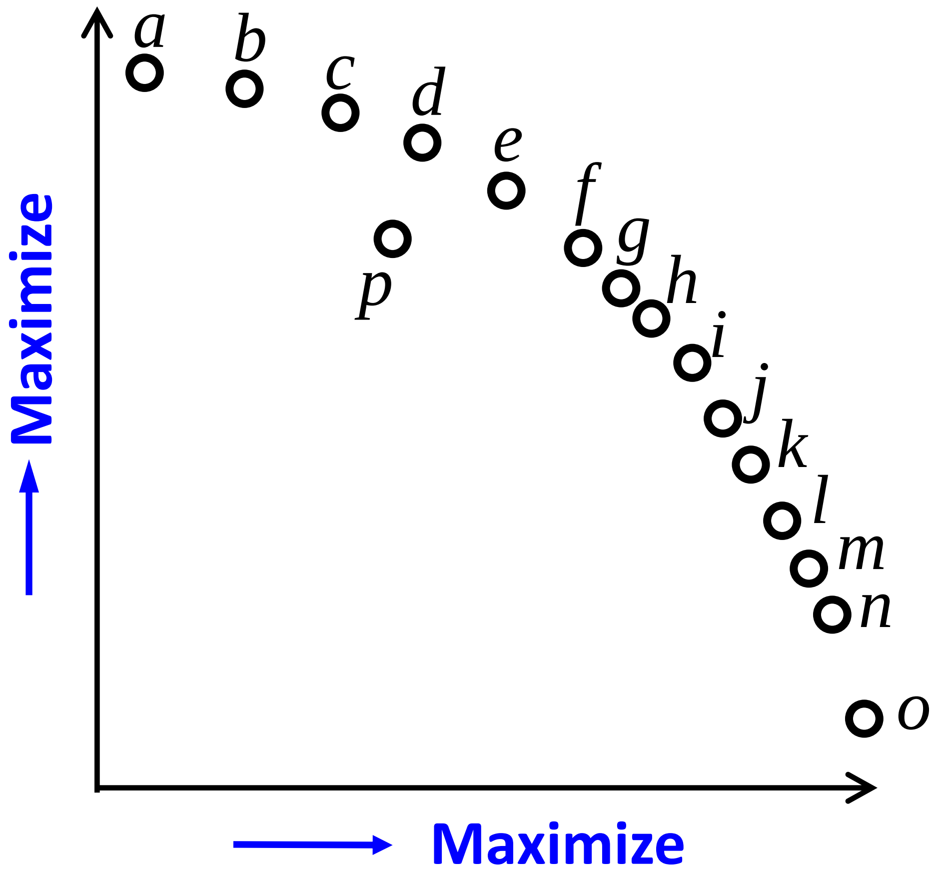


Difficulties of NSGA-II

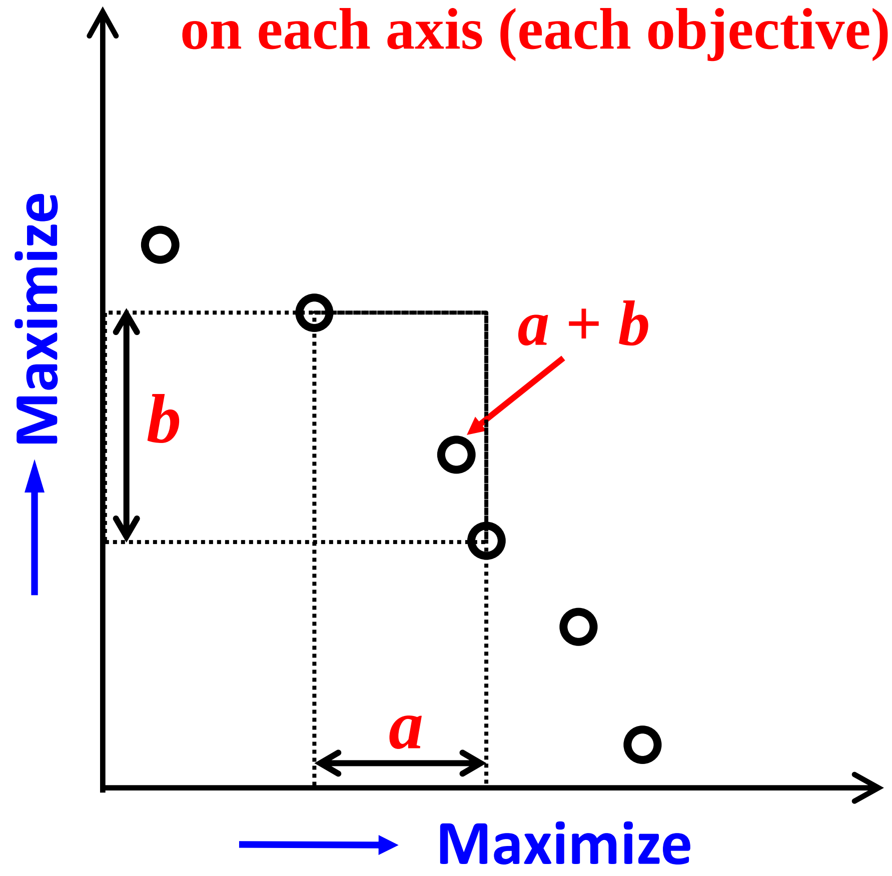
- (+   )-style framework
- Projection-based crowding distance calculation
- Pareto dominance-based main fitness criterion

Question: Next population (8 solutions): ? ? ? ? ? ? ? ? .

Generation Update: (8 + 8)



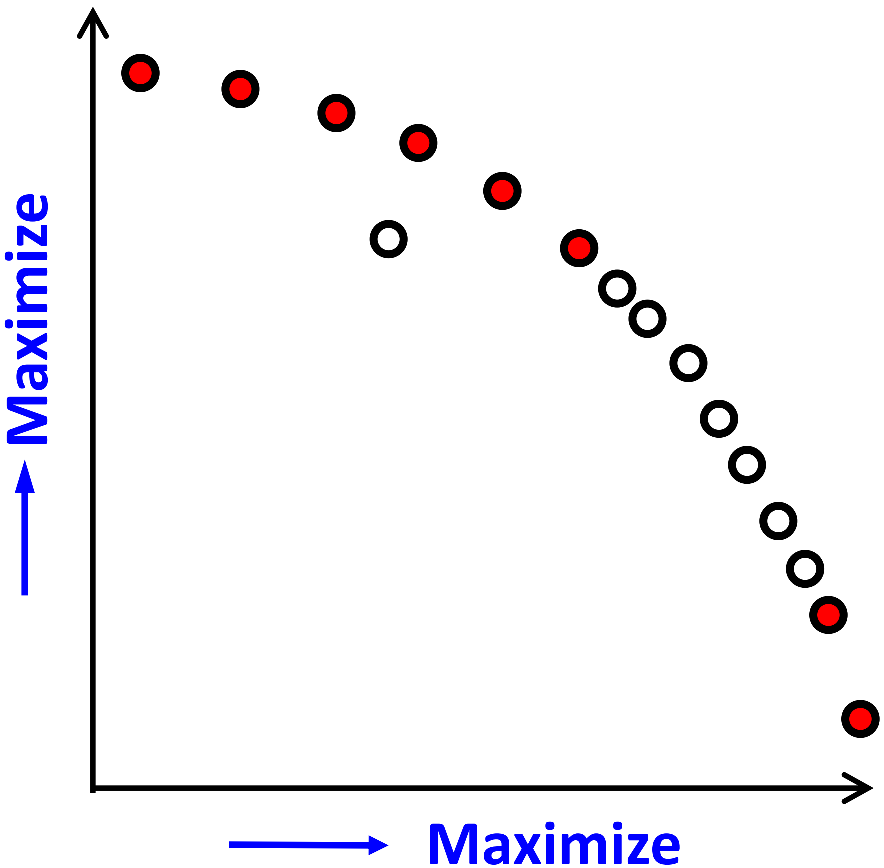
Crowding distance calculation on each axis (each objective)



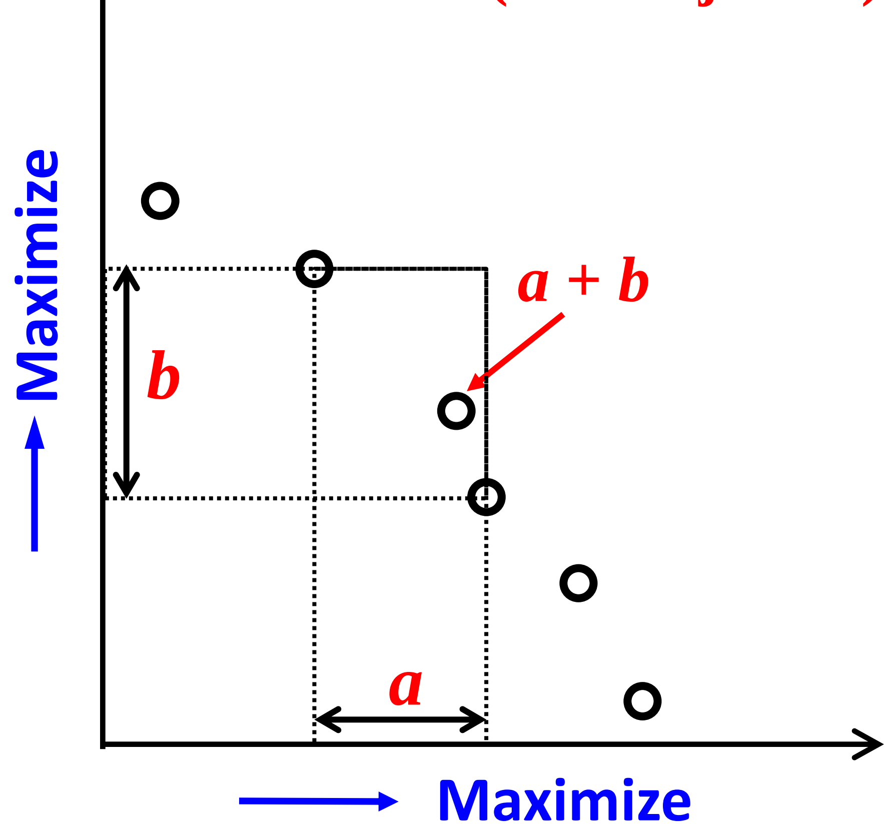
Difficulties of NSGA-II

- (+)-style framework
- **Projection-based crowding distance calculation**
- Pareto dominance-based main fitness criterion

Generation Update: (8 + 8)



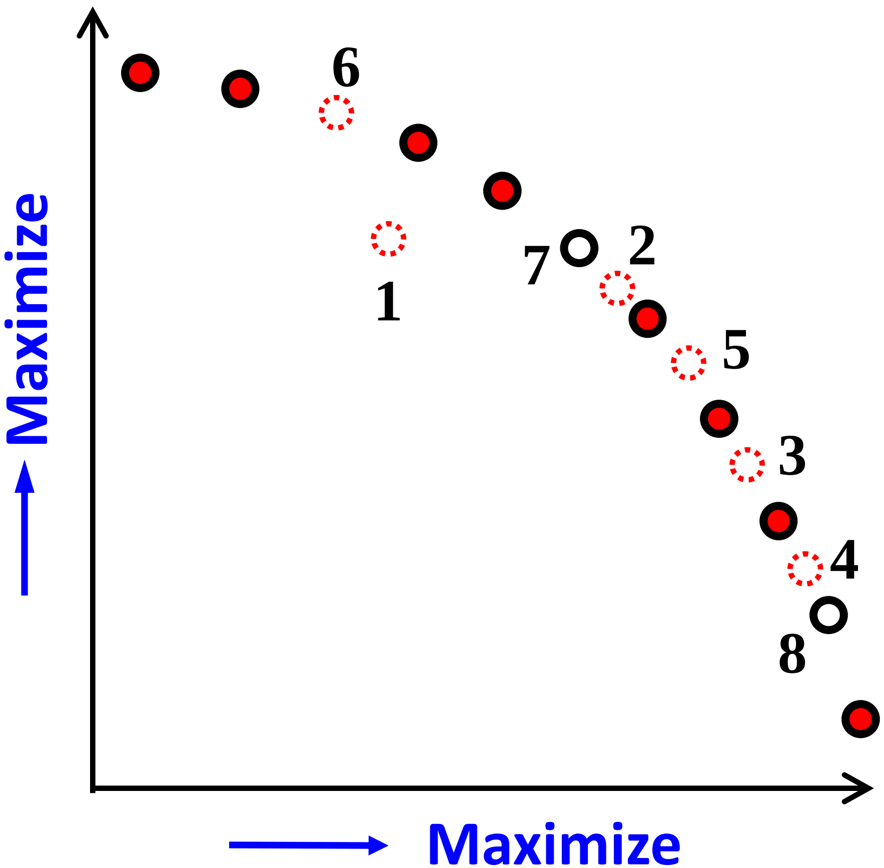
Crowding distance calculation on each axis (each objective)



Difficulties of NSGA-II

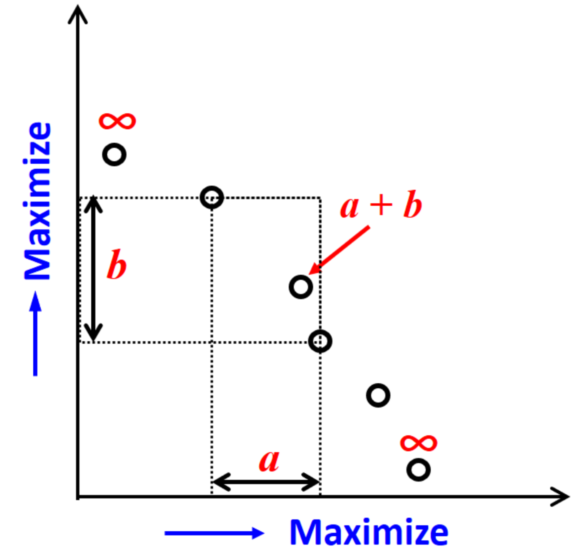
- (+)-style framework ==> Recalculation or (+1)-styl
- Projection-based crowding distance calculation
- Pareto dominance-based main fitness criterion

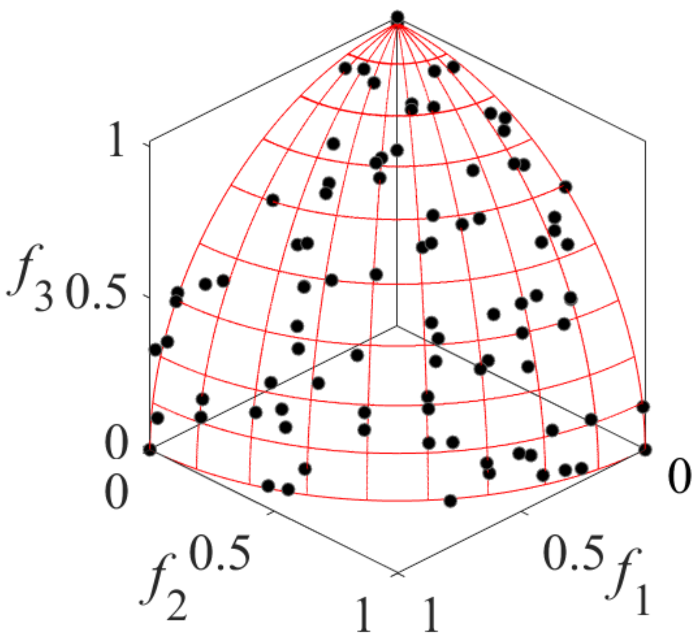
Generation Update: (8 + 8)



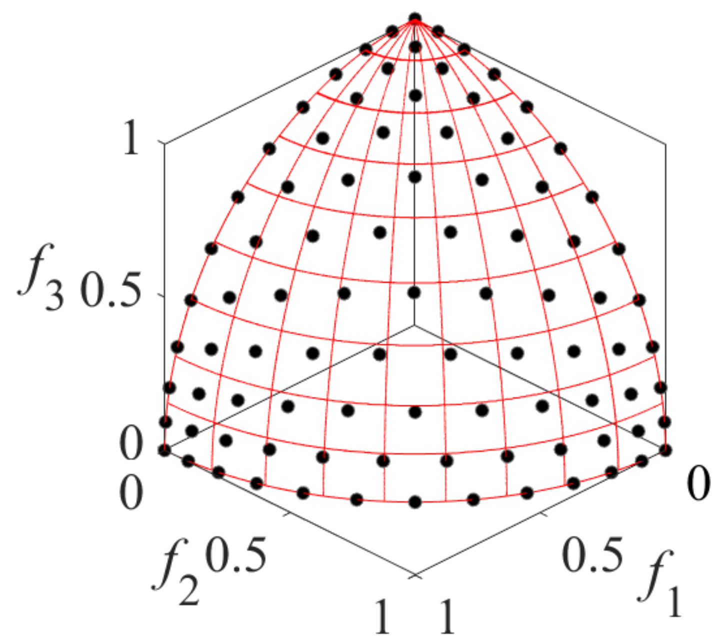
Crowding Distance:

Crowding distance maximization leads to a uniformly distributed solution set on the Pareto front in the case of two objectives. However, it does not lead to a good solution set as shown in the following figures in the case of three or more objectives.





Larger crowding distance solutions (Obtained by NSGA-II)

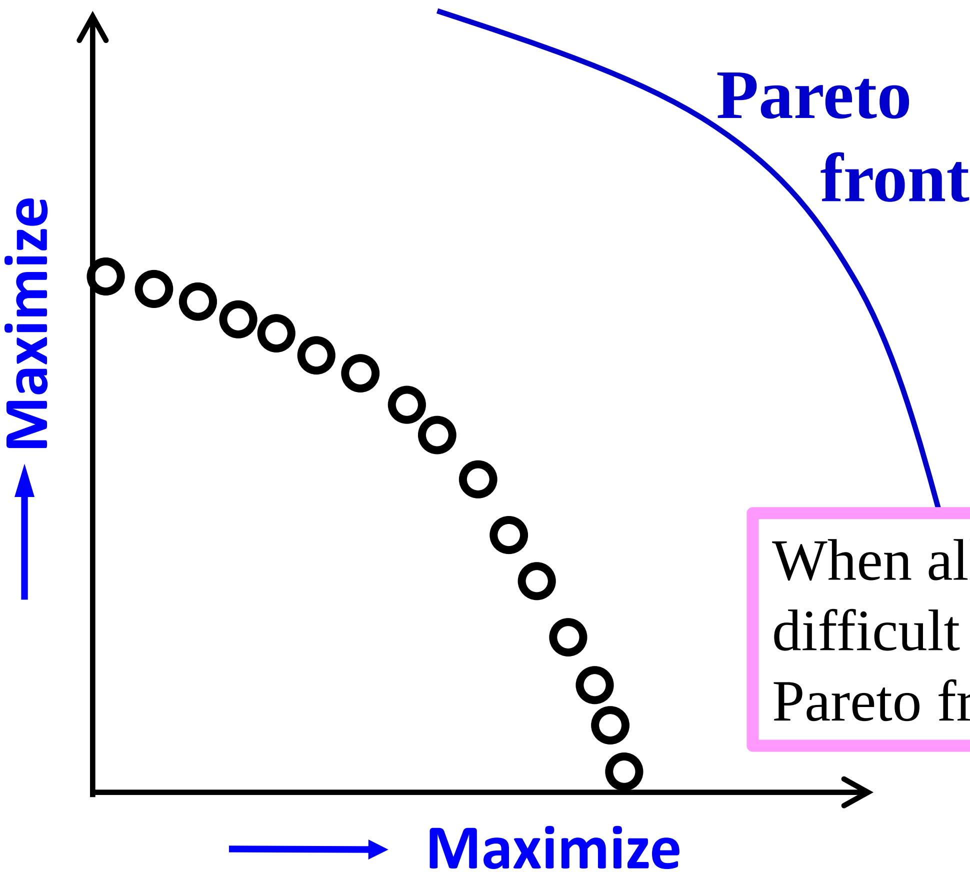


Small crowding distance solutions (Obtained by MOEA/D)

Difficulties of NSGA-II

- (+   )-style framework
- Projection-based crowding distance calculation
- **Pareto dominance-based main fitness criterion**

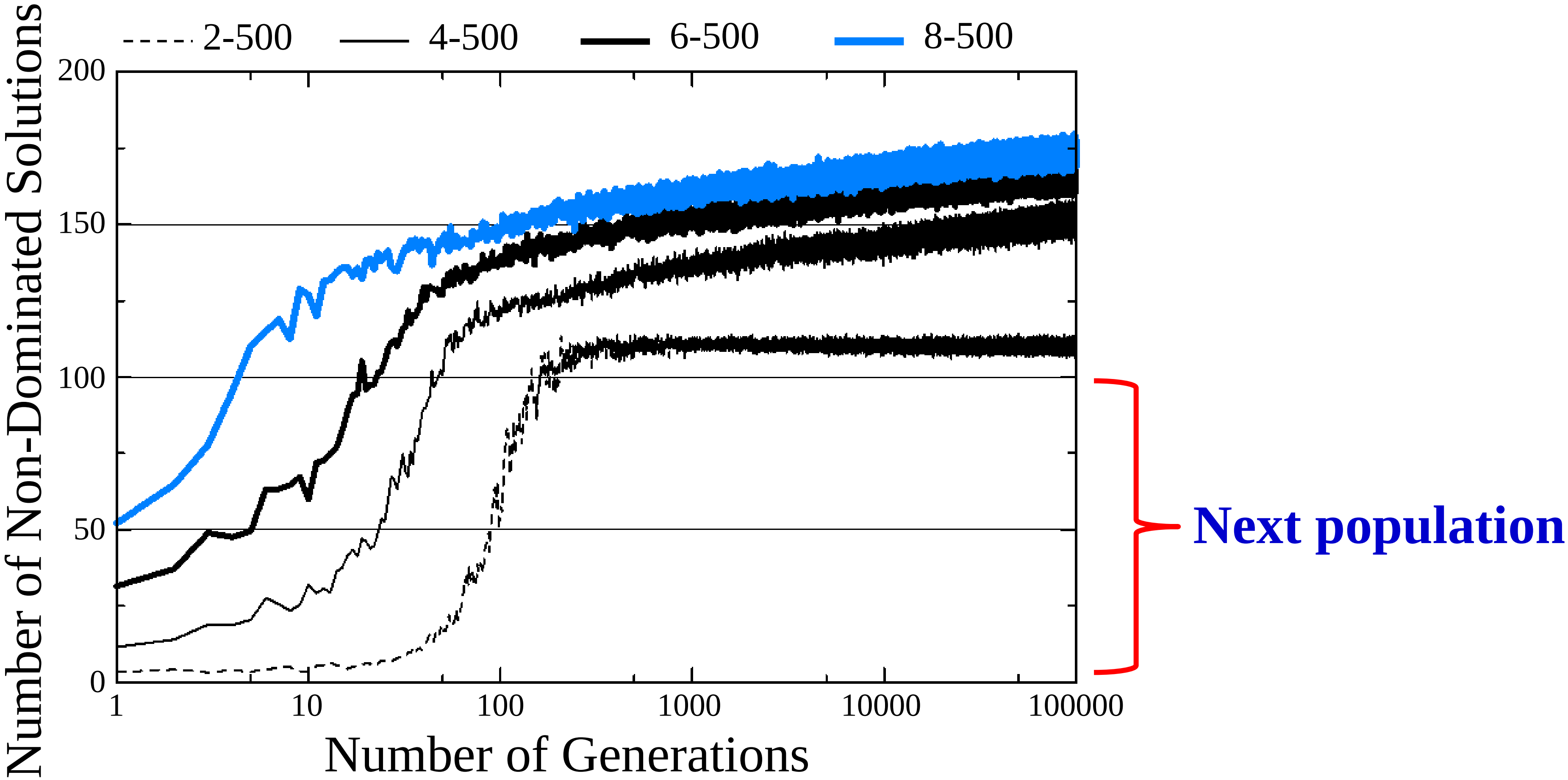
Generation Update: (8 + 8)



When all solutions are non-dominated, it is difficult to push the population towards the Pareto front.

Number of Non-Dominated Solutions in NSGA-II

NSGA-II on knapsack problems (Population size: 100, Offspring size: 100)

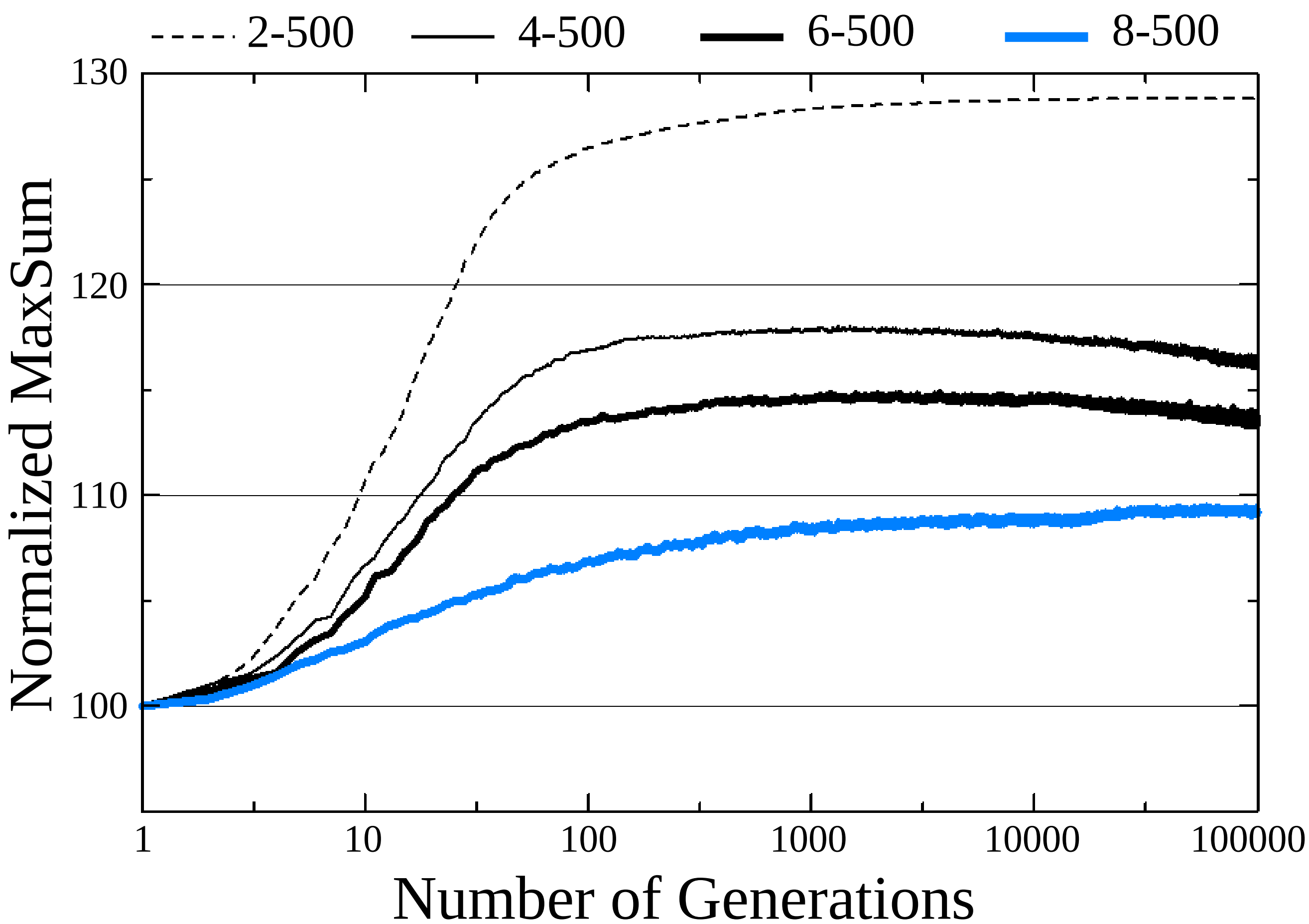


Number of non-dominated solution in the merged population

H. Ishibuchi, N. Tsukamoto and Y. Nojima, "Evolutionary many-objective optimization: A short review," Proc. of IEEE CEC 2008.

Sum of Objectives in NSGA-II

NSGA-II on knapsack problems (Population size: 100, Offspring size: 100)

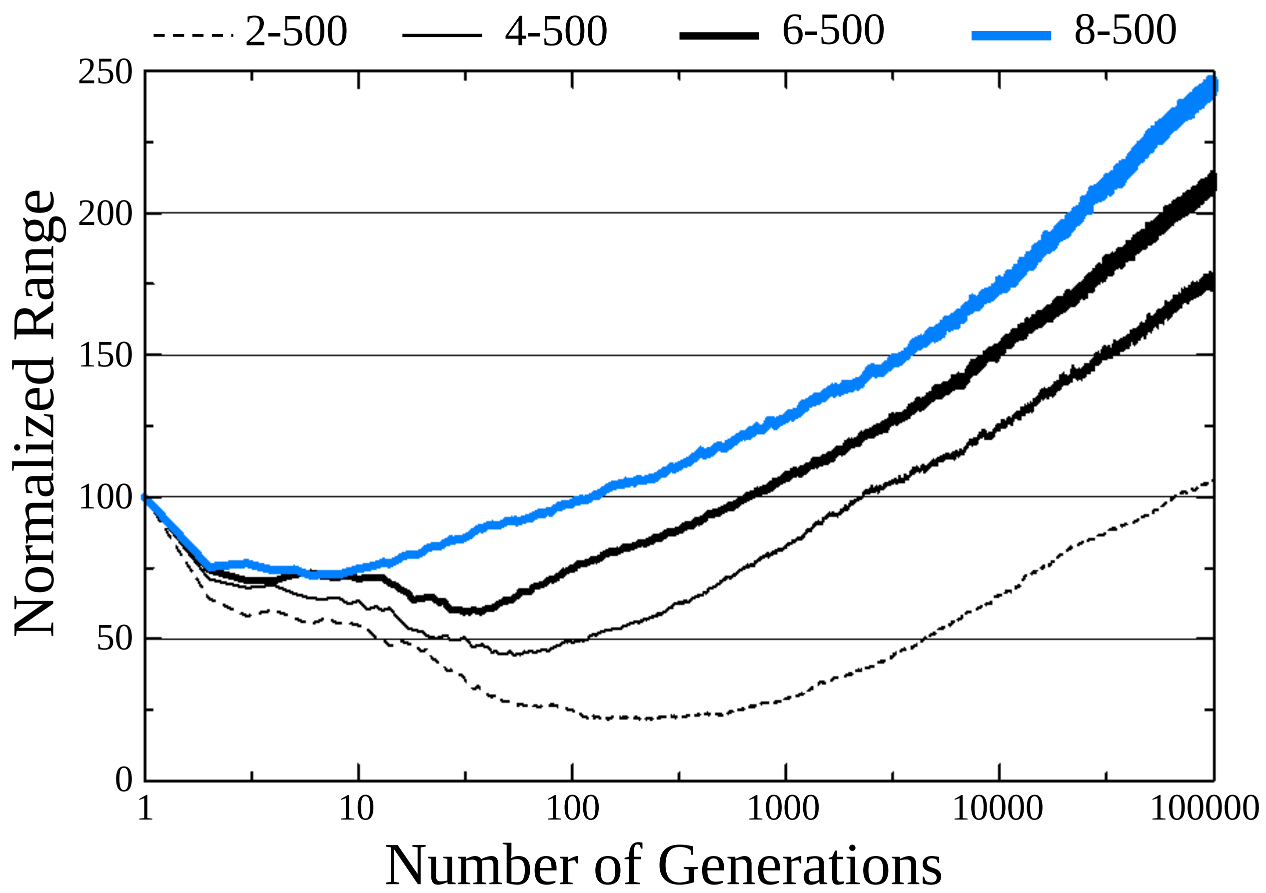


Convergence Performance around the Center: $f_1(x) + f_2(x) + \ldots$

H. Ishibuchi, N. Tsukamoto and Y. Nojima, "Evolutionary many-objective optimization: A short review," Proc. of IEEE CEC 2008.

Range (Maximum Spread) of Population in NSGA-II

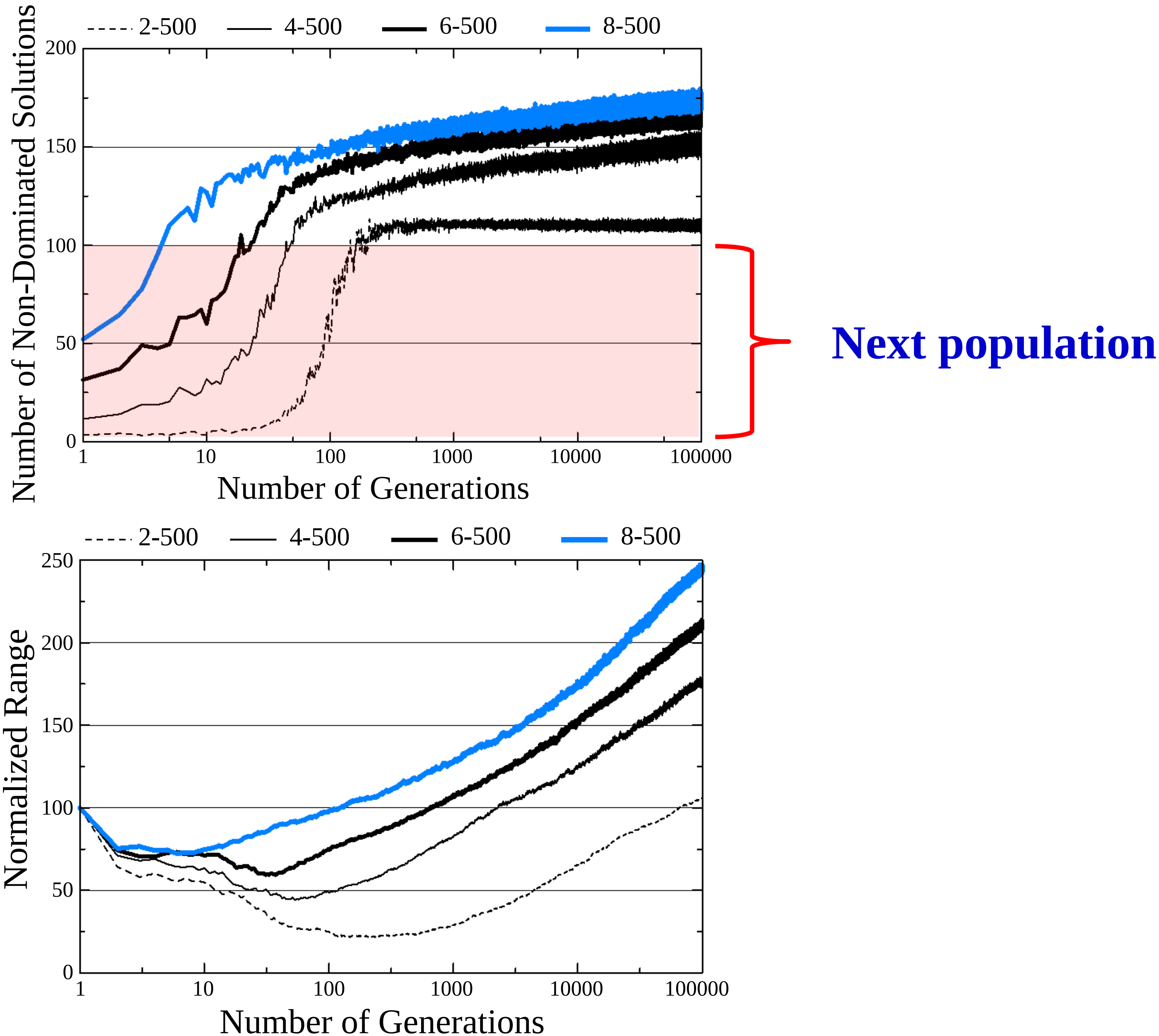
NSGA-II on knapsack problems (Population size: 100, Offspring size: 100)



Diversity of Population (Range):

$$|\mathrm{Max}\{f_1(x)\} - \mathrm{Min}\{f_1(x)\}| + |\mathrm{Max}\{f_2(x)\} - \mathrm{Min}\{f_2(x)\}| + \ldots$$

H. Ishibuchi, N. Tsukamoto and Y. Nojima, “Evolutionary many-objective optimization: A short review,” Proc. of IEEE CEC 2008.

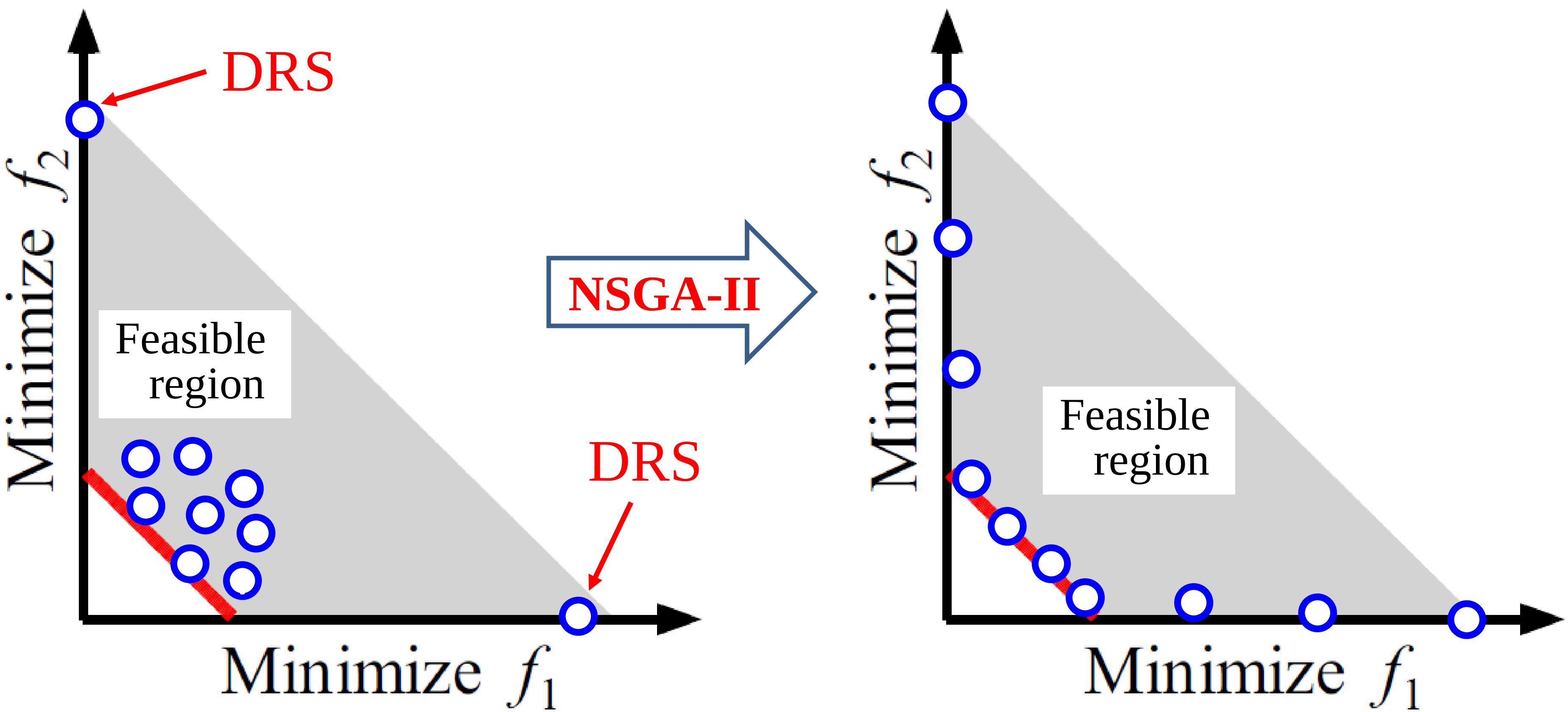


2-500
4-500
6-500
8-500
Number of Non-Dominated Solutions
200
150
100
50
0
1
10
100
1000
10000
100000
Number of Generations
Next population
2-500
4-500
6-500
8-500
Normalized Range
250
200
150
100
50
0
1
10
100
1000
10000
100000
Number of Generations

Dominance Resistant Solutions

Dominance Resistant solutions (DRSs):

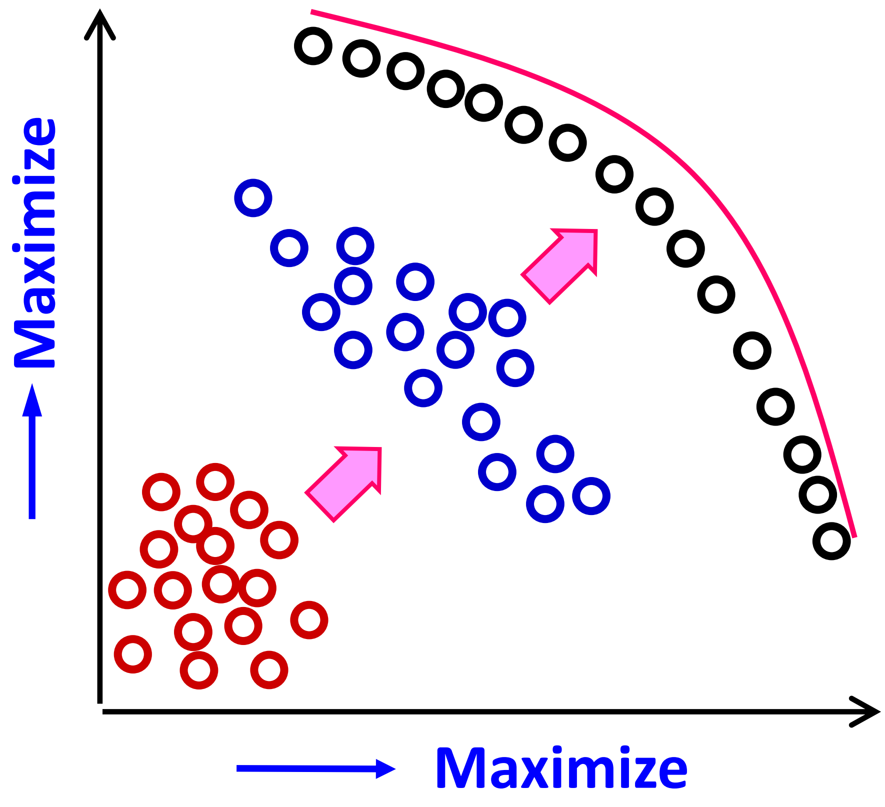
Solutions which have some very good objective values and other very bad objective values. Dominance resistant solutions are far away from the Pareto front. However, they are not likely to be dominated by other solutions due to some very good objective values.



Difficulties of NSGA-II

- (+)-style framework
- Projection-based crowding distance calculation **(Uniformity?)**
- Pareto dominance-based main fitness criterion **(Convergence?)**

These two difficulties are not clearly observed when we are solving only two-objective problems.

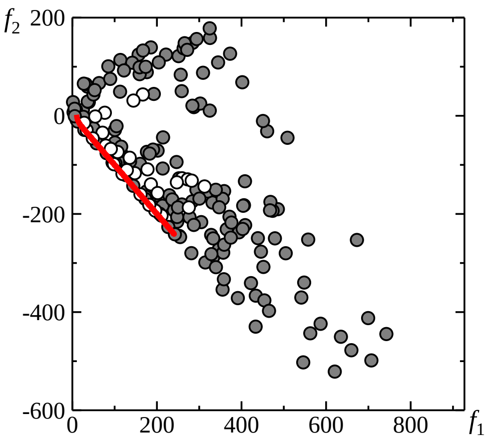


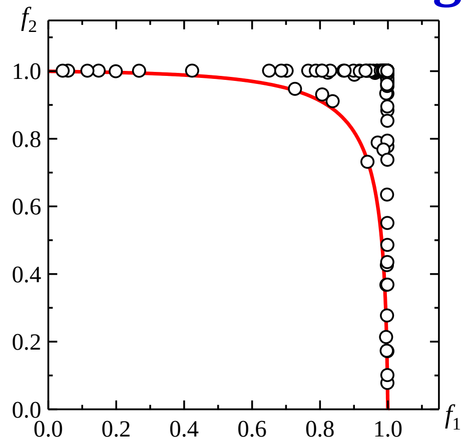
Relation between Test Problems and EMO Algorithms

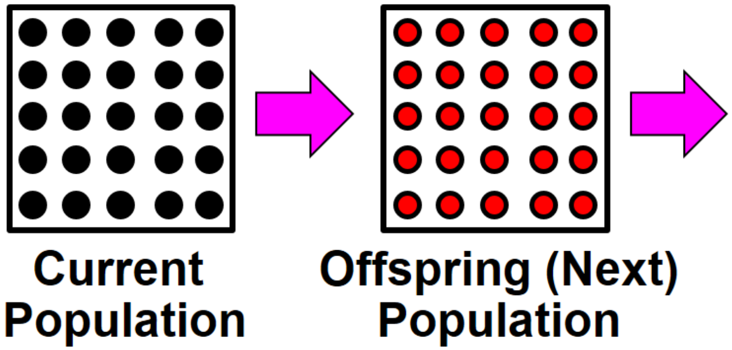
Around 1995

Simple two-objective problems (no need of strong convergence)

Non-elitist Pareto dominance-based algorithms



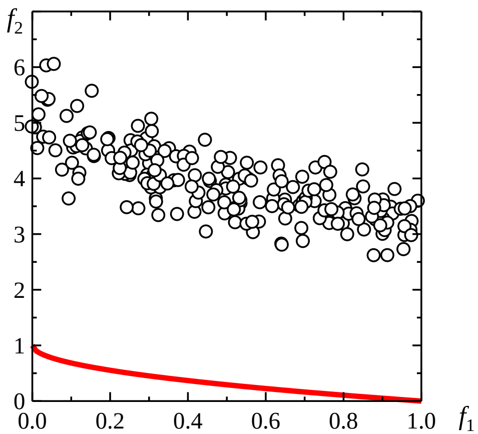


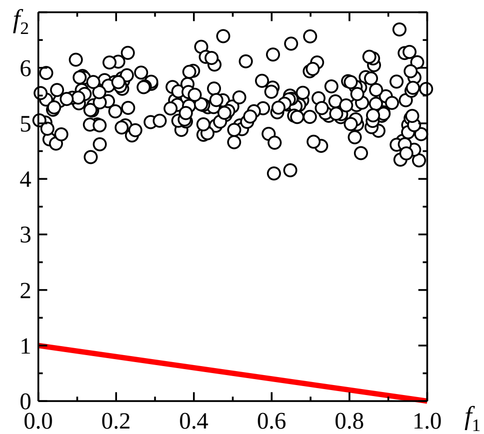


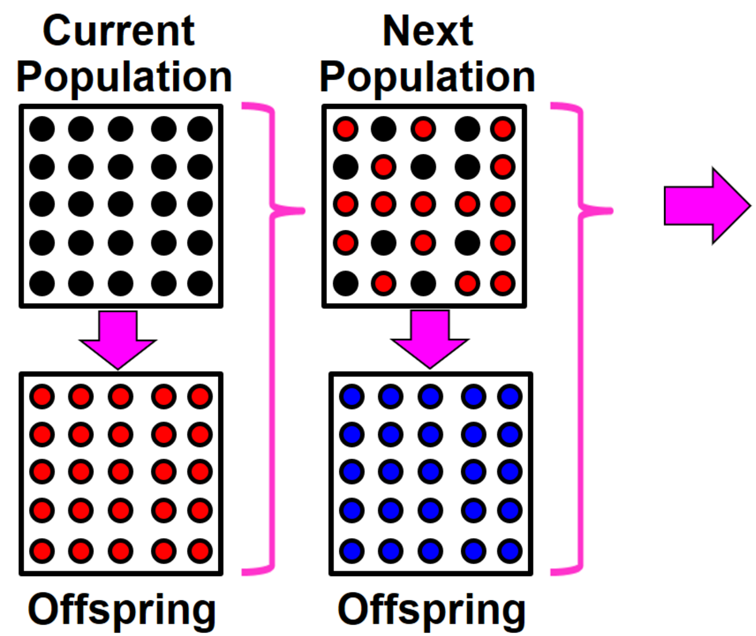
Around 2000

Difficult two-objective problems

Elitist Pareto dominance-based algorithms







Relation between Test Problems and EMO Algorithms

Around 1995

Simple two-objective problems (no need of strong convergence)

Non-elitist Pareto dominance-based algorithms: MOGA, NPGA, NSGA

These algorithms work well on simple two-objective problems

Around 2000

Difficult two-objective problems: ZDT

Elitist Pareto dominance-based algorithms: SPEA, SPEA2, NSGA-II

These algorithms work well on two-objective problems

Relation between Test Problems and EMO Algorithms

Around 1995

Simple two-objective problems (no need of strong convergence)

Non-elitist Pareto dominance-based algorithms: MOGA, NPGA, NSGA

These algorithms work well on simple two-objective problems

Around 2000

Difficult two-objective problems: ZDT

Elitist Pareto dominance-based algorithms: SPEA, SPEA2, NSGA-II

These algorithms work well on two-objective problems

2000s

Easy scalable test problems (the number of objectives can be arbitrarily specified: *m*-objective problems): DTLZ and WFG

DTLZ Test Problems (2002)
WFG Test Problems (2006)

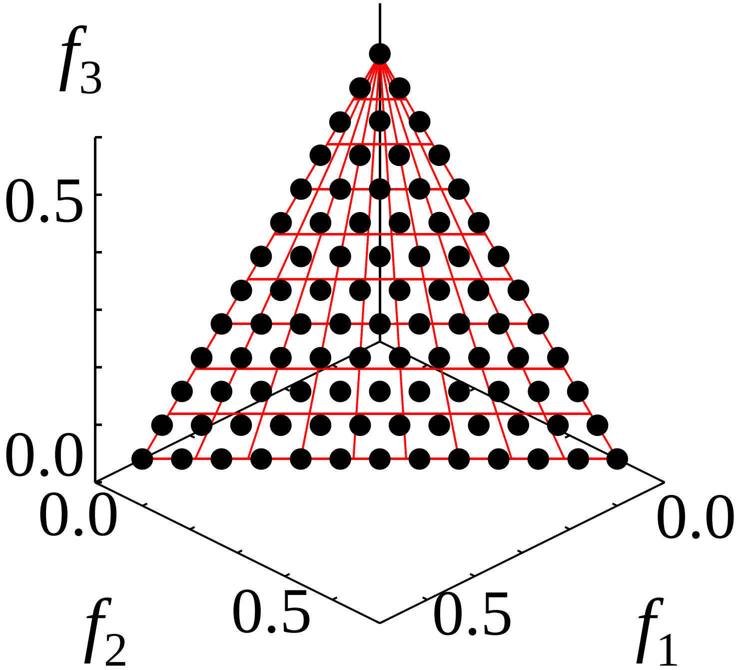
Scalable Test Problems

The number of objectives can be arbitrarily specified.

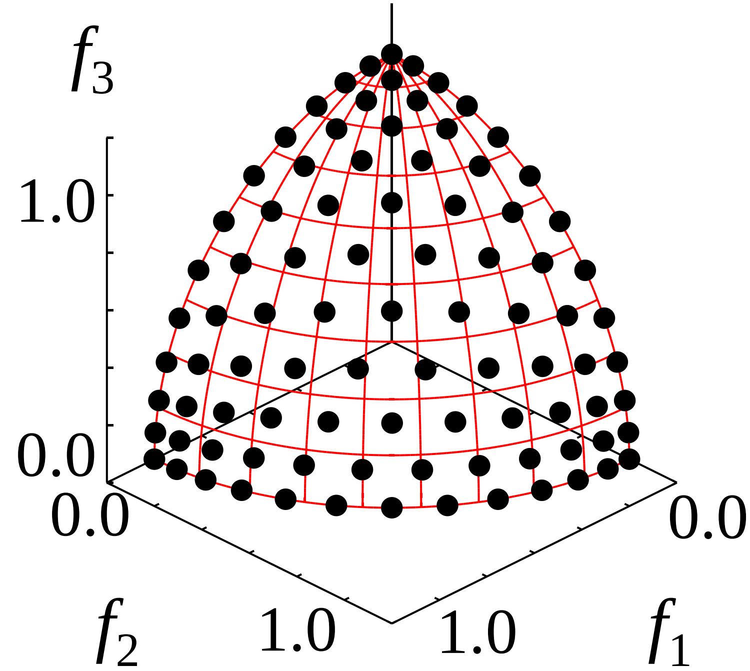
DTLZ K. Deb, L. Thiele, M. Laumanns, and E. Zitzler, "**Scalable multi-objective optimization test problems**," IEEE CEC 2002.

WFG S. Huband, P. Hingston, L. Barone, and L. While, "**A review of multiobjective test problems and a scalable test problem toolkit**," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 5, pp. 477-506, 2006.

Pareto Fronts of Test Problems with Three Objectives



DTLZ 1

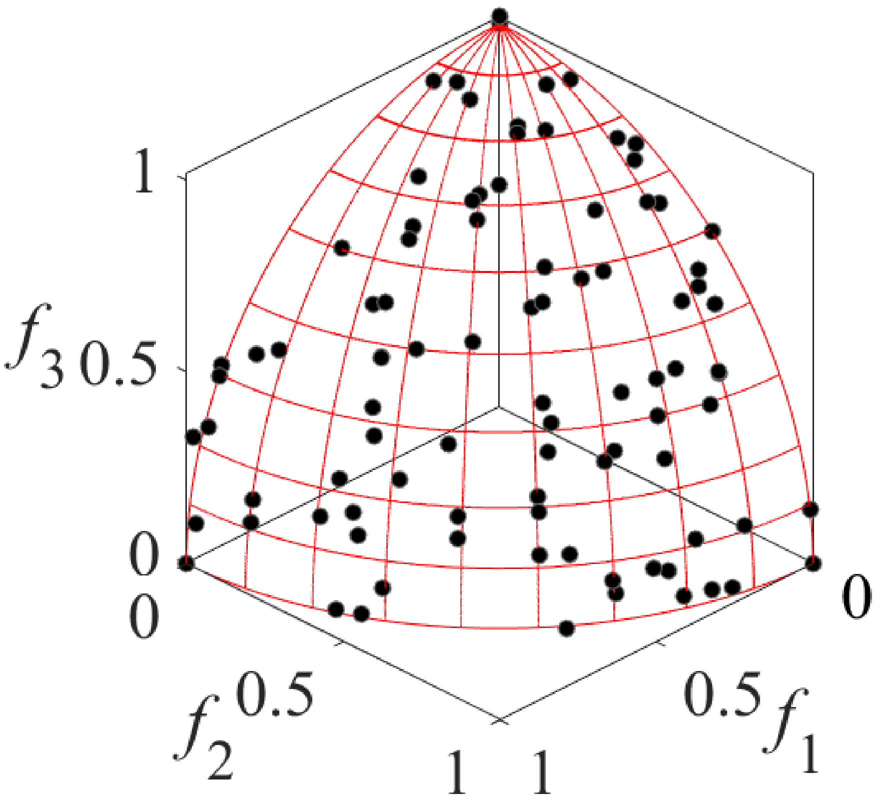


DTLZ 2, DTLZ 3, DTLZ 4
Normalized WFG 4-9.

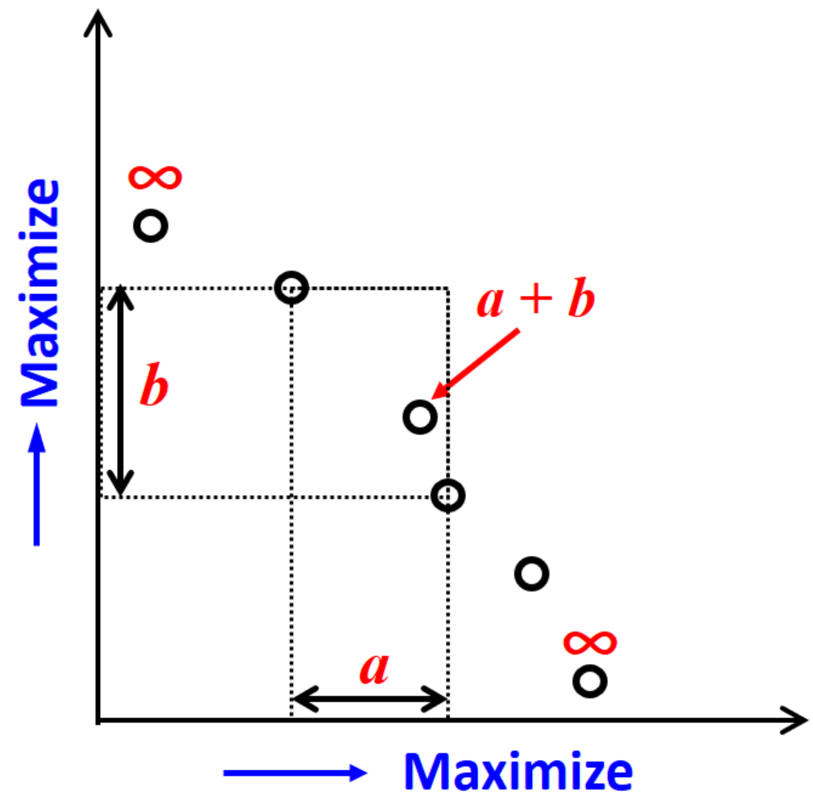
Important Issue:

To find a set of well-distributed solutions over the entire Pareto front.

Results of NSGA-II



Final population (obtained solutions).



Crowding distance calculation.

It is difficult to find a well-distributed solution set for multi-objective problems with three or more objectives.

Relation between Test Problems and EMO Algorithms

Around 1995

Simple two-objective problems (no need of strong convergence)

Non-elitist Pareto dominance-based algorithms: MOGA, NPGA, NSGA

These algorithms work well on simple two-objective problems

Around 2000

Difficult two-objective problems: ZDT

Elitist Pareto dominance-based algorithms: SPEA, SPEA2, NSGA-II

These algorithms work well on two-objective problems

2000s

Easy scalable test problems (the number of objectives can be arbitrarily specified: *m*-objective problems): DTLZ and WFG

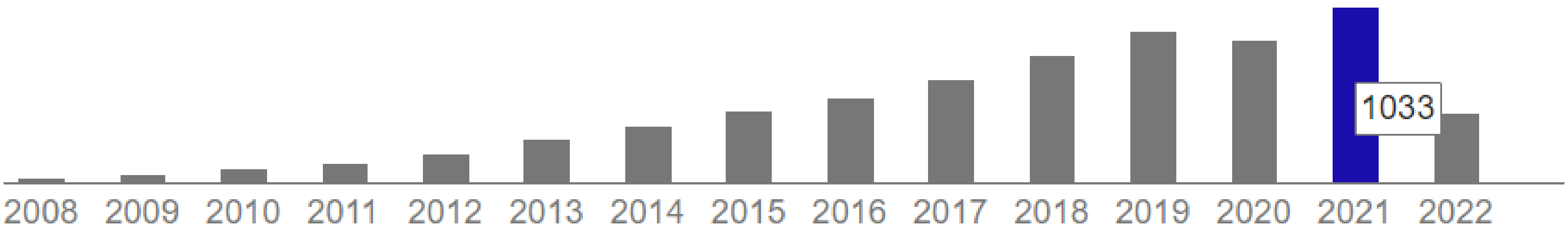
Decomposition-based algorithms: MOEA/D

Indicator-based algorithms: SMS-EMOA

MOEA/D (IEEE TEVC 2007)

Multi-Objective Evolutionary Algorithm based on Decomposition

Cited by 6449



MOEA/D: A multiobjective evolutionary algorithm based on decomposition
Q Zhang, H Li - IEEE Transactions on evolutionary computation, 2007
Cited by 6449 Related articles All 17 versions

Qingfu Zhang

Chair Professor, FIEEE, City University of Hong Kong
Verified email at cityu.edu.hk - Homepage

evolutionary computation multiobjective optimization

Decomposition-Based EMO Algorithm

Basic Idea:

A single multi-objective problem

==> A number of single-objective problems

Minimization of d_1 and d_2

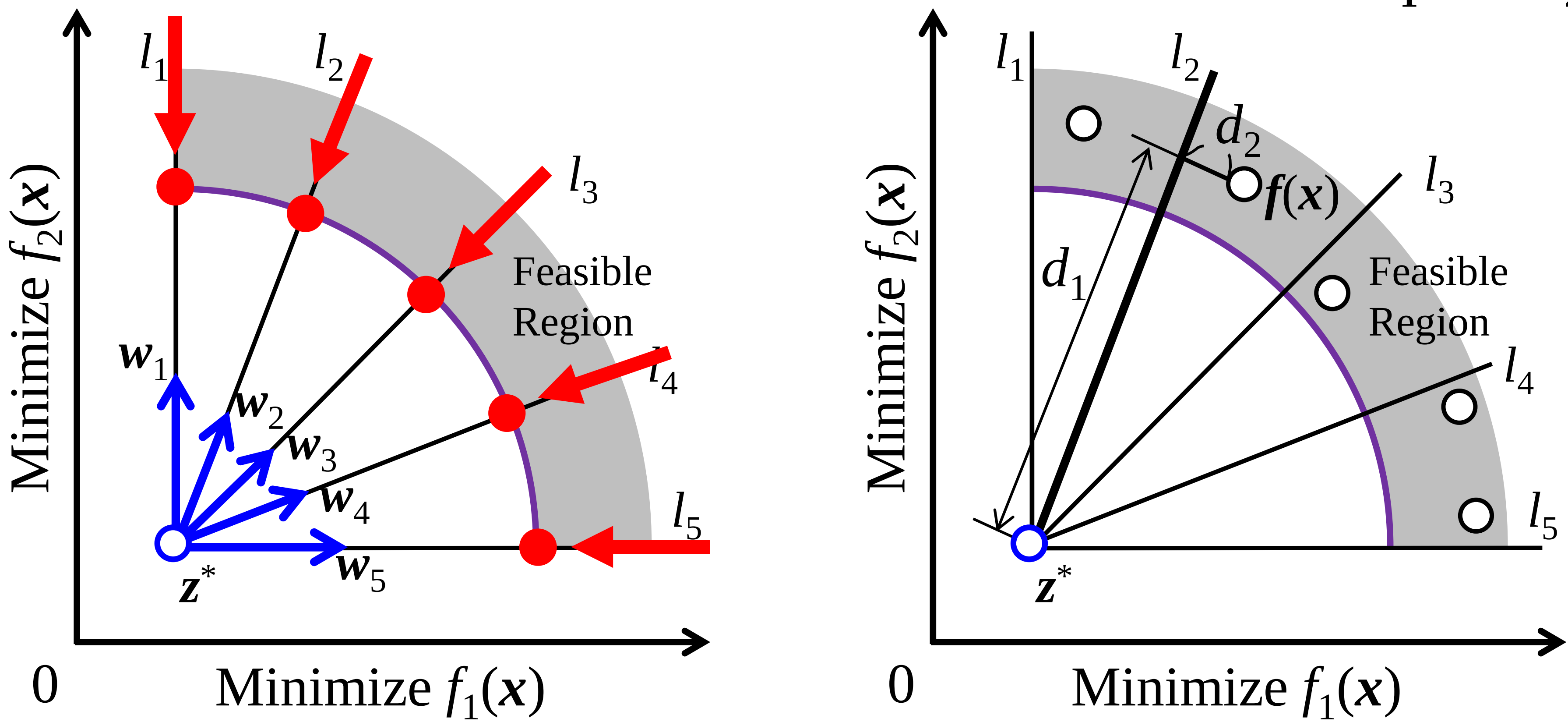
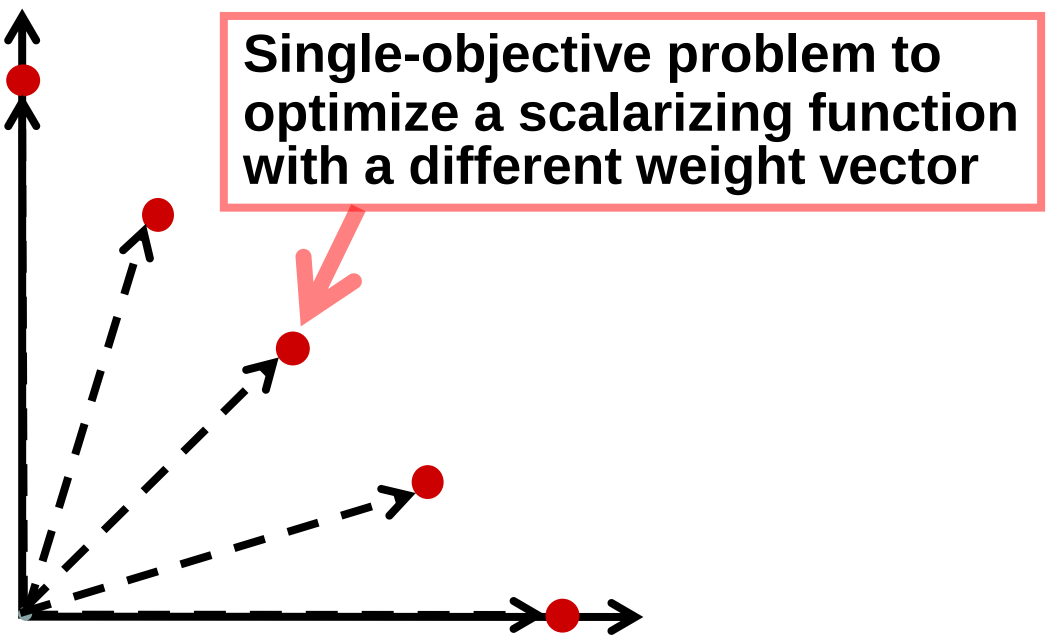


Illustration for a minimization problem.

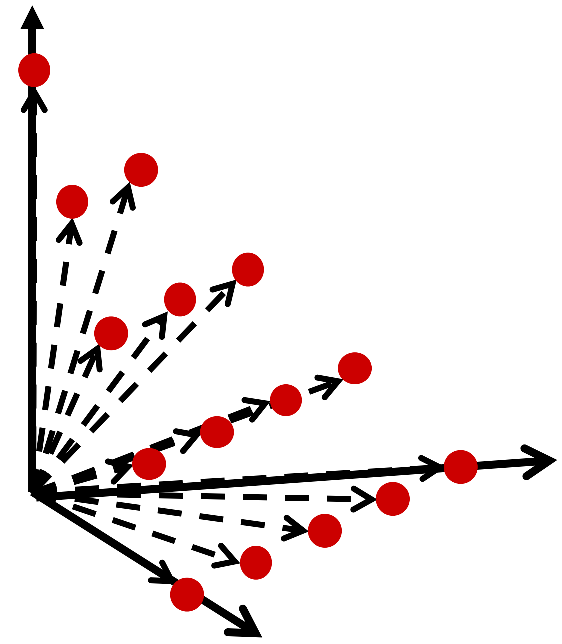
Basic Idea of MOEA/D

Q. Zhang and H. Li (IEEE TEVC 2007)

Decomposition: A multi-objective problem is handled as a set of scalarizing function optimization problems with different weight vectors. Uniformly distributed weight vectors were used in the original study (2007).



Weight vectors (2-objective case)



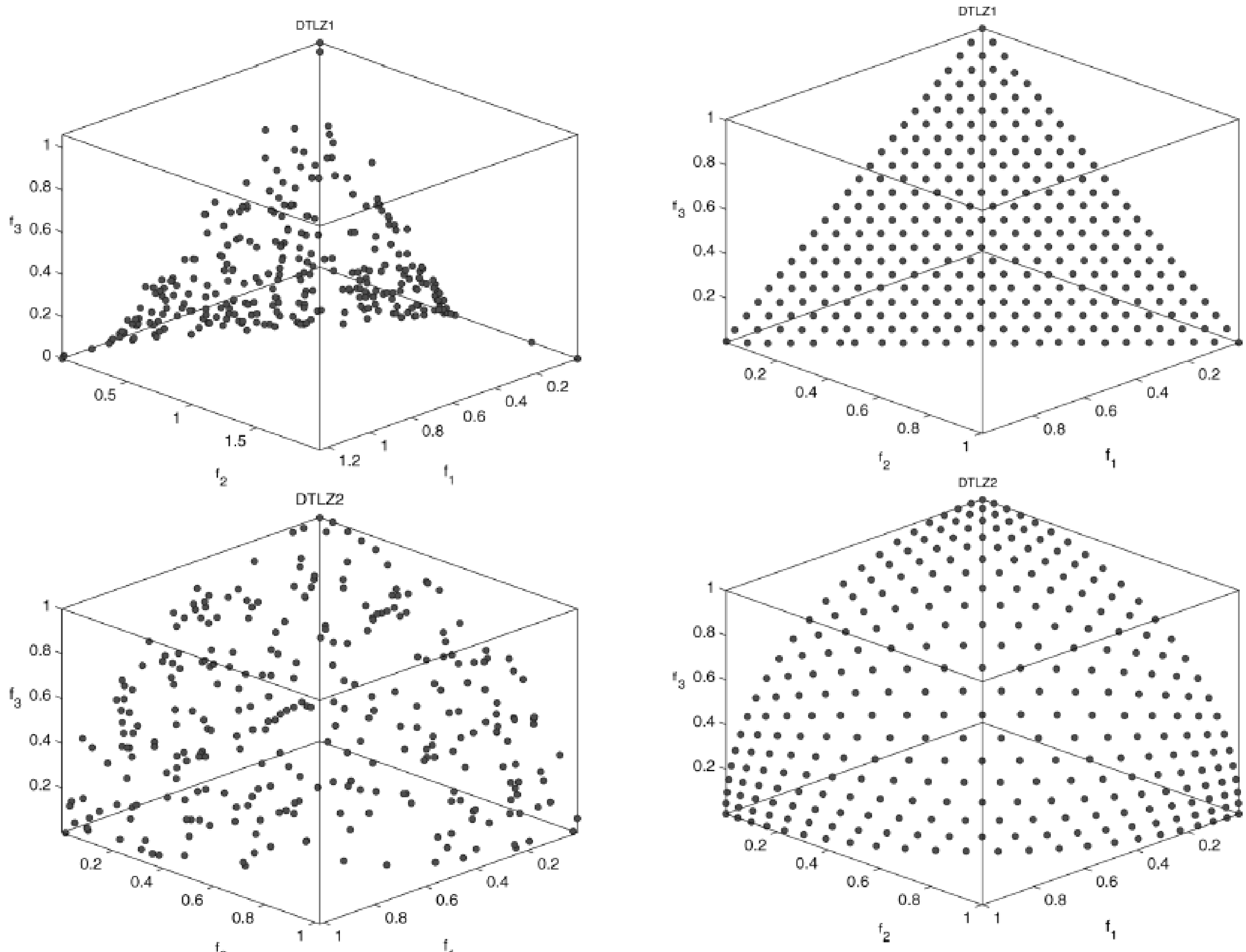
Weight vectors (3-objective case)

MOEA/D (IEEE TEVC 2007)

Multi-Objective Evolutionary Algorithm based on Decomposition

DTLZ1

DTLZ2



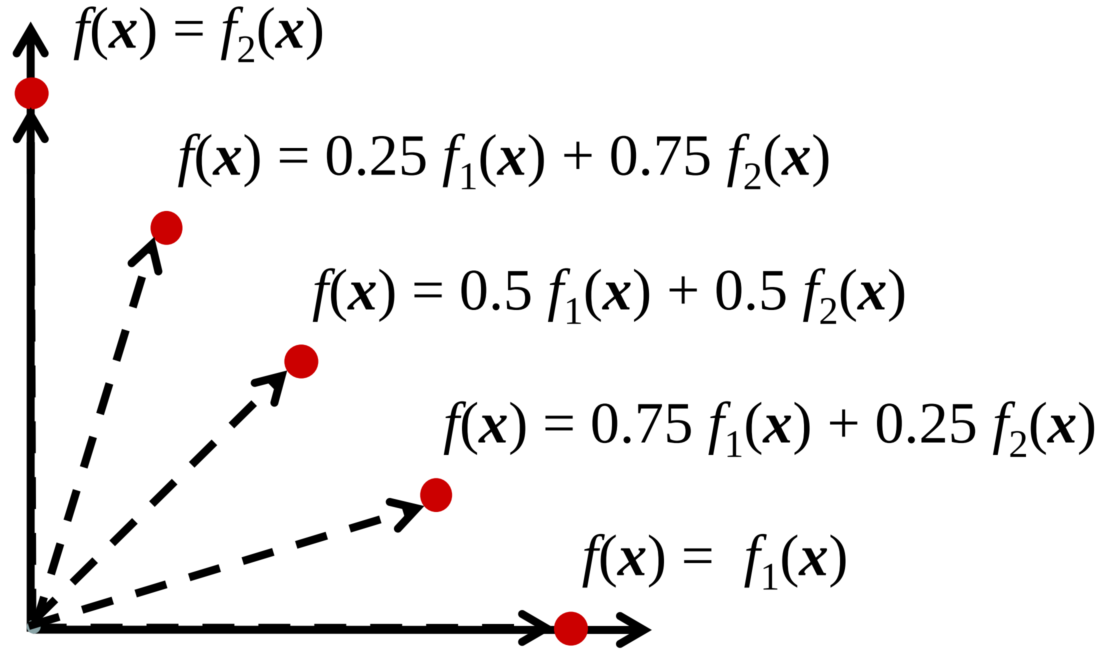
NSGA-II

MOEA/D

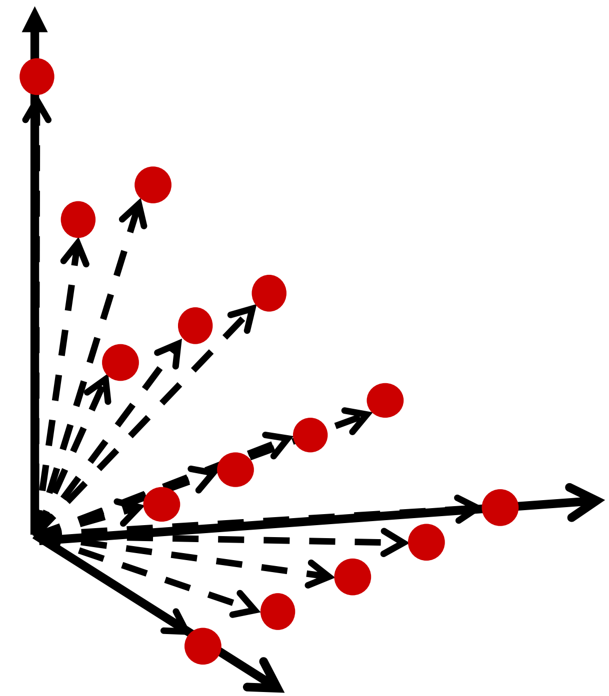
Basic Idea of MOEA/D

Q. Zhang and H. Li (IEEE TEVC 2007)

Decomposition: A multi-objective problem is handled as a set of scalarizing function optimization problems with different weight vectors. Uniformly distributed weight vectors were used in the original study (2007).



Weighted Sum



Scalarizing Functions in MOEA/D

Q. Zhang and H. Li (IEEE TEVC 2007)

In the original study of MOEA/D (IEEE TEVC 2007), the following three functions were examined.

Weight vector: $\boldsymbol{w} = (w_1, w_2, \ldots, w_m)$

1. Weighted Sum:

$$g^{WS} = \sum_{i=1}^{m} w_i \cdot f_i(\boldsymbol{x})$$

2. Weighted Tchebycheff:

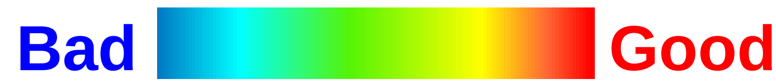
$$g^{TE} = \max_{i=1,2,\ldots,m} \left\{ w_i \cdot | z_i^* - f_i(\boldsymbol{x}) | \right\}$$

3. PBI (Penalty-based Boundary Intersection):

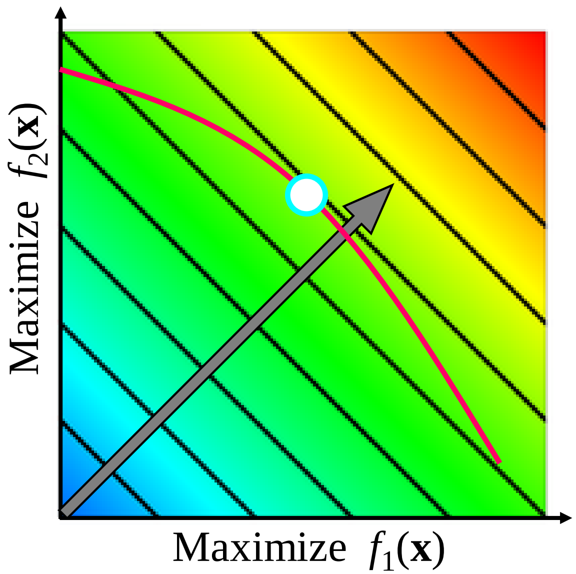
Scalarizing Functions in MOEA/D

Their Contour Lines

Contour line of each scalarizing function for the weighted vector λ = (0.5, 0.5)

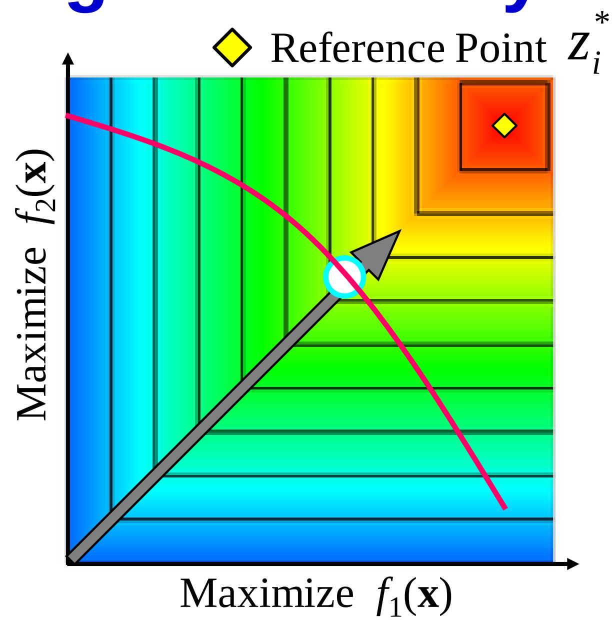


Weighted Sum



$$g^{WS} = \sum_{i=1}^{m} w_i \cdot f_i(\boldsymbol{x})$$

Weighted Tchebycheff



$$g^{TE} = \max_{i=1,2,\ldots,m} \left\{ w_i \cdot | z_i^* - f_i(\boldsymbol{x}) | \right\}$$

***m*: Number of objectives**

Scalarizing Functions in MOEA/D

Q. Zhang and H. Li (IEEE TEVC 2007)

In the original study of MOEA/D (IEEE TEVC 2007), the following three functions were examined.

3. PBI (Penalty-based Boundary Intersection):

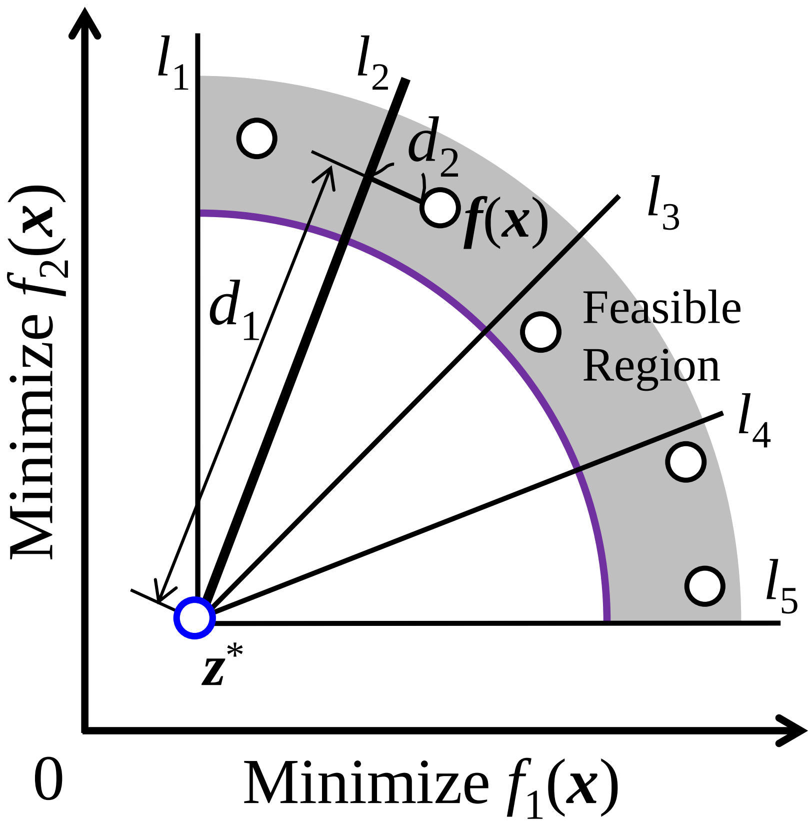
$$g^{PBI} = d_1 + \theta \cdot d_2$$

$$d_1 = \left\| (\boldsymbol{z}^* - \boldsymbol{f}(\boldsymbol{x}))^T \boldsymbol{w} \right\| / \|\boldsymbol{w}\|$$

$$d_2 = \left\| \boldsymbol{f}(\boldsymbol{x}) - (\boldsymbol{z}^* - d_1 \boldsymbol{w} / \|\boldsymbol{w}\|) \right\|$$

Parameter Setting in PBI:

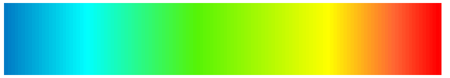
$$\theta = 5$$



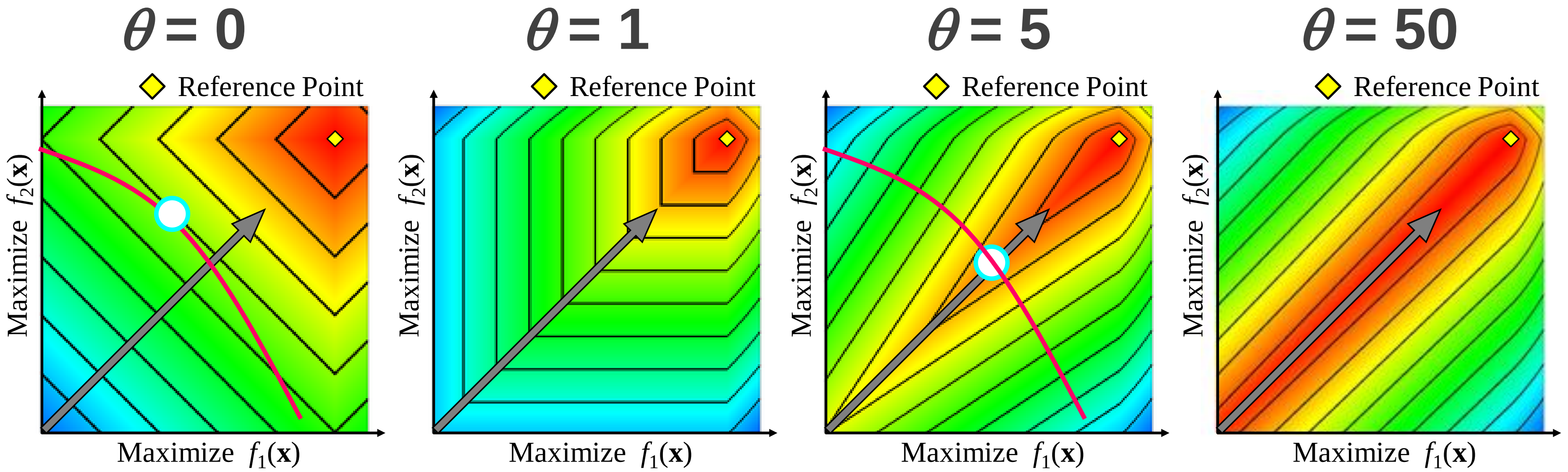
Scalarizing Functions in MOEA/D

Their Contour Lines

Contour line of each scalarizing function for the weighted vector *w* = (0.5, 0.5)

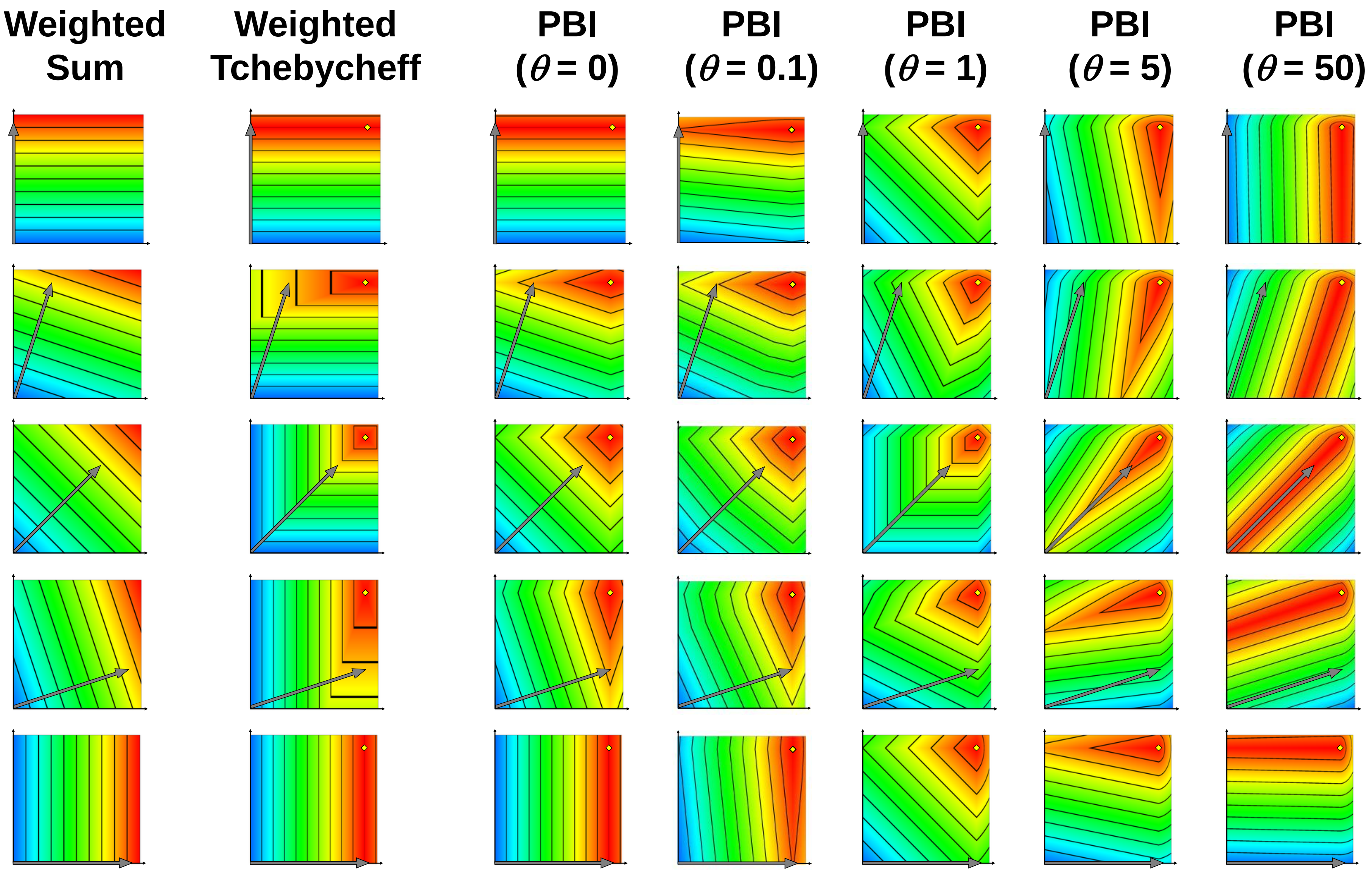


PBI (Penalty-based Boundary Intersection)



$$g^{PBI} = d_1 + \theta \cdot d_2$$

Contour Lines of Scalarizing Functions for Various Weight Vectors

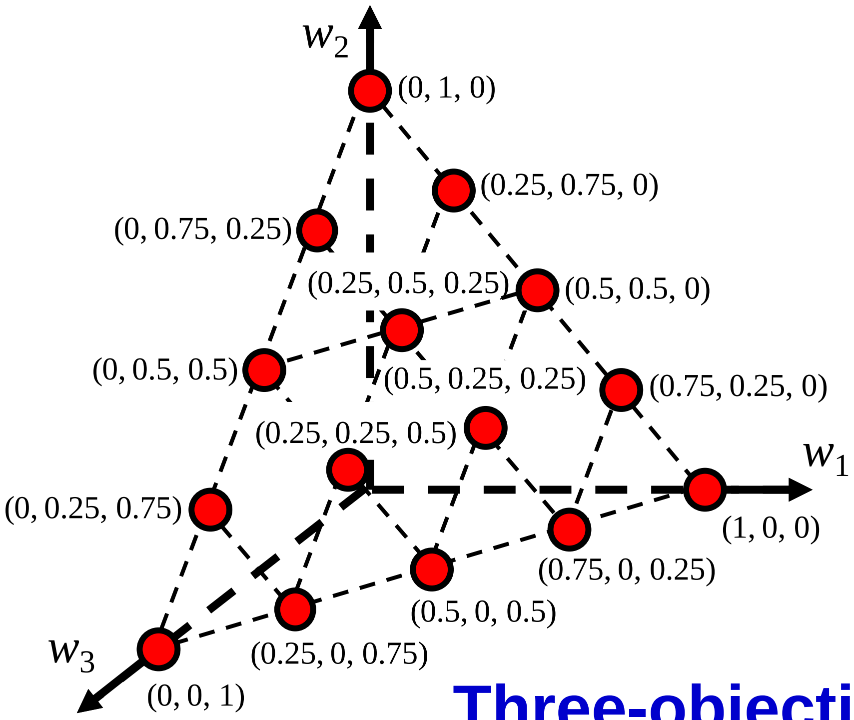


Uniformly Distributed Weight Vectors

$$w_1 + w_2 + \ \dots \ + w_m = 1$$ **(for an *m*-objective problem)**

$$w_i \in \{0/H, 1/H, \ \dots, \ H/H\}, \quad i = 1, 2, \dots, m$$

(*H*: Integer)



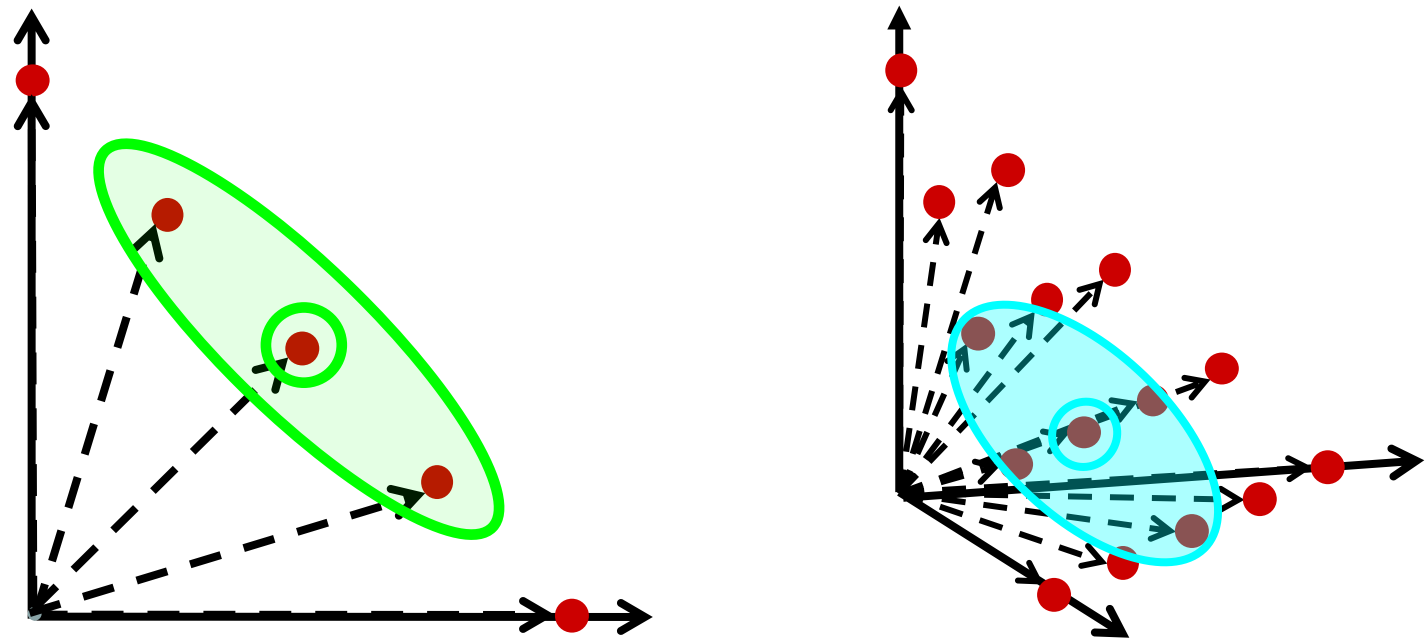
Three-objective case *H* = 4

Mechanisms in MOEA/D

Local Selection and Local Comparison

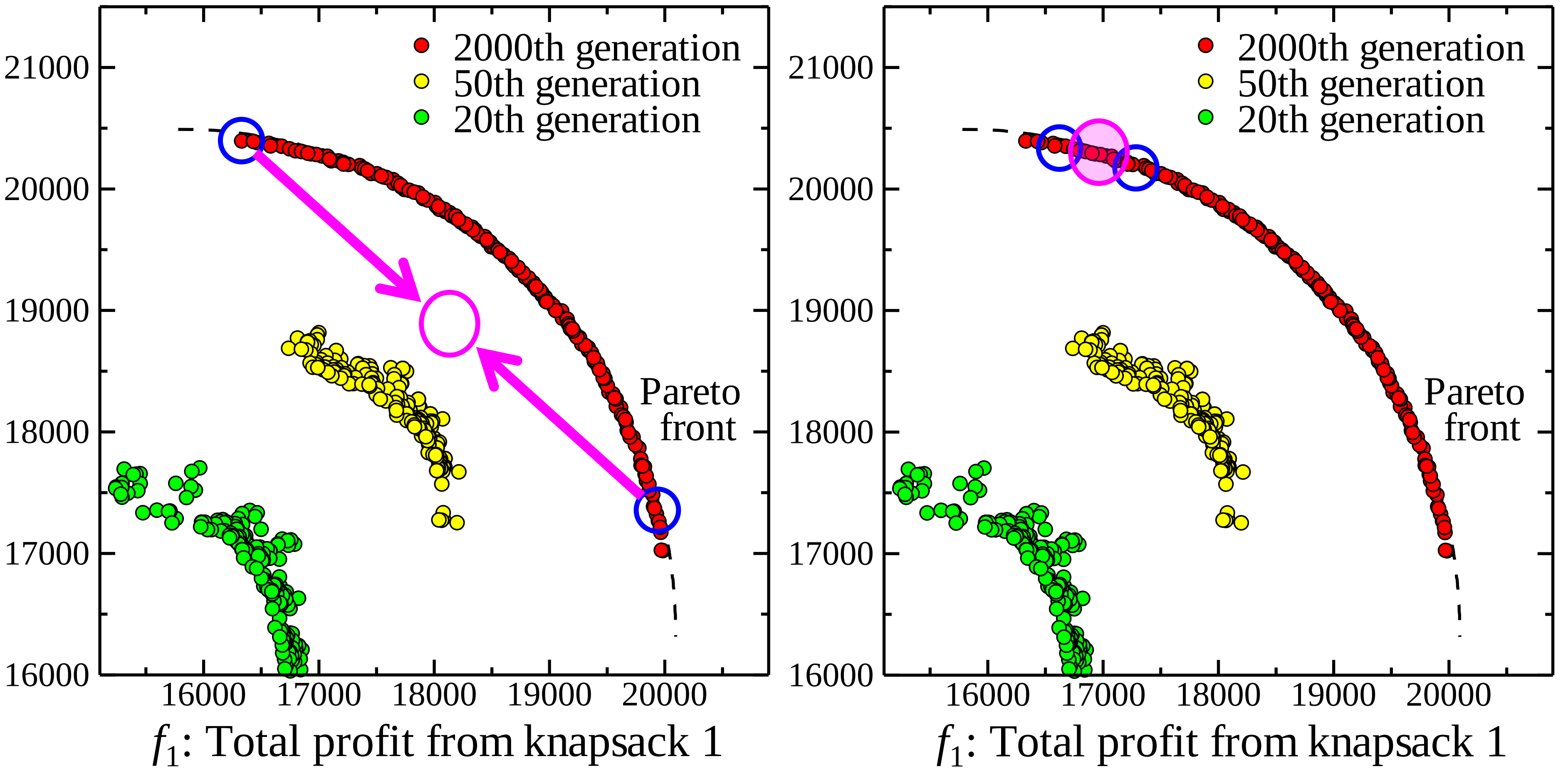
1. Local Selection: Choice of Similar Parents

Neighbors are defined for each weight vector by the distance between weight vectors. Local selection as in cellular algorithms is used for parent selection.



(a) Two-objective case (b) Three-objective case

Recombination of Similar Parents



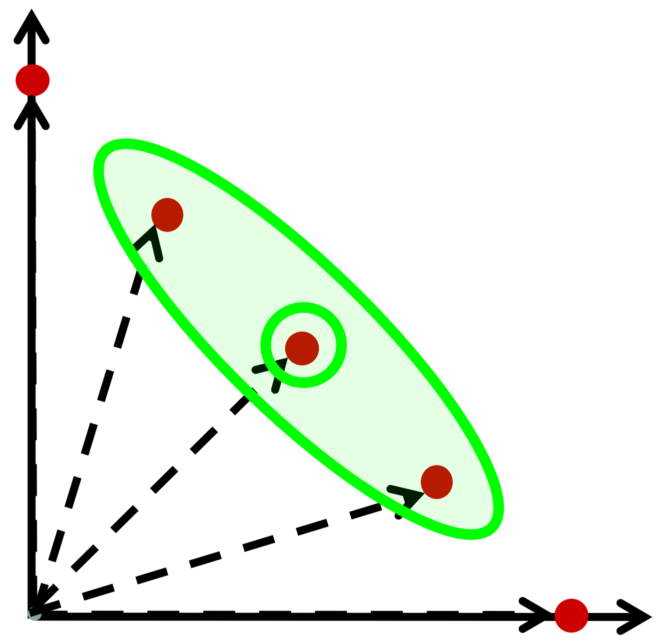
Different Parents **Similar Parents**

Mechanisms in MOEA/D

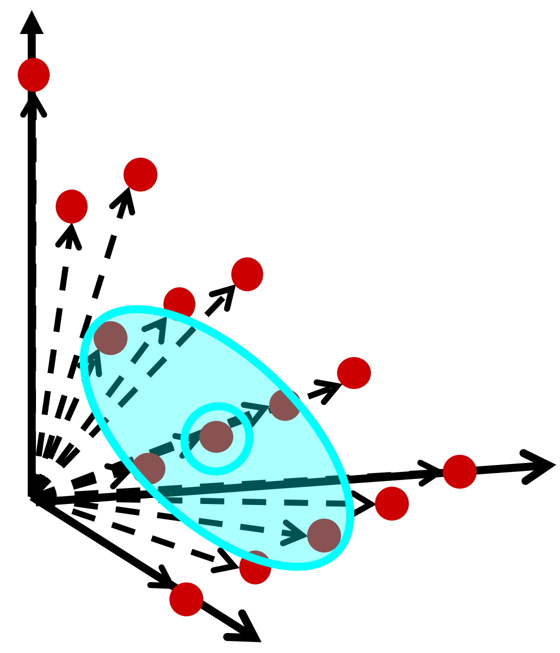
Local Selection and Local Comparison

2. Local Comparison: Multiple Replacement

A newly generated offspring is compared with each of its neighbors. All inferior neighbors are replaced with the new offspring.



(a) Two-objective case



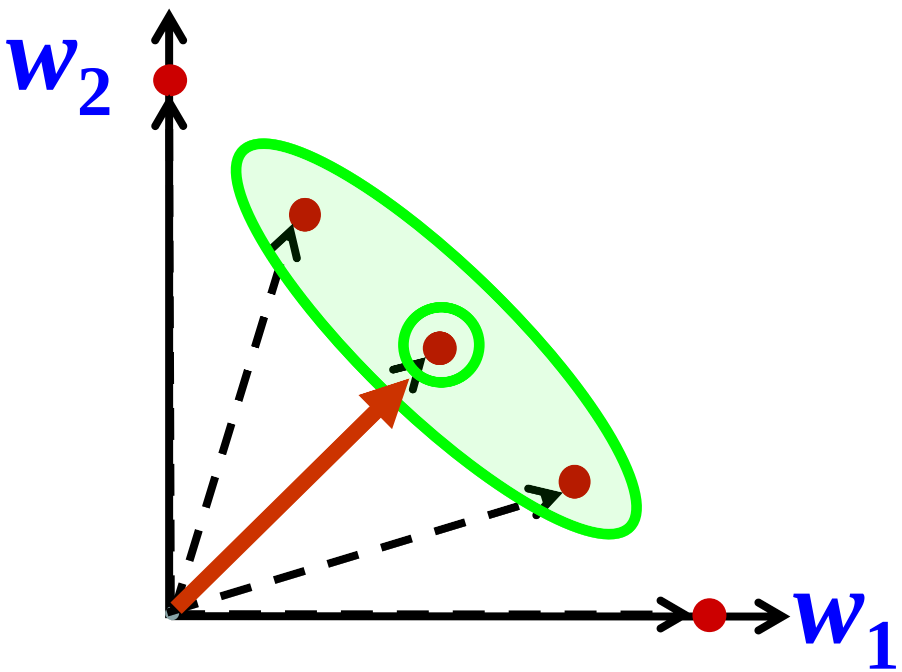
(b) Three-objective case

Mechanisms in MOEA/D

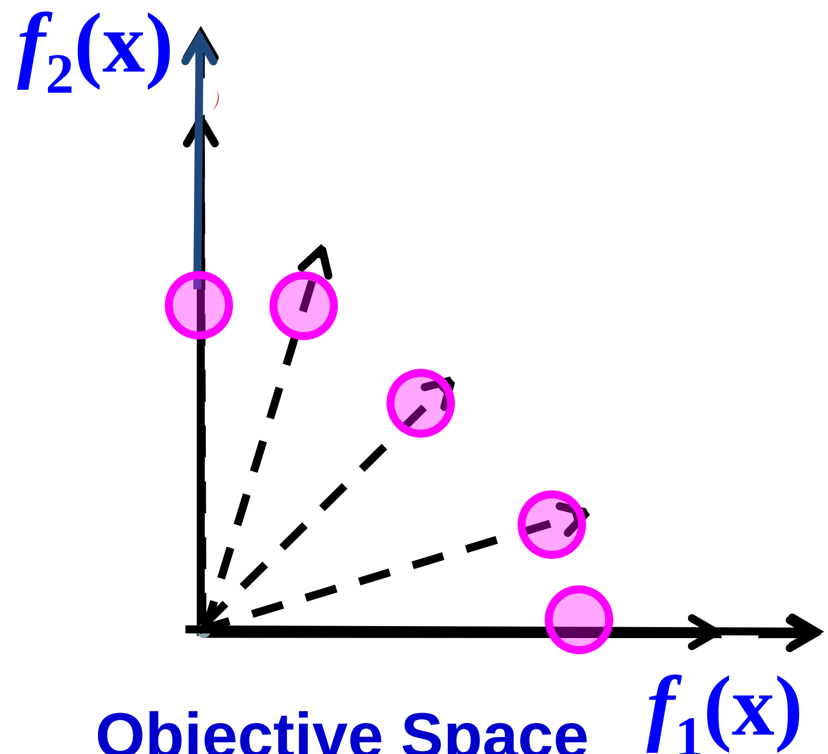
Local Selection and Local Comparison

2. Local Comparison: Multiple Replacement

A newly generated offspring is compared with each of its neighbors. All inferior neighbors are replaced with the new offspring.



Weight Vector Space



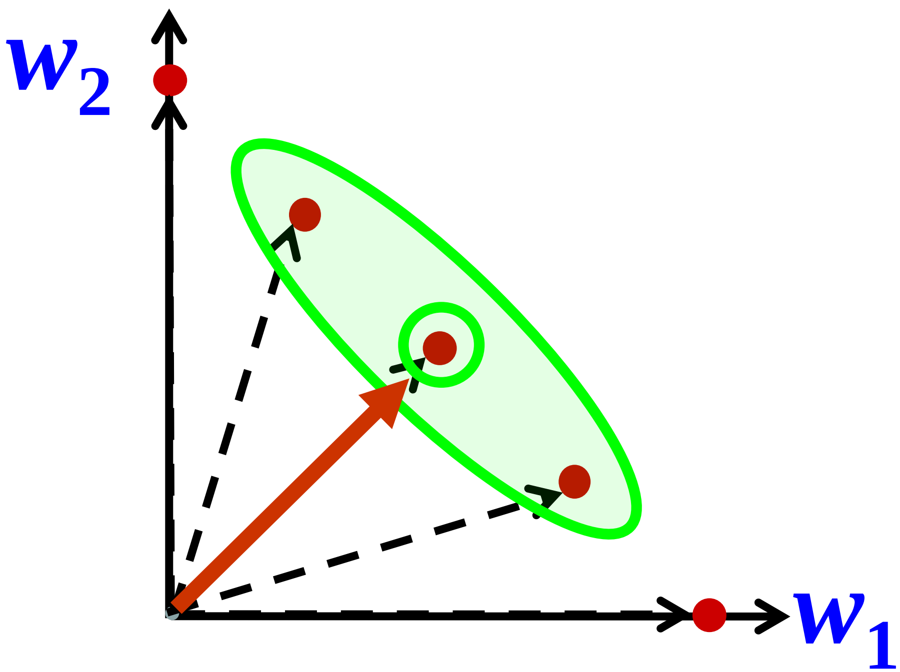
Objective Space

Mechanisms in MOEA/D

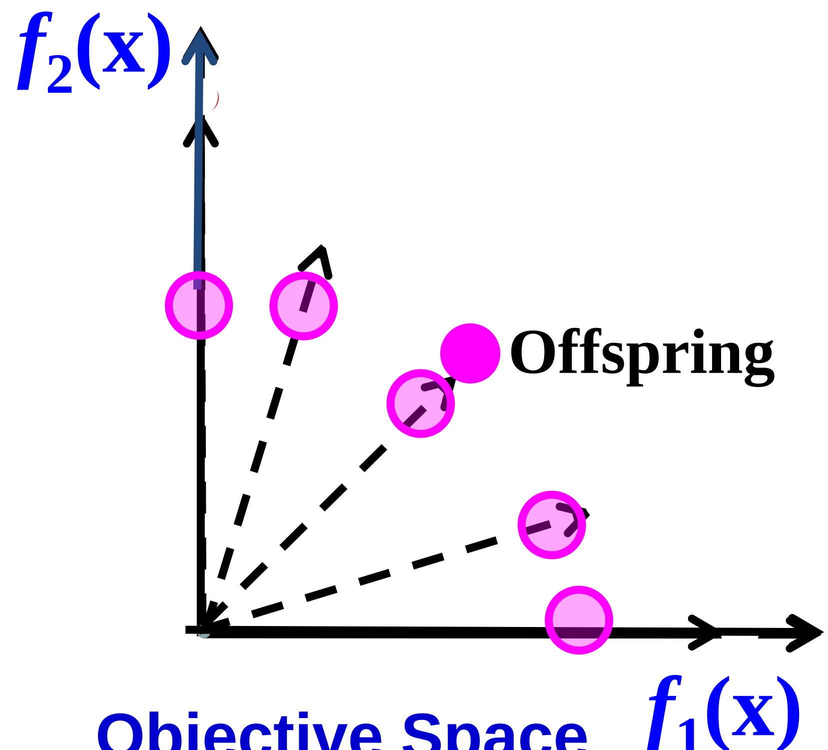
Local Selection and Local Comparison

2. Local Comparison: Multiple Replacement

A newly generated offspring is compared with each of its neighbors. All inferior neighbors are replaced with the new offspring.



Weight Vector Space



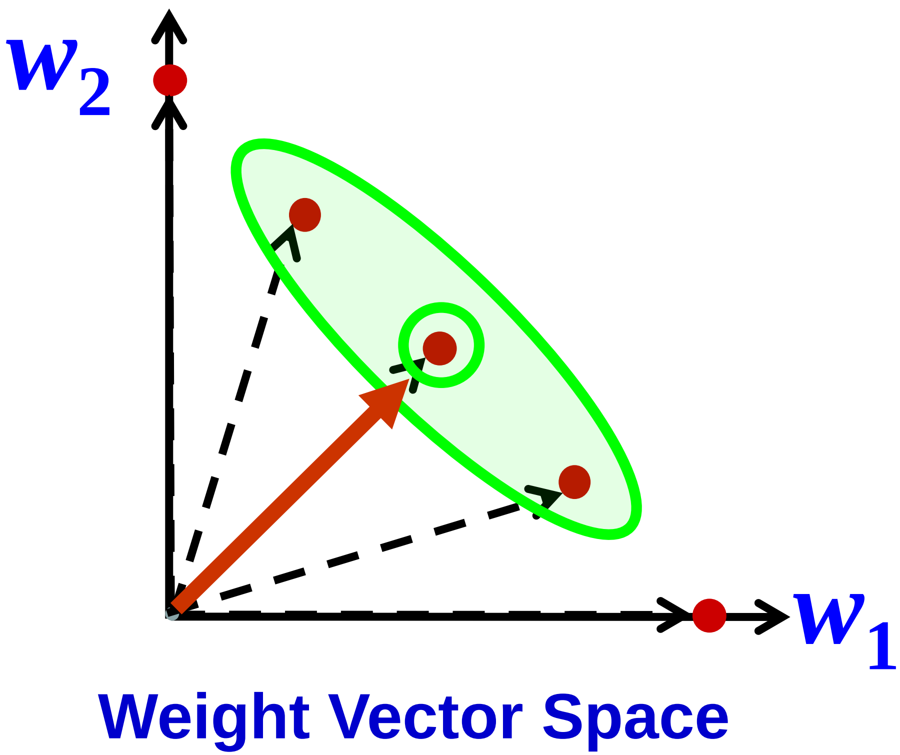
Objective Space

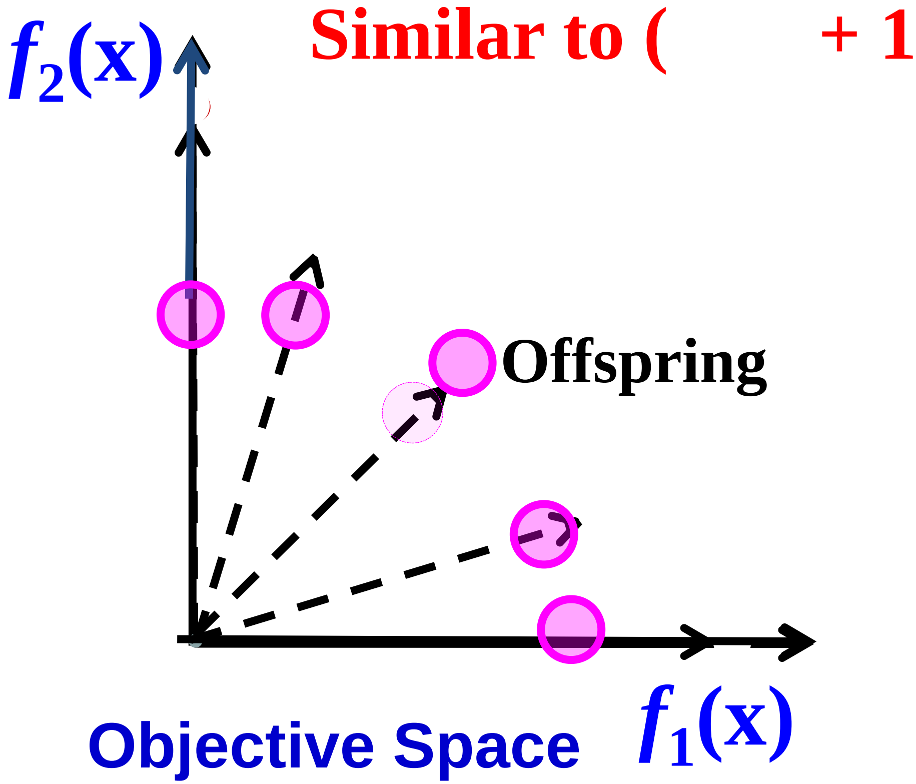
Mechanisms in MOEA/D

Local Selection and Local Comparison

2. Local Comparison: Multiple Replacement

A newly generated offspring is compared with each of its neighbors. All inferior neighbors are replaced with the new offspring.



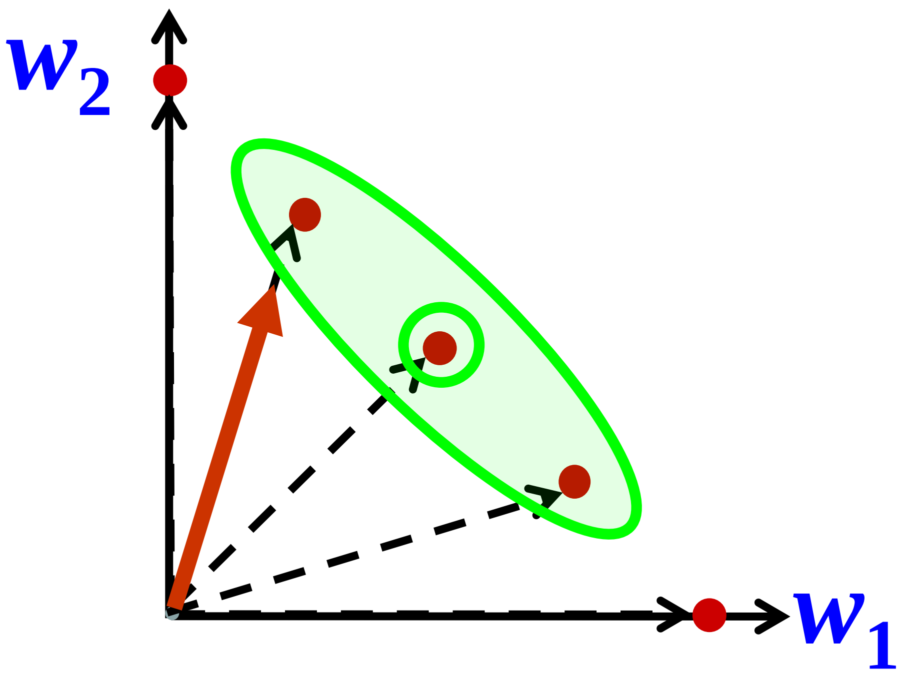


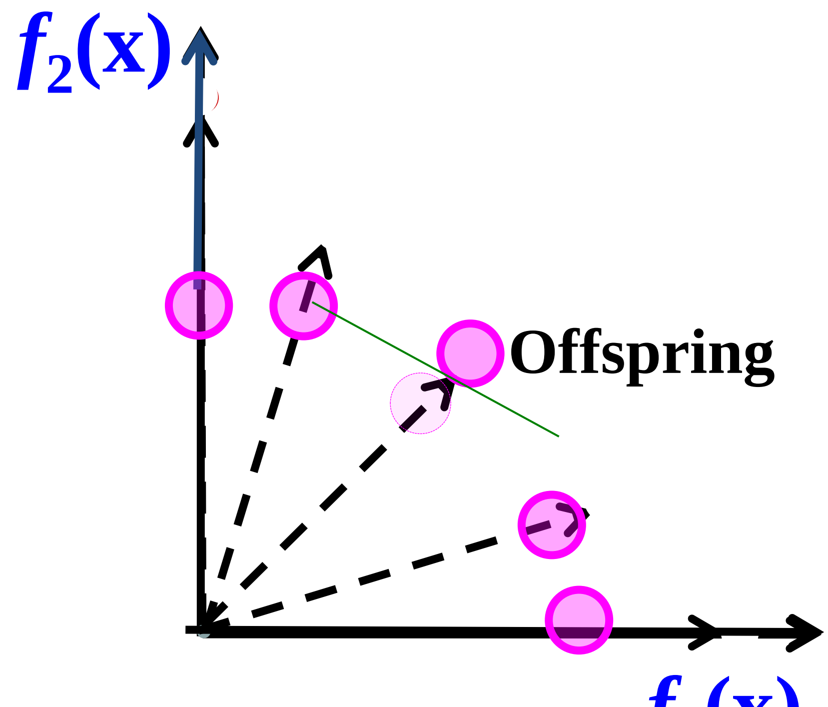
Mechanisms in MOEA/D

Local Selection and Local Comparison

2. Local Comparison: Multiple Replacement

A newly generated offspring is compared with each of its neighbors. All inferior neighbors are replaced with the new offspring.





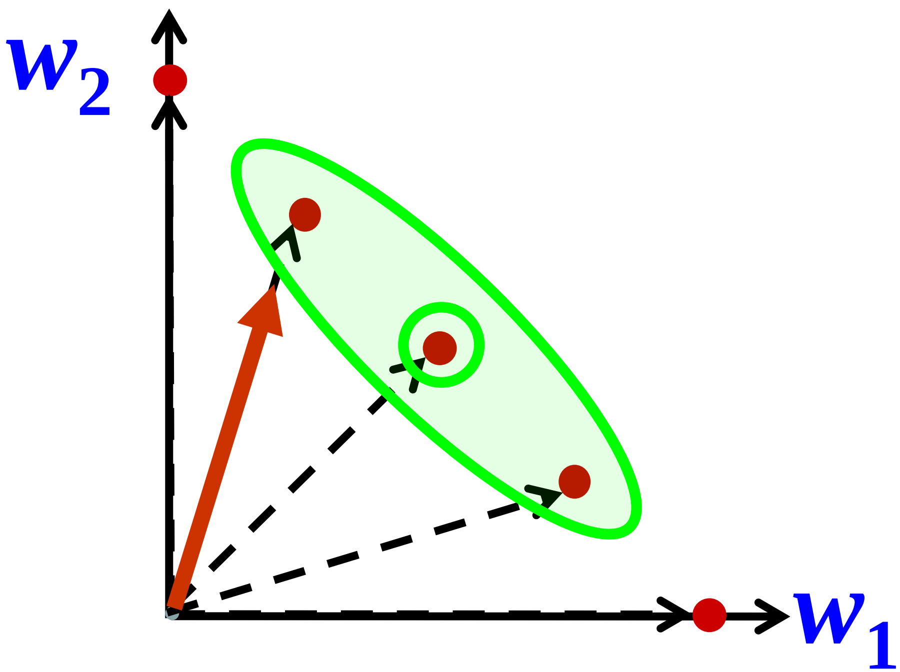
Mechanisms in MOEA/D

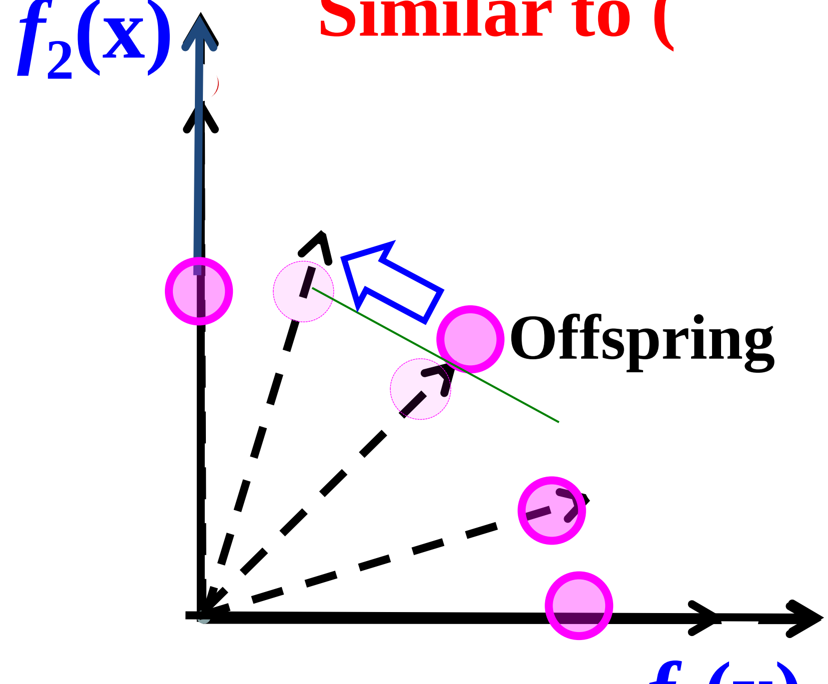
Local Selection and Local Comparison

2. Local Comparison: Multiple Replacement

A newly generated offspring is compared with each of its neighbors. All inferior neighbors are replaced with the new offspring.

Similar to (+ 1



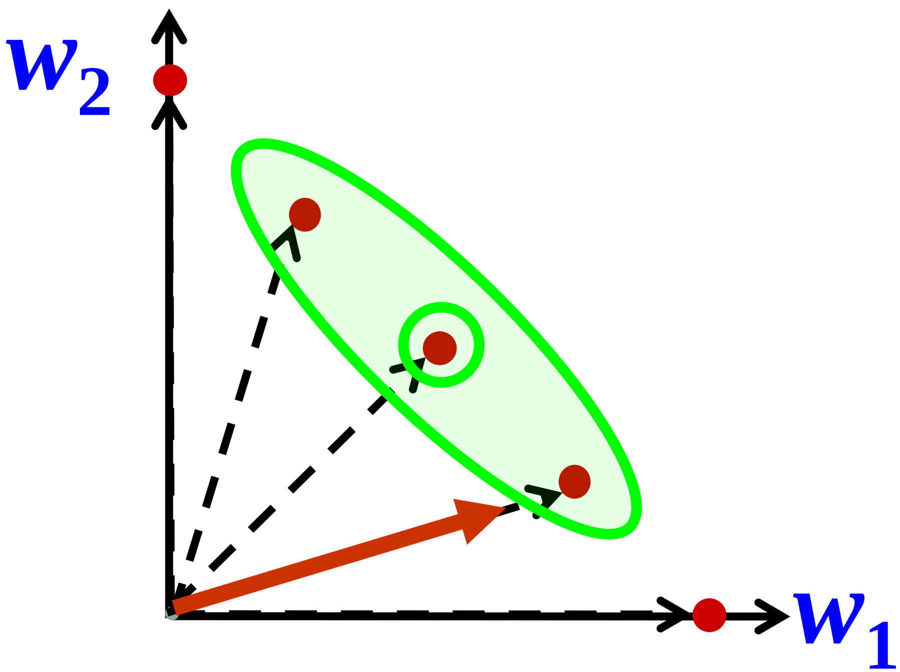


Mechanisms in MOEA/D

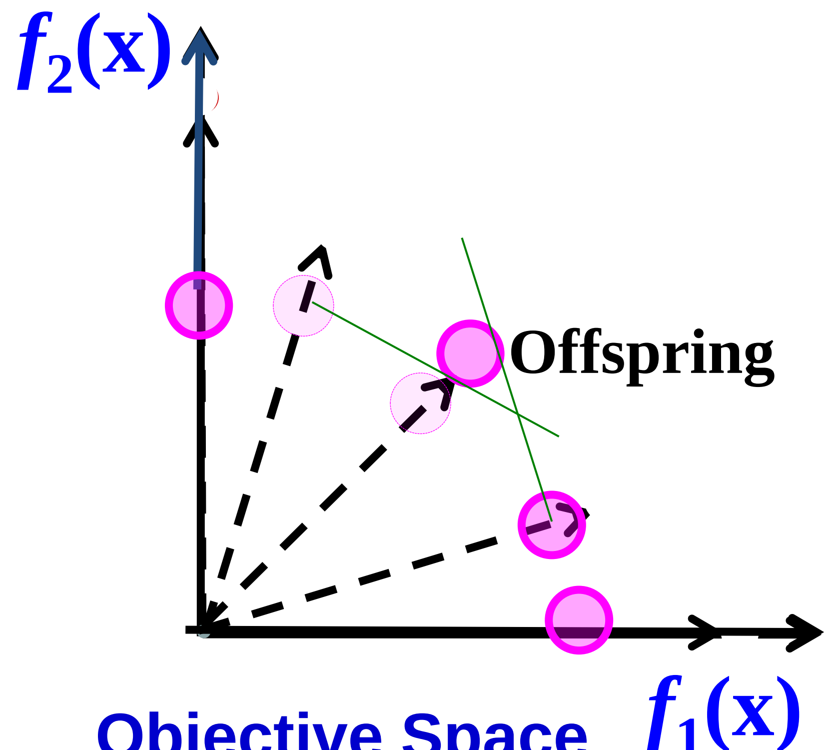
Local Selection and Local Comparison

2. Local Comparison: Multiple Replacement

A newly generated offspring is compared with each of its neighbors. All inferior neighbors are replaced with the new offspring.



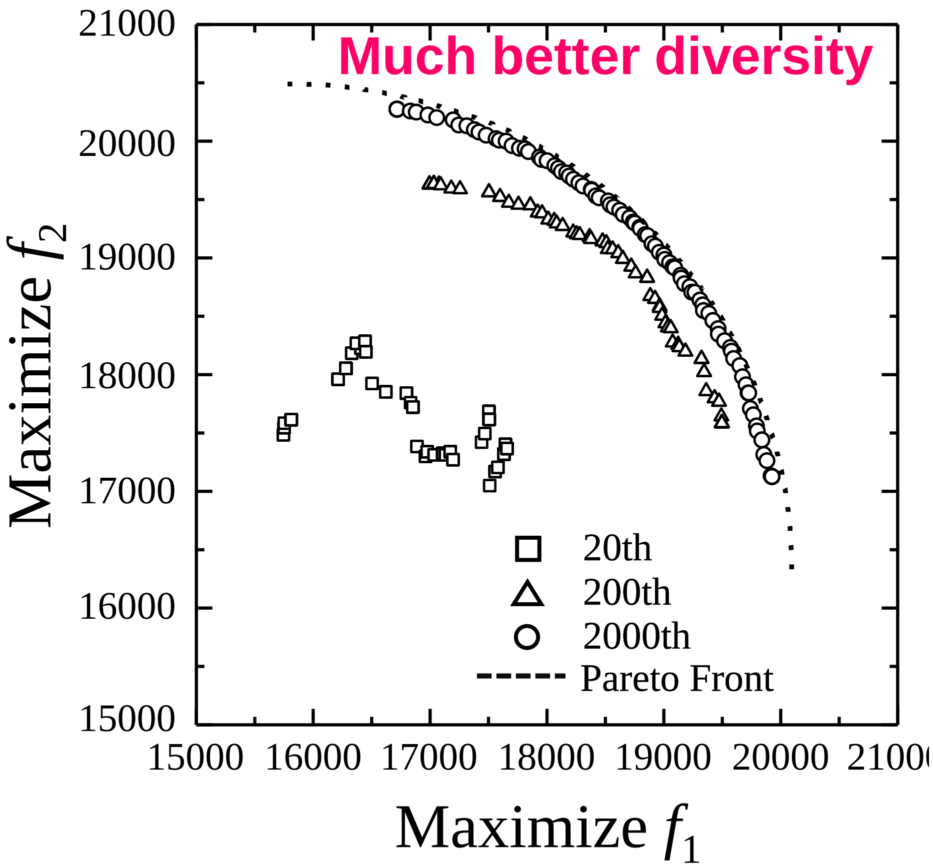
Weight Vector Space



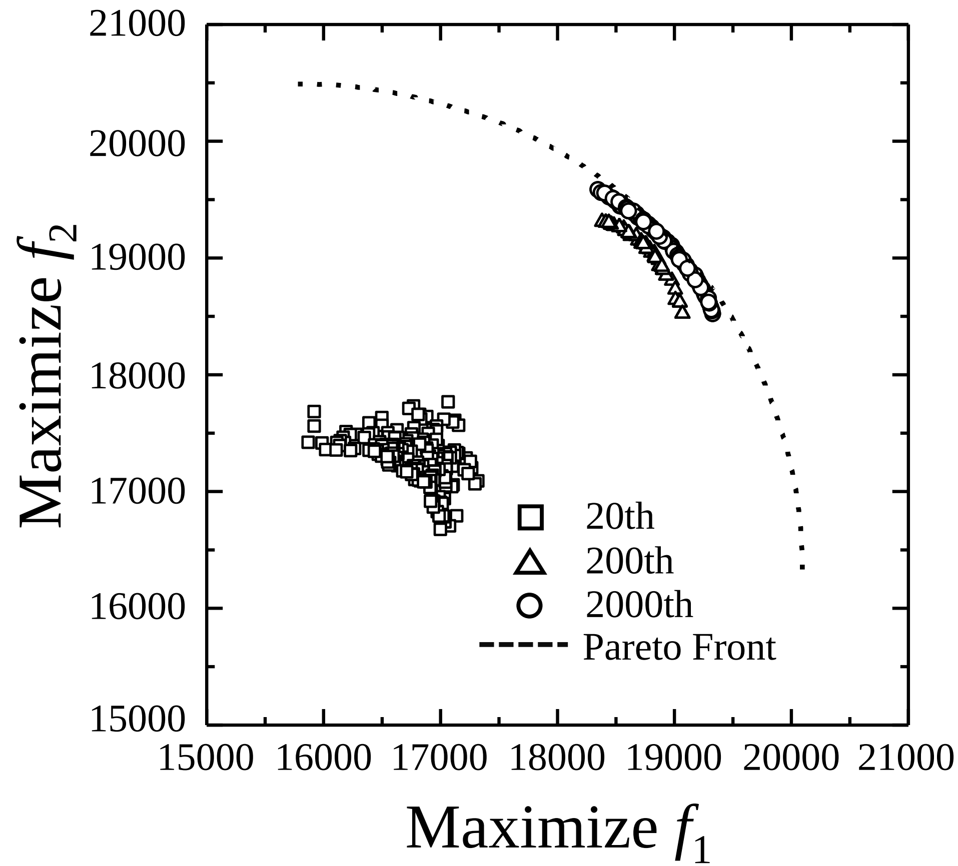
Objective Space

Experimental Results

Two-Objective 500-Item Knapsack Problem



MOEA/D



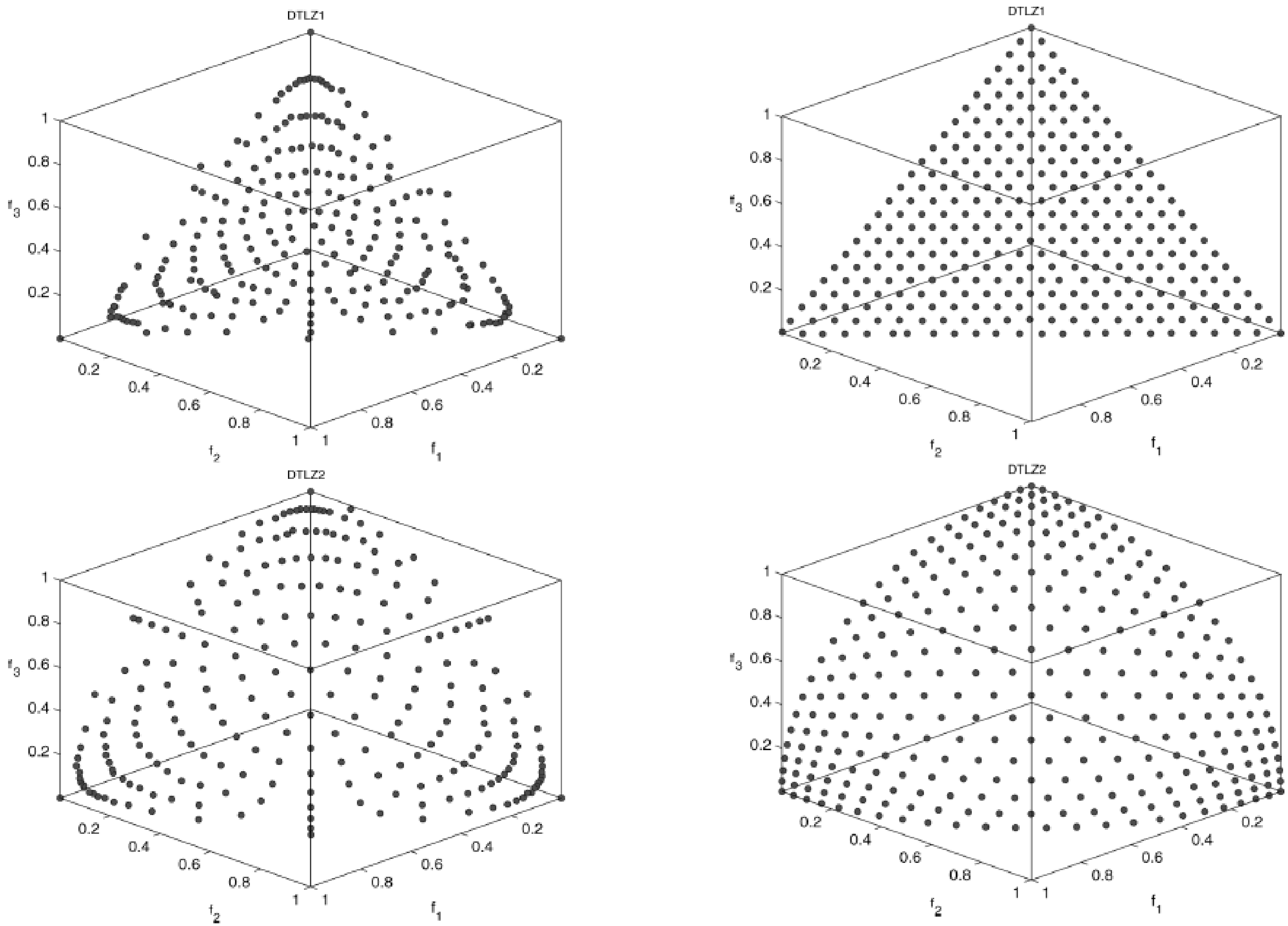
NSGA-II

MOEA/D (IEEE TEVC 2007)

Multi-Objective Evolutionary Algorithm based on Decomposition

DTLZ1

DTLZ2



MOEA/D Tchebycheff

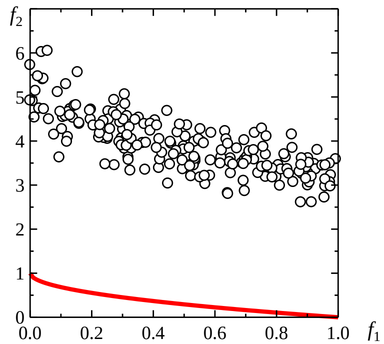
MOEA/D PBI Function

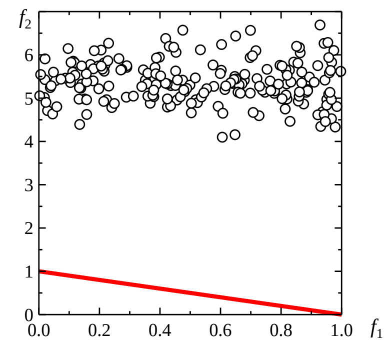
Scalarizing Functions in MOEA/D Paper (2007)

1. **Weighted Sum & Tchebycheff** for Very Difficult Problems
   Knapsack problems with 2-4 objectives and 250-750 items
   Better results are obtained from the weighted sum function.

2. **Tchebycheff** for Difficult Two-Objective Problems: **ZDT**
   Good results are obtained by the Tchebycheff function.

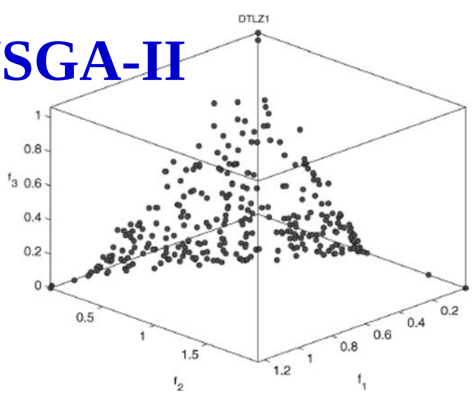
Tchebycheff Function
Property: Any Pareto solution can be obtained by the Tchebycheff function.



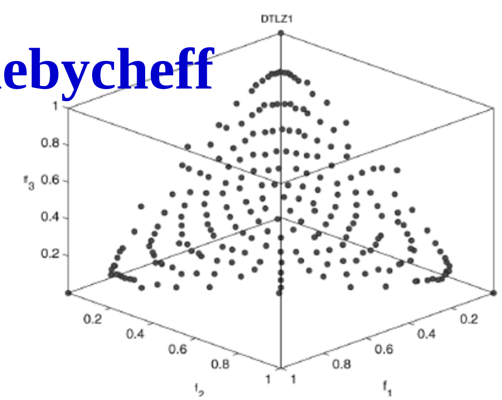


3. **Tchebycheff & PBI** for Easy Three-Objective Problems: **DTLZ**
   Better results are obtained from the PBI function.

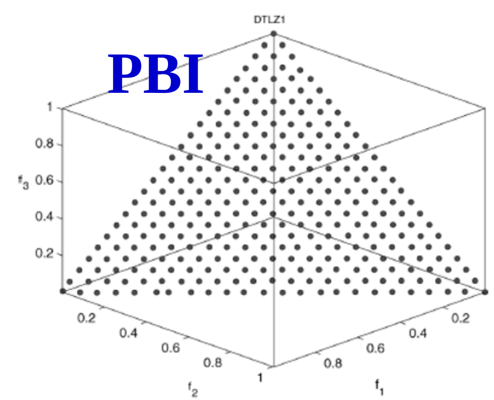
NSGA-II



Tchebycheff



PBI



Scalarizing Functions in MOEA/D Paper (2007)

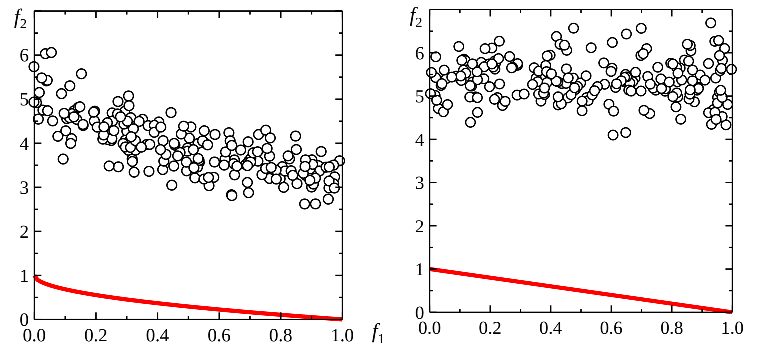
1. Four-Objective Knapsack Problems

Strong convergence is needed ==> **Use of the weighted sum**

2. Two-Objective ZDT Problems

Not very difficult ==> **Use of the Tchebycheff function**

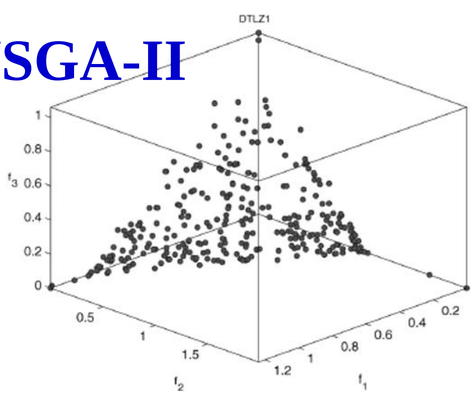
> **Tchebycheff Function**
> **Property:** Any Pareto solution can be obtained by the Tchebycheff function.



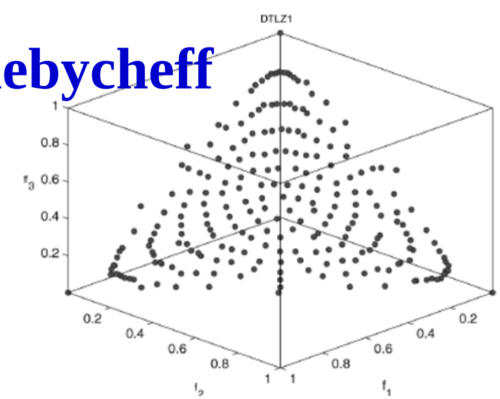
3. Three-Objective DTLZ Problems

Easy convergence ==> **Use of the PBI function for uniformity**

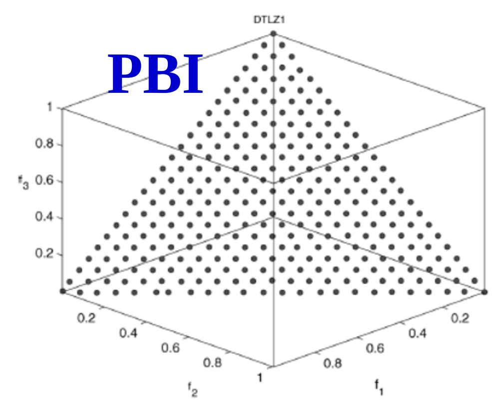
NSGA-II



Tchebycheff



PBI



Importance of Test Problem Choice

If they used only four-objective knapsack problems as test problems:
Only the weighted sum would be used.

Comparison Results on Knapsack Problems (Ishibuchi et al., IEEE TEVC 2015)

TABLE II. RELATIVE AVERAGE HYPERVOLUME. THE REFERENCE POINT (0, 0, ..., 0) IS FAR FROM THE PARETO FRONT.

EMO Algorithm	2-500	4-500	6-500	8-500	10-500
NSGA-II	96.5	86.2	77.8	72.0	65.5
MOEA/D: WS	100.0	**100.0**	**100.0**	**100.0**	**100.0**
MOEA/D: Tchebycheff	100.7	99.7	94.0	90.1	87.7
MOEA/D: PBI (0.01)	99.9	99.7	99.8	**100.0**	99.7
MOEA/D: PBI (0.05)	99.4	98.6	98.5	98.4	98.5
MOEA/D: PBI (0.1)	98.8	96.9	96.6	96.0	95.9
MOEA/D: PBI (0.5)	92.7	82.7	79.6	77.0	74.3
MOEA/D: PBI (1.0)	96.1	78.1	68.0	66.0	63.0
MOEA/D: PBI (5)	100.9	89.3	73.8	67.4	61.9

Importance of Test Problem Choice

If they used only four-objective knapsack problems as test problems:

Only the weighted sum would be used.

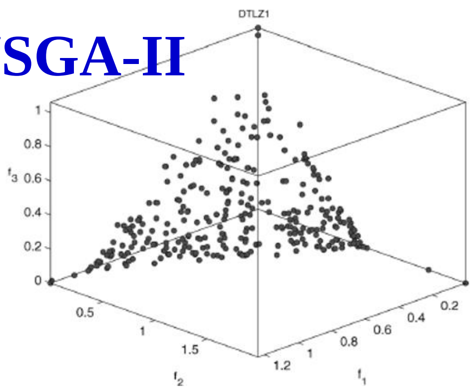
==> The algorithm is not applicable to most test problems such as DTLZ and WFG (since they have concave Pareto fronts).

Importance of Test Problem Choice

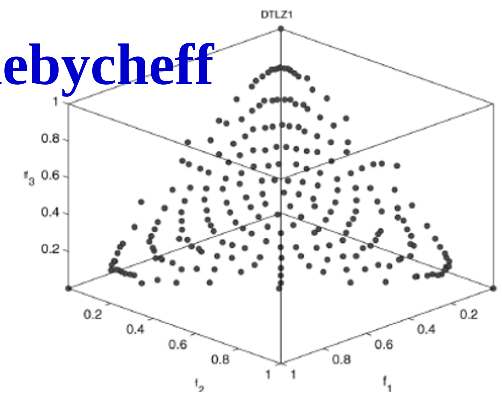
If they used only four-objective knapsack problems as test problems:
Only the weighted sum would be used.
==> The algorithm is not applicable to most test problems such as DTLZ and WFG (since they have concave Pareto fronts).

If they used only three-objective DTLZ test problems:
Only the PBI function would be used.
==> The usefulness of the Tchebycheff function cannot be demonstrated. The algorithm is not always useful for many-objective problems.

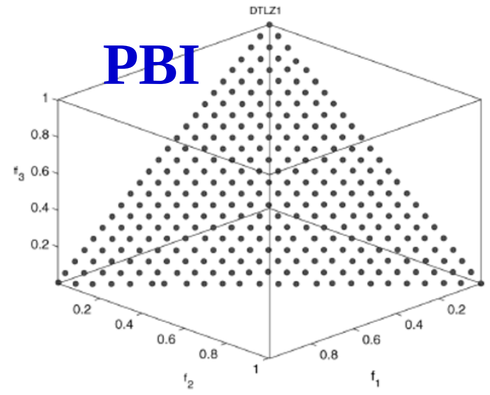
NSGA-II



Tchebycheff



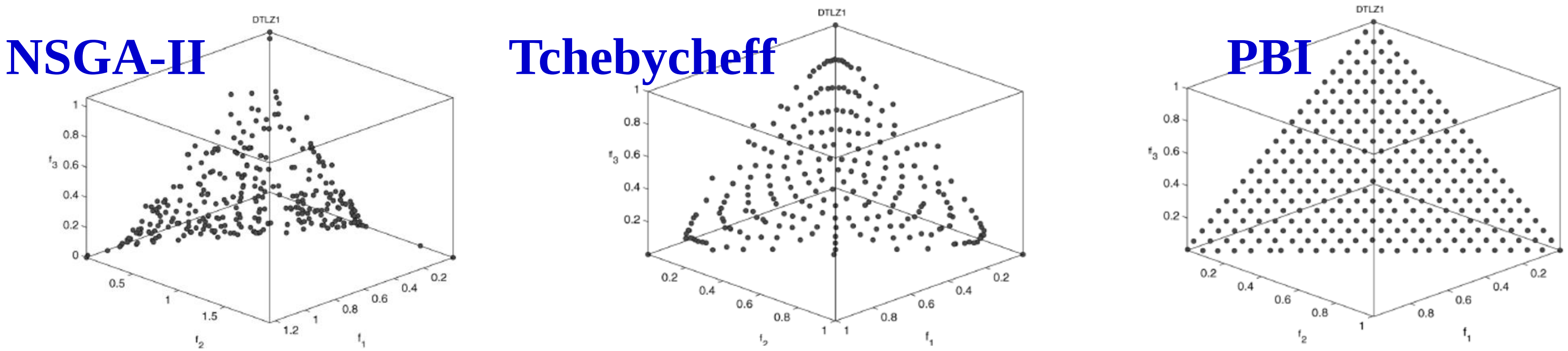
PBI



Importance of Test Problem Choice

If they used only four-objective knapsack problems as test problems:
Only the weighted sum would be used.
==> The algorithm is not applicable to most test problems such as DTLZ and WFG (since they have concave Pareto fronts).

If they used only three-objective DTLZ test problems:
Only the PBI function would be used.
==> The usefulness of the Tchebycheff function cannot be demonstrated.
The algorithm is not always useful for many-objective problems.



If they used only two-objective ZDT test problems:
Only the Tchebycheff function would be used.
==> Good solution distributions cannot be demonstrated.
The algorithm is not always useful for many-objective problems.

Comparison Results on Knapsack Problems (Ishibuchi et al., IEEE TEVC 2015)

TABLE II. RELATIVE AVERAGE HYPERVOLUME. THE REFERENCE POINT (0, 0, ..., 0) IS FAR FROM THE PARETO FRONT.

EMO Algorithm	2-500	4-500	6-500	8-500	10-500
NSGA-II	96.5	86.2	77.8	72.0	65.5
MOEA/D: WS	100.0	**100.0**	**100.0**	**100.0**	**100.0**
MOEA/D: Tchebycheff	100.7	99.7	94.0	90.1	87.7
MOEA/D: PBI (0.01)	99.9	99.7	99.8	**100.0**	99.7
MOEA/D: PBI (0.05)	99.4	98.6	98.5	98.4	98.5
MOEA/D: PBI (0.1)	98.8	96.9	96.6	96.0	95.9
MOEA/D: PBI (0.5)	92.7	82.7	79.6	77.0	74.3
MOEA/D: PBI (1.0)	96.1	78.1	68.0	66.0	63.0
MOEA/D: PBI (5)	100.9	89.3	73.8	67.4	61.9

If they used only two-objective ZDT test problems:

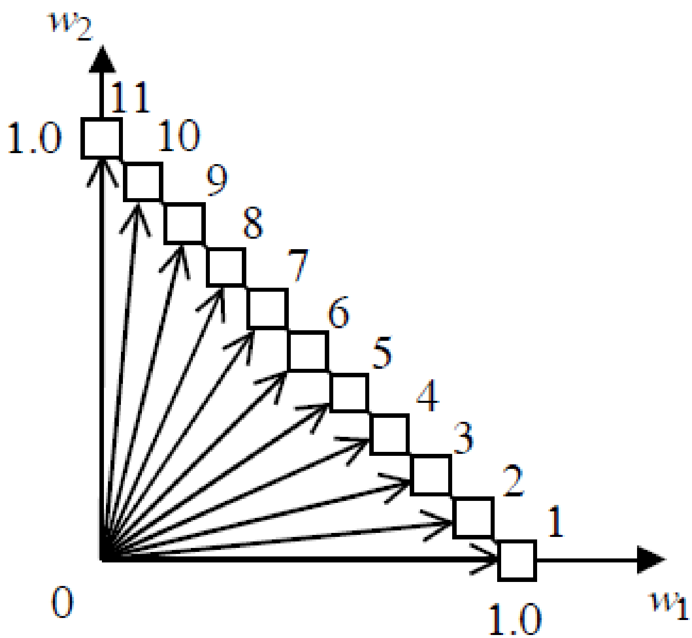
Only the Tchebycheff function would be used.

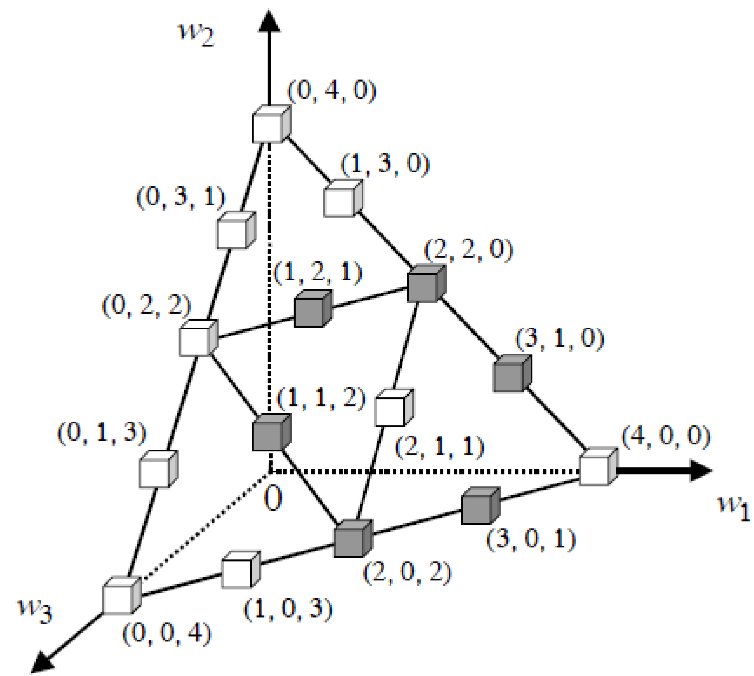
==> Good solution distributions cannot be demonstrated.

The algorithm is not always useful for many-objective problems.

A Similar Algorithm to MOEA/D: Cellular MOGA

T. Murata, H. Ishibuchi, M. Gen: Specification of Genetic Search Directions in Cellular Multi-objective Genetic Algorithms. EMO 2001 (1st EMO Conference).





Almost the same framework as MOEA/D

Difference 1: Only the flowshop scheduling problems were used.

Difference 2: Only the weighted sum was used.

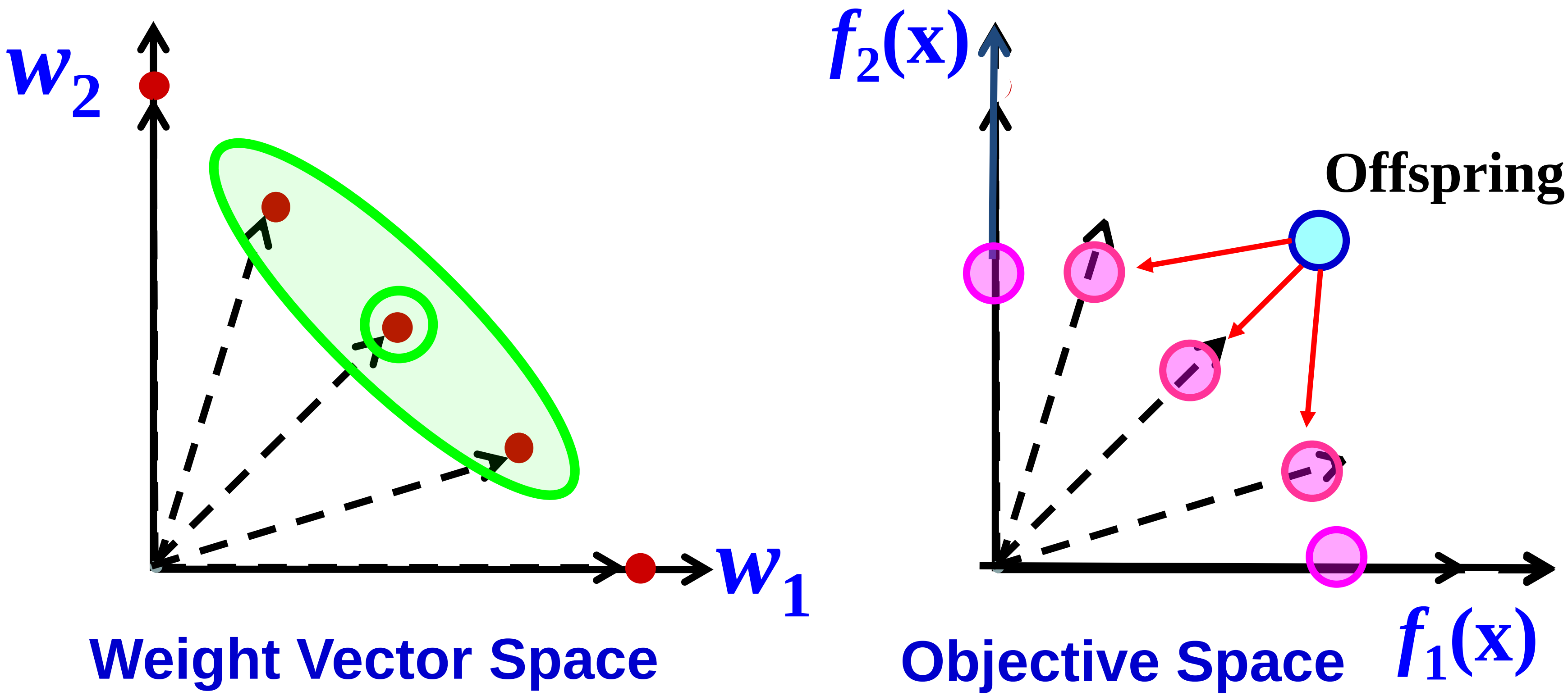
Difference 3: Only a single solution could be replaced.

Difference 4: The algorithm name was not general (very special)

Potential Difficulties of MOEA/D

1. A new solution can be better than all neighbors.

Many solutions in the current population can be replaced with a single offspring. ==> strong convergence, weak diversity.

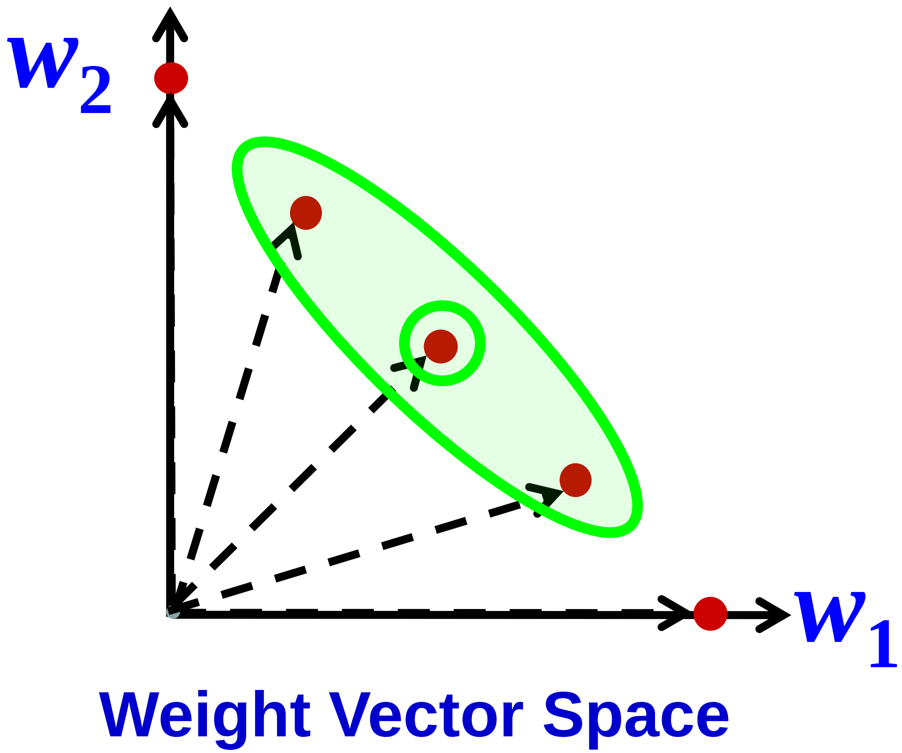


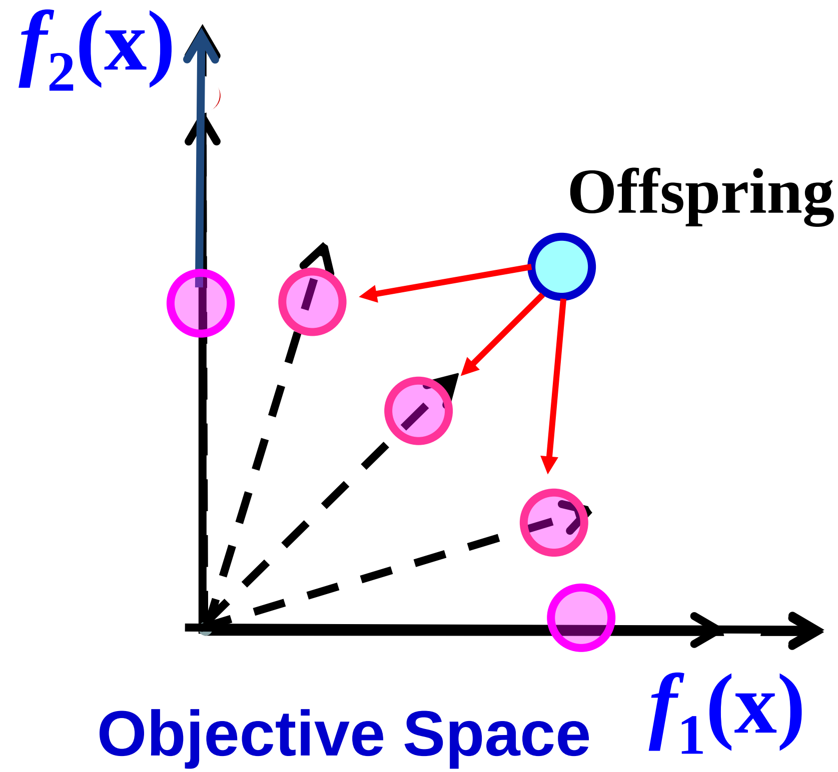
Potential Difficulties of MOEA/D

1. A new solution can be better than all neighbors.

Many solutions in the current population can be replaced with a single offspring. ==> strong convergence, weak diversity.

Modification: Only a single solution (or only a small number of solutions) can be replaced with a single offspring.

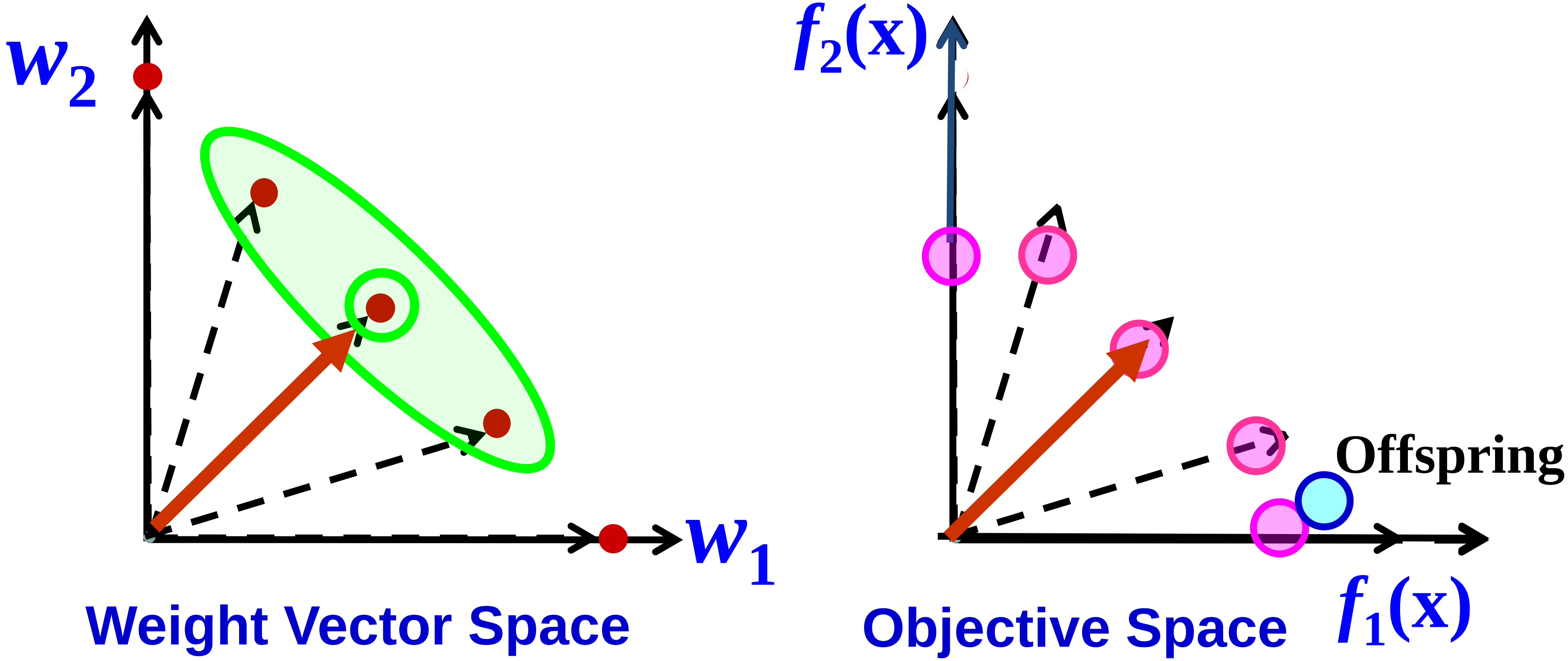




Potential Difficulties of MOEA/D

2. A new solution can be far away from the neighborhood.

A new solution is not always close to its parents.

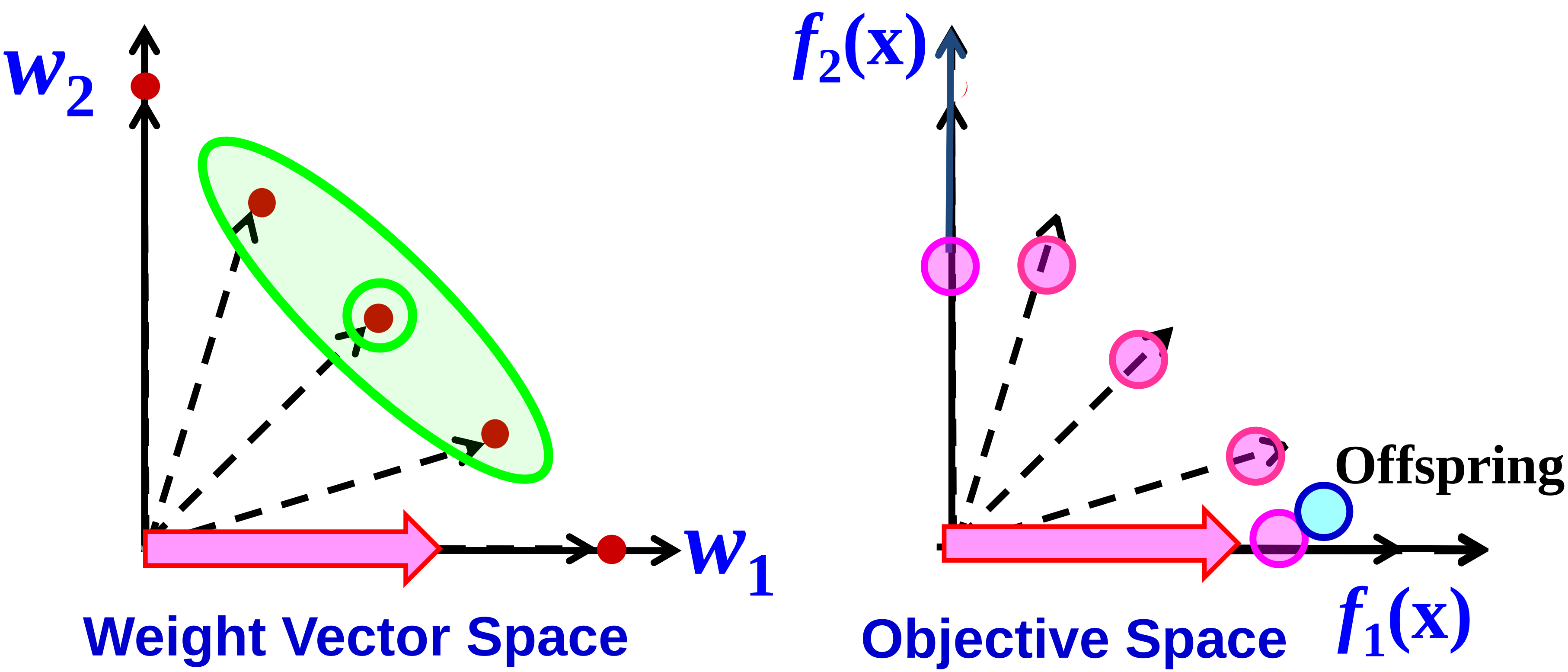


Potential Difficulties of MOEA/D

2. A new solution can be far away from the neighborhood.

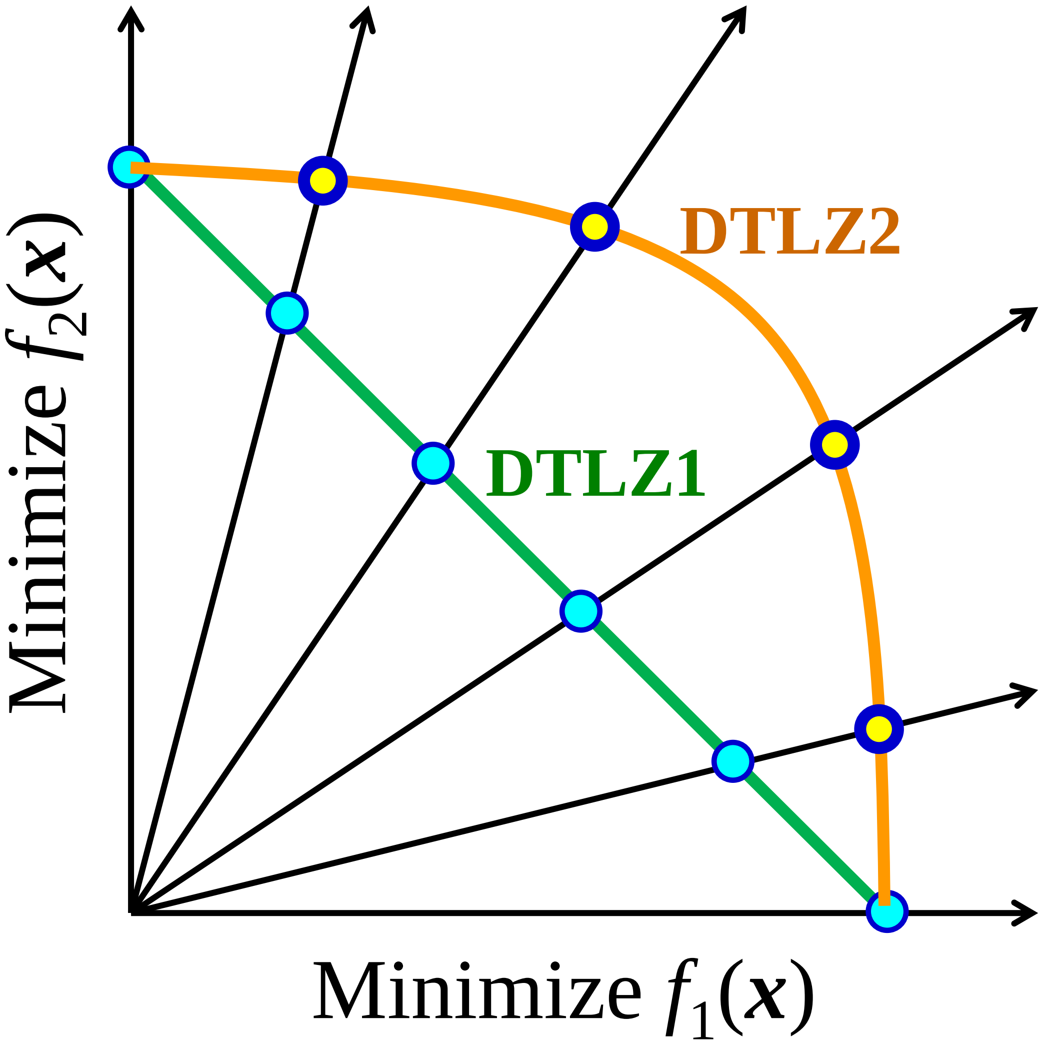
A new solution is not always close to its parents.

Modification: The closest weight vector to the new solution is identified, and the replacement is examined in the neighborhood of the closest weight vector.



Potential Difficulties of MOEA/D

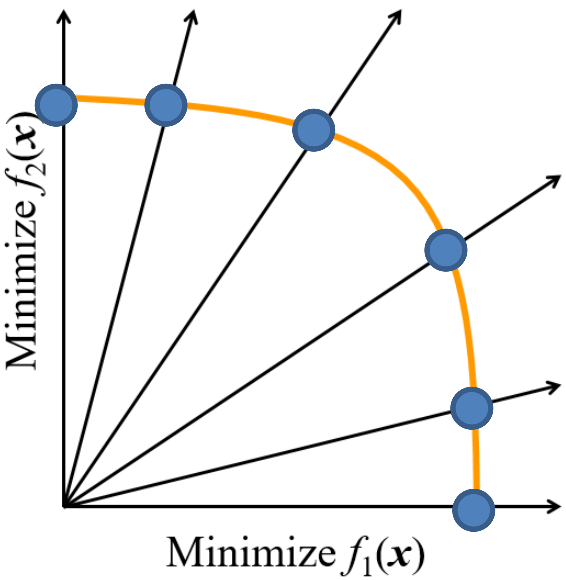
3. Uniform weight vectors are not always appropriate.



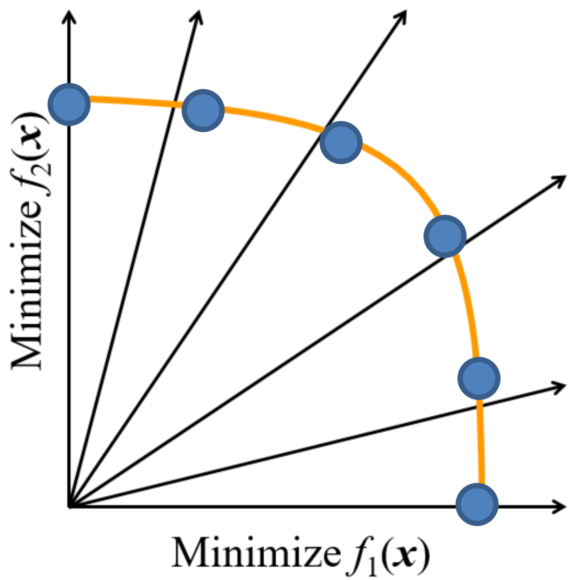
Potential Difficulties of MOEA/D

3. Uniform weight vectors are not always appropriate.

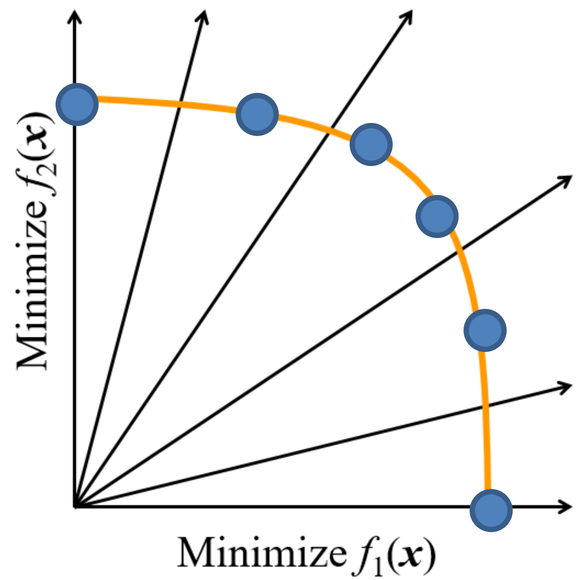
Q. Which is the best solution set among the following three sets ?



(A)



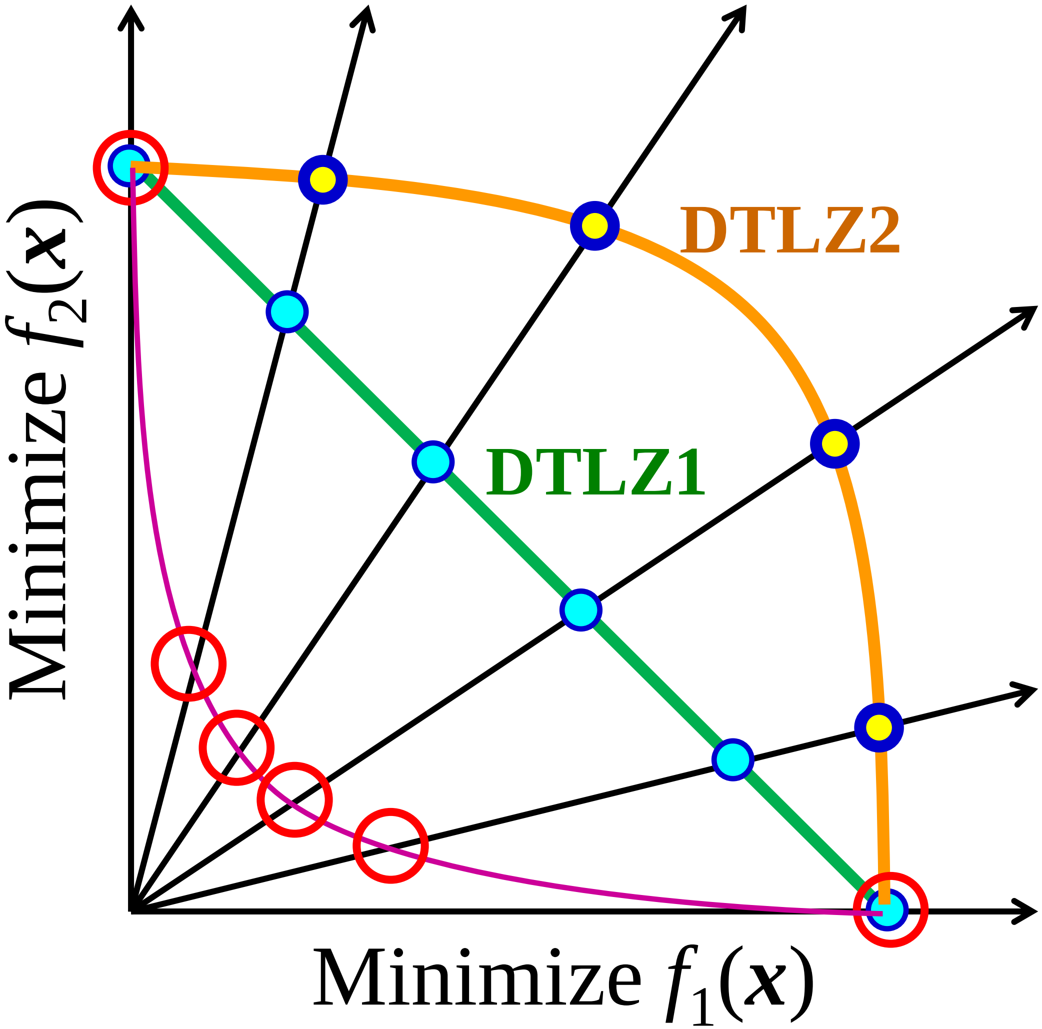
(B)



(C)

Potential Difficulties of MOEA/D

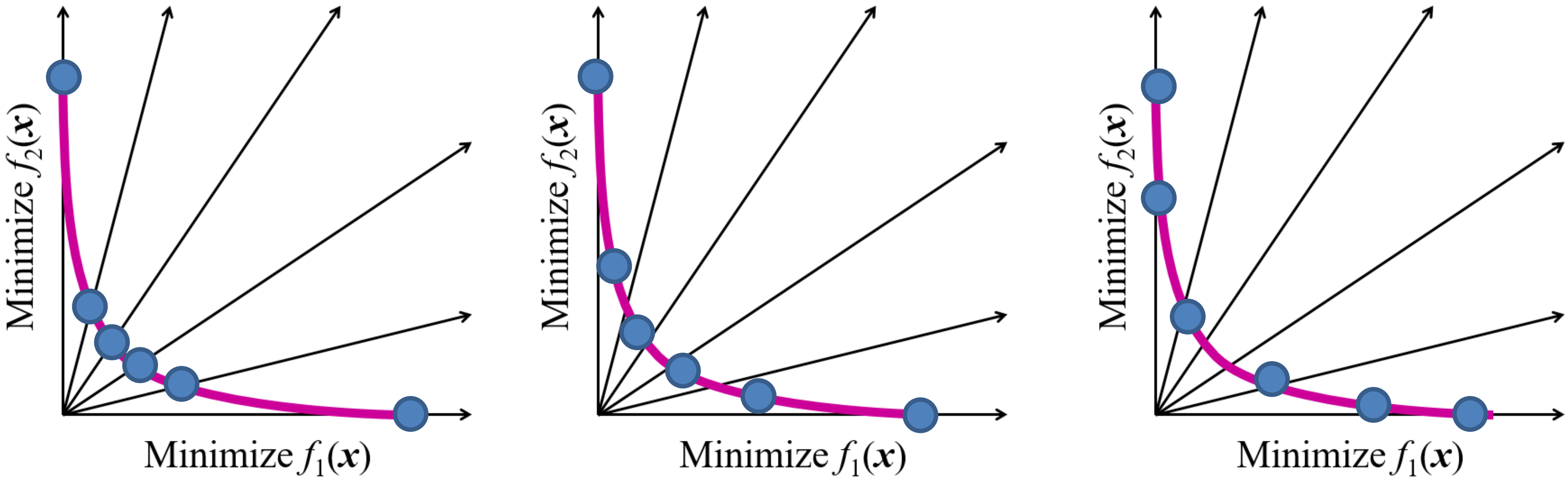
3. Uniform weight vectors are not always appropriate.



Potential Difficulties of MOEA/D

3. Uniform weight vectors are not always appropriate.

Q. Which is the best solution set among the following three sets ?

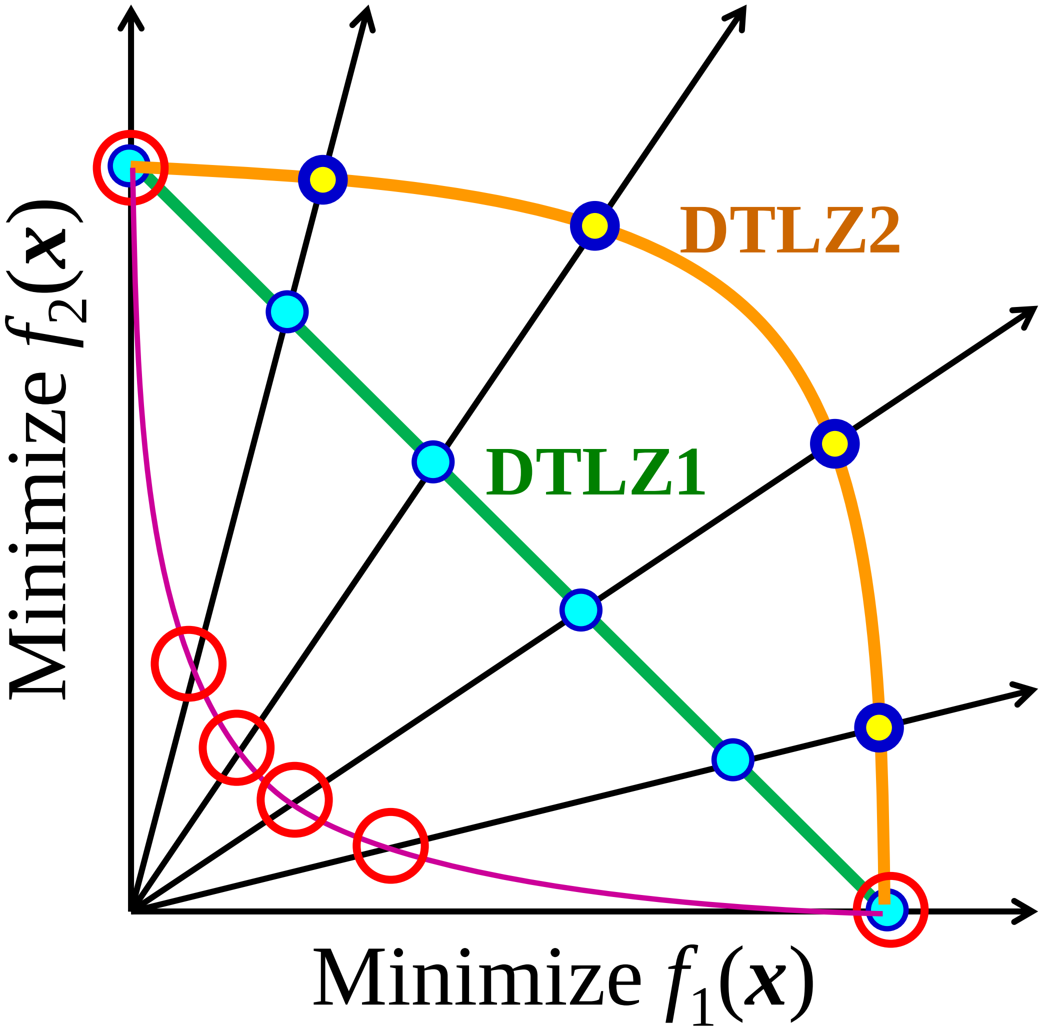


(A) **(B)** **(C)**

Potential Difficulties of MOEA/D

3. Uniform weight vectors are not always appropriate.

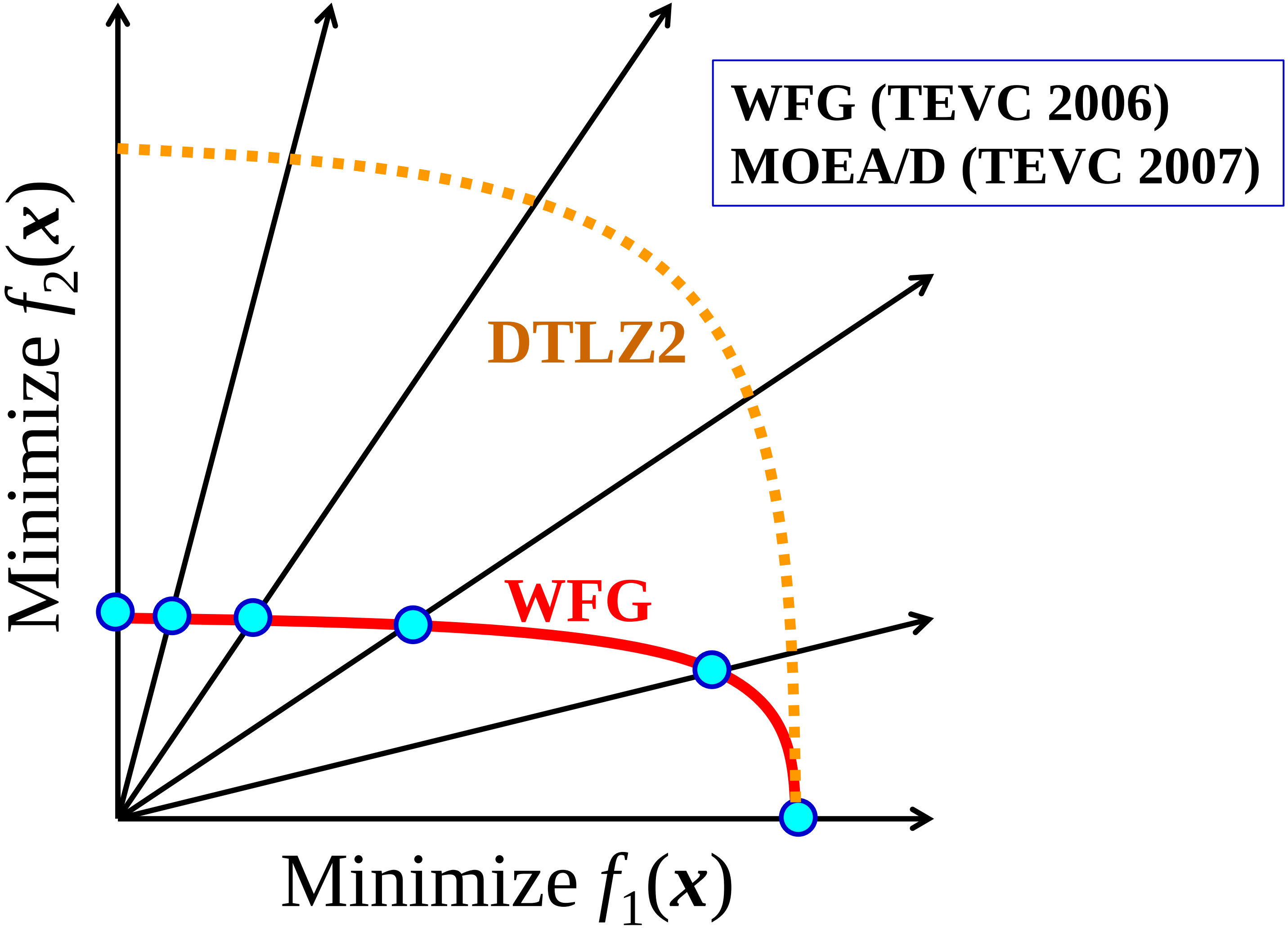
Modification: Adaptation of weight vectors.



Potential Difficulties of MOEA/D

3. Uniform weight vectors are not always appropriate.

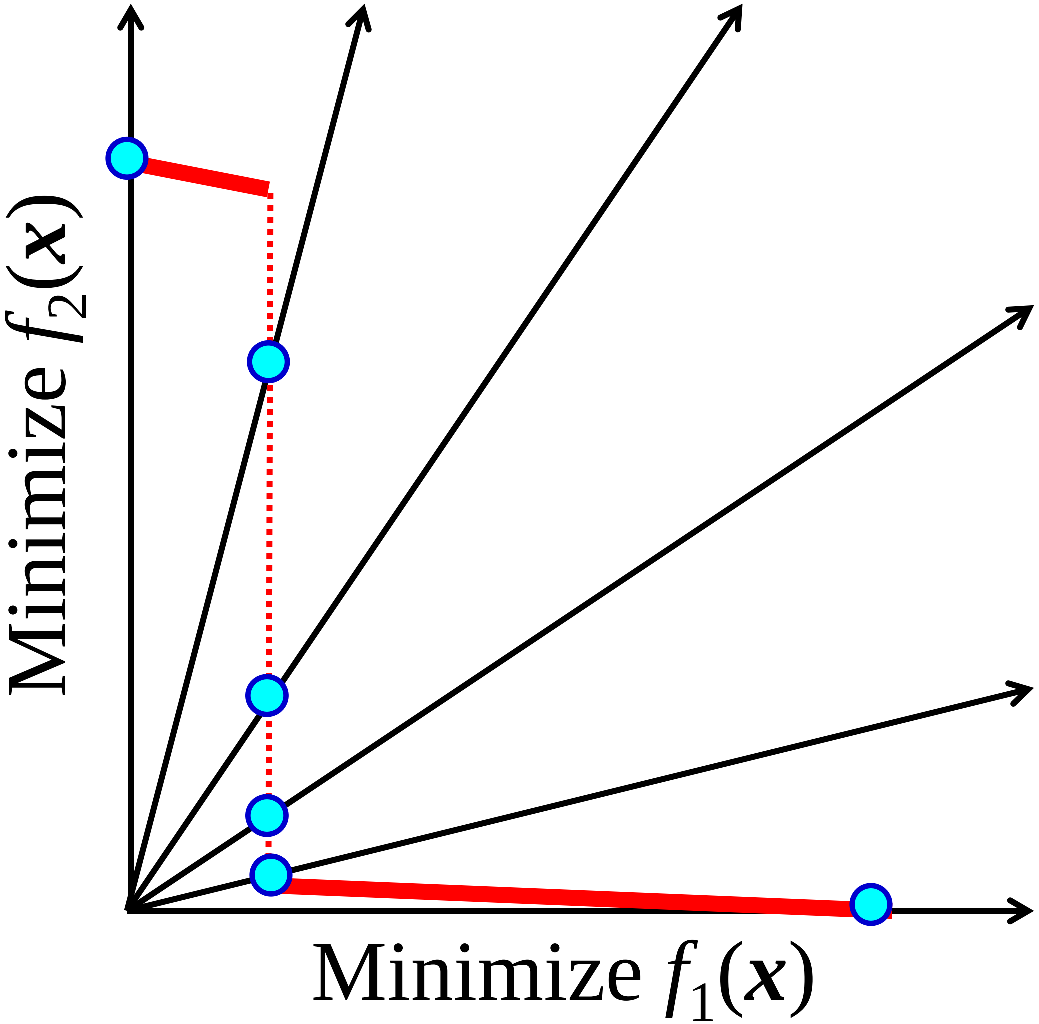
Modification: Normalization of the objective space.



Potential Difficulties of MOEA/D

3. Uniform weight vectors are not always appropriate.

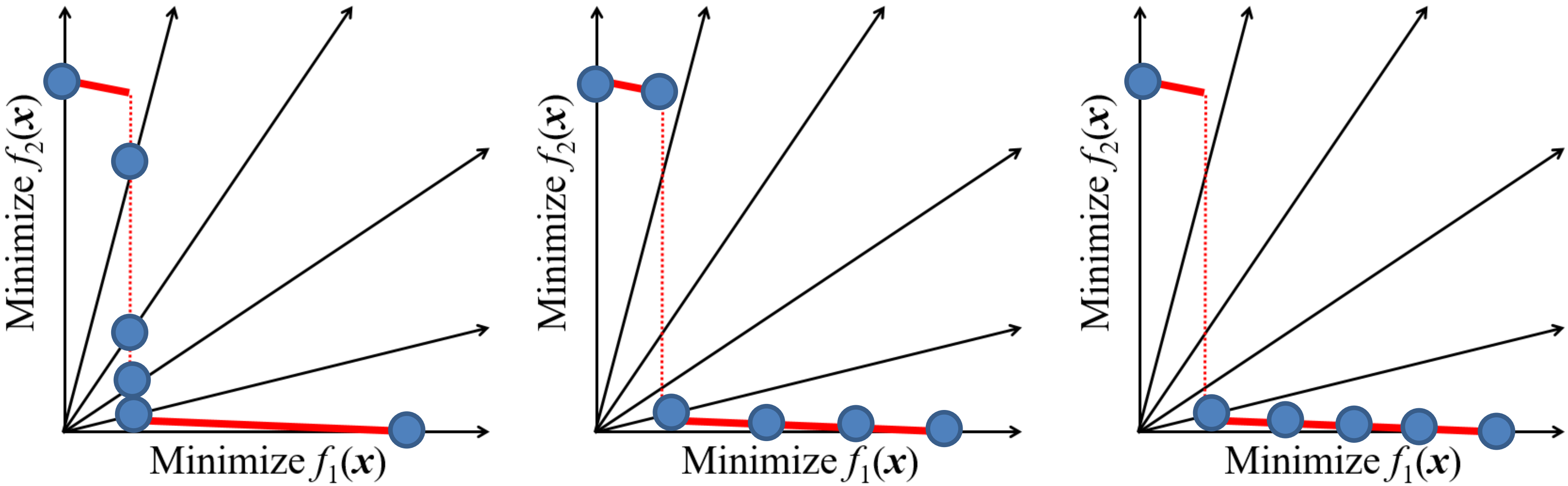
Modification: Adaptation of weight vectors.



Potential Difficulties of MOEA/D

3. Uniform weight vectors are not always appropriate.

Q. Which is the best solution set among the following three sets ?



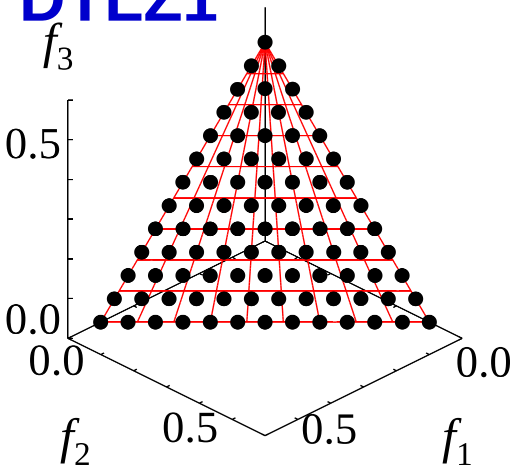
(A) **(B)** **(C)**

Potential Difficulties of MOEA/D

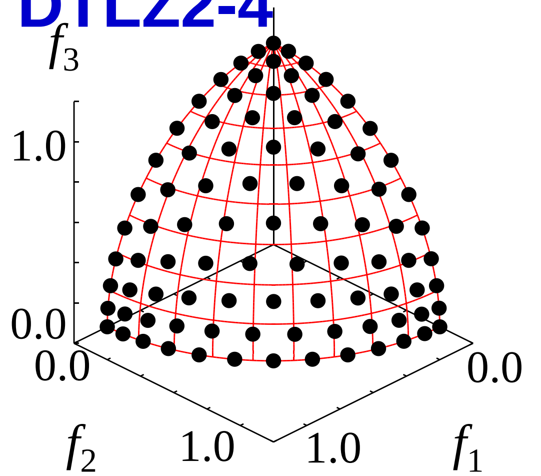
3. Uniform weight vectors are not always appropriate.

Modification: Adaptation of weight vectors.

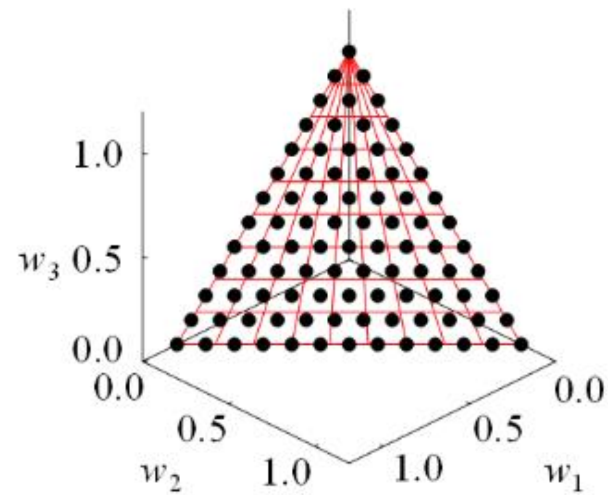
DTLZ1



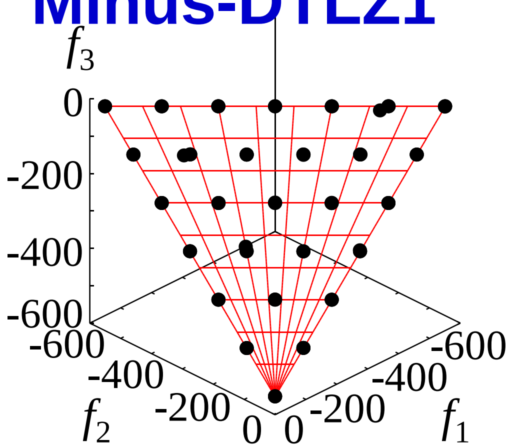
DTLZ2-4



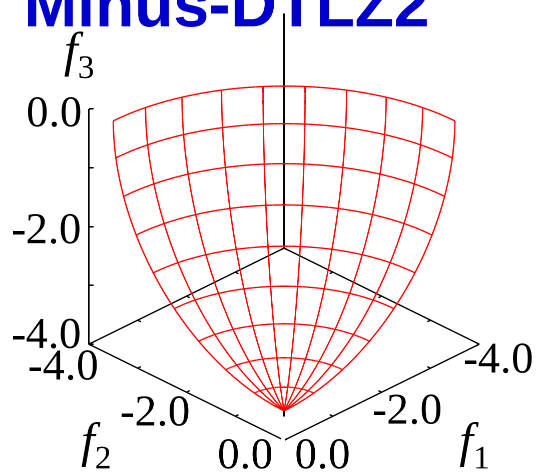
Weight Vectors



Minus-DTLZ1



Minus-DTLZ2



Potential Difficulties of MOEA/D

4. No inside weight vectors for many-objective problems.

$$w_1 + w_2 + \cdots + w_m = 1$$

$$w_i \in \{0/H, 1/H, \ldots, H/H\},\ \ i = 1, 2, ..., m .$$

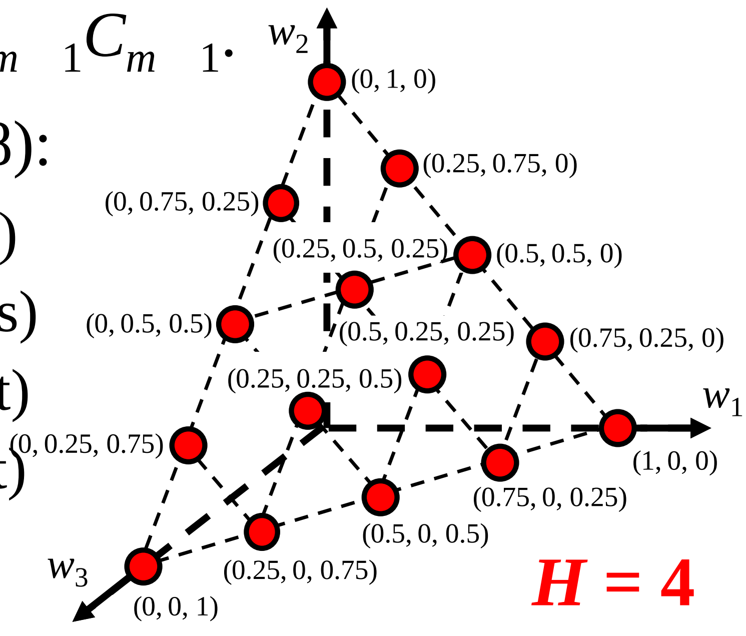
(1) When $H = m$, only one vector $(1/m, ..., 1/m)$ is an inside vector.
All the other vectors have at least one zero element ($w_i = 0$).

(2) When $H < m$, no vectors are inside vectors.

(3) The total number of weight vectors is ${}_{H+m-1}C_{m-1}$.

For example, in 8-objective problems ($m = 8$):

$H = 4$: ${}_{11}C_7 = 330$ (at least four zero elements)
$H = 5$: ${}_{12}C_7 = 792$ (at least three zero elements)
$H = 6$: ${}_{13}C_7 = 1{,}716$ (at least two zero element)
$H = 7$: ${}_{14}C_7 = 3{,}432$ (at least one zero element)
$H = 8$: ${}_{15}C_7 = 6{,}435$ (one inside vector)

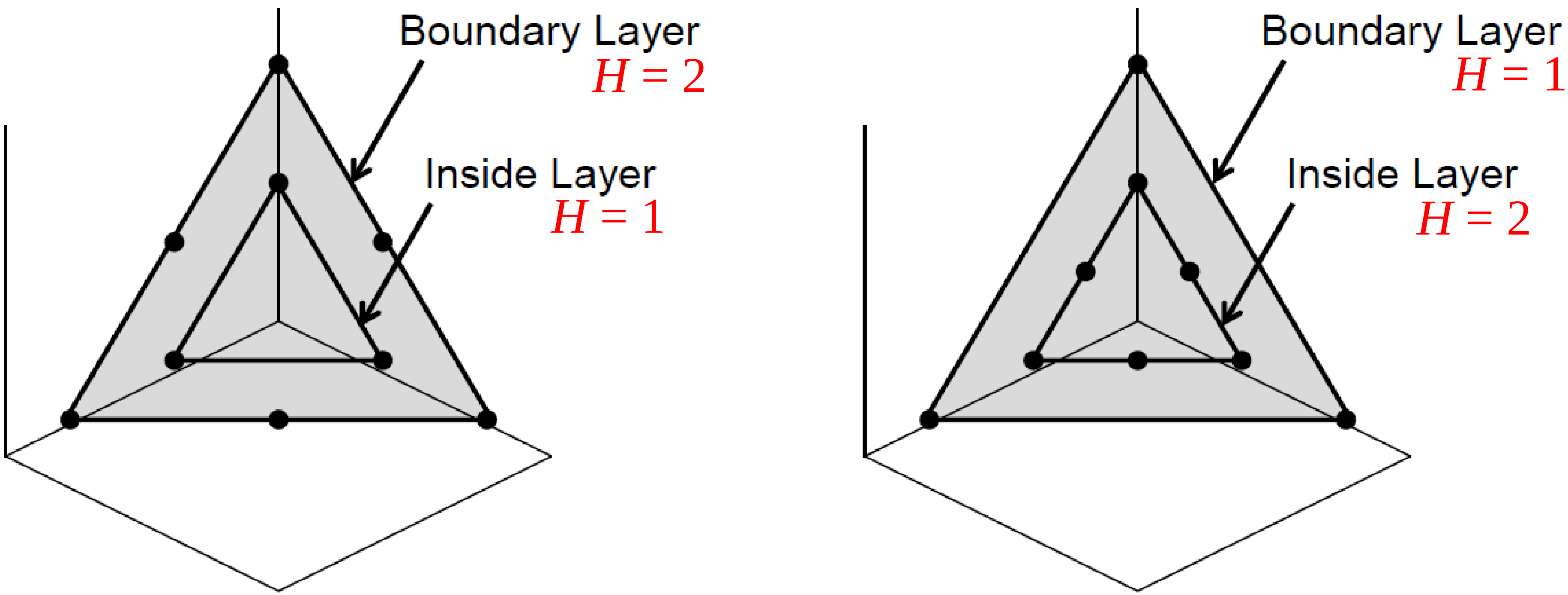


In 2007, many-objective research was not popular.

Potential Difficulties of MOEA/D

4. No inside weight vectors for many-objective problems.

Modification: Two-layer method (NSGA-III, 2014).



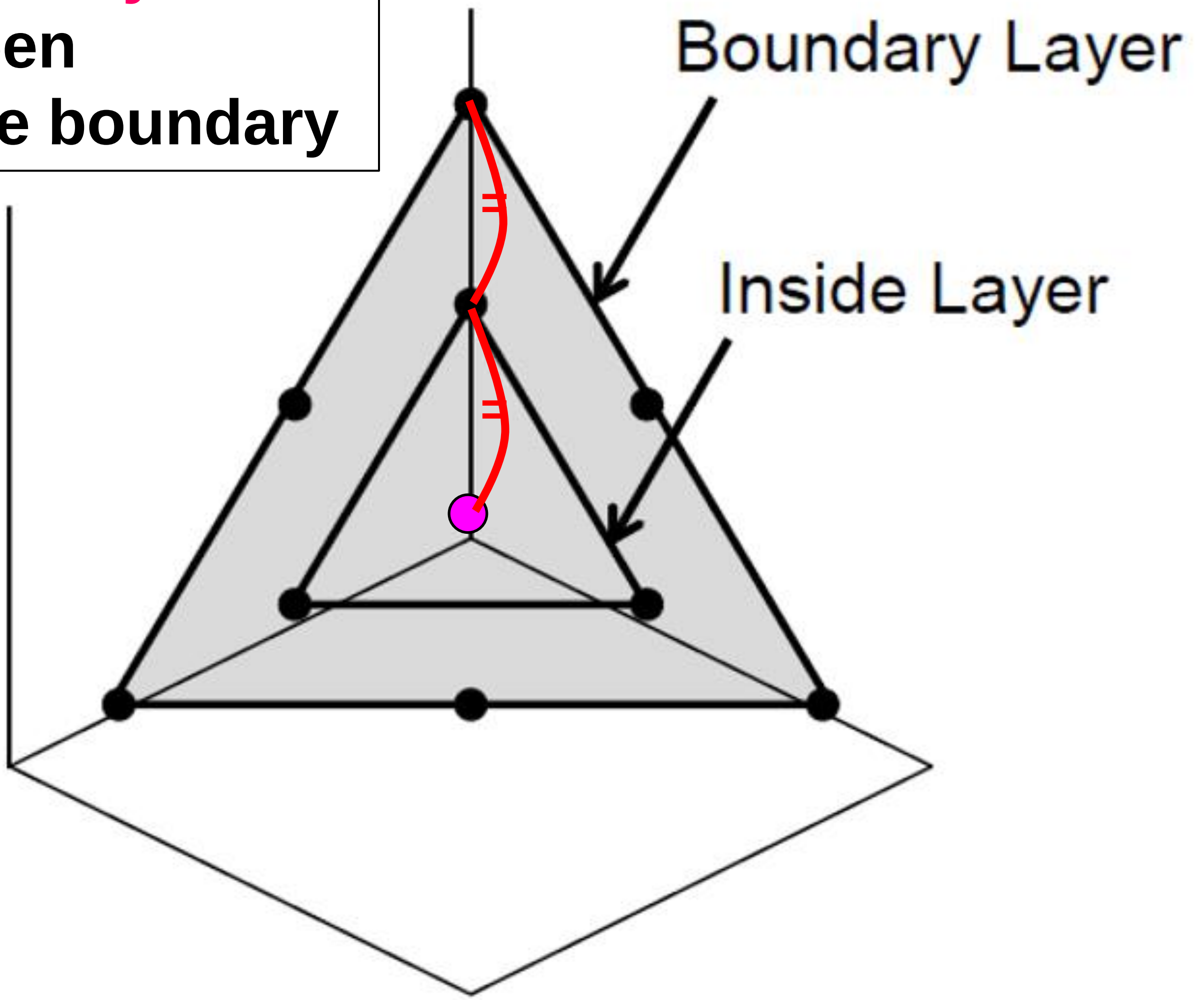
(a) 6 boundary and 3 inside vectors. (b) 3 boundary and 6 inside vectors.

Potential Difficulties of MOEA/D

4. No inside weight vectors for many-objective problems.

Modification: Two-layer method (NSGA-III, 2014).

Location of the inside layer:
In the middle between the center and the boundary



Center vector:
$(1/m, 1/m, ..., 1/m)$

Relation between Test Problems and EMO Algorithms

Around 1995

Simple two-objective problems (no need of strong convergence)

Non-elitist Pareto dominance-based algorithms: MOGA, NPGA, NSGA

These algorithms work well on simple two-objective problems

Around 2000

Difficult two-objective problems: ZDT

Elitist Pareto dominance-based algorithms: SPEA, SPEA2, NSGA-II

These algorithms work well on two-objective problems

2000s

Easy scalable test problems (the number of objectives can be arbitrarily specified: *m*-objective problems): DTLZ and WFG

Decomposition-based algorithms: MOEA/D

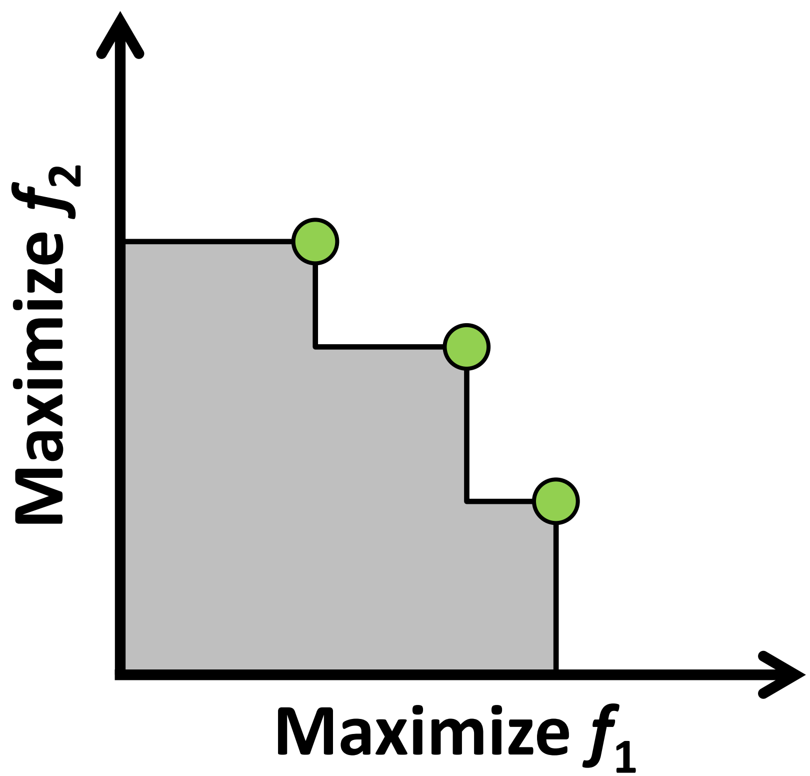
Indicator-based algorithms: SMS-EMOA

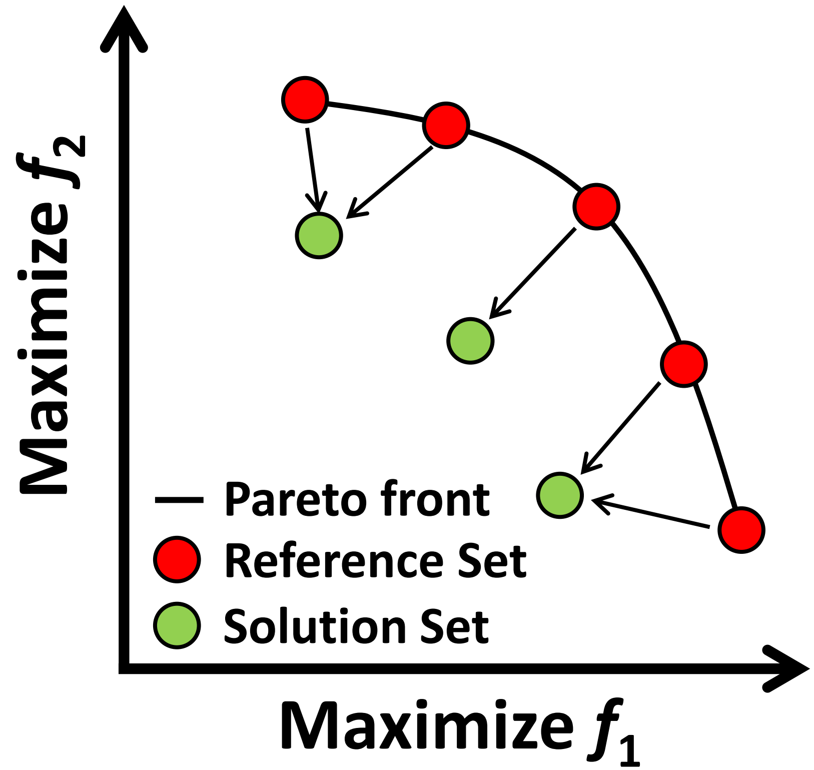
Indicator-Based EMO Algorithms

Basic Idea:

Since an EMO algorithm is evaluated by an indicator, it may be a good idea to directly optimize the indicator instead of trying to find well-distributed solutions over the entire Pareto front.

==> The EMO algorithm will be evaluated as the best algorithm.





Indicator-Based Algorithm: **SMS-EMOA (HV-based Algorithm)**

Michael Emmerich

FOLLOW

Associate Professor, LIACS, Leiden University & Lead AI Scientist at SILO.AI

Verified email at liacs.nl - Homepage

Multicriteria Optimization Complex Networks Gaussian Processes Sustainable Design Chemoinformatics

TITLE	CITED BY	YEAR
SMS-EMOA: Multiobjective selection based on dominated hypervolume N Beume, B Naujoks, M Emmerich European Journal of Operational Research 181 (3), 1653-1669	1710	2007
An EMO algorithm using the hypervolume measure as selection criterion M Emmerich, N Beume, B Naujoks International Conference on Evolutionary Multi-Criterion Optimization, 62-76	665	2005
Single-and multiobjective evolutionary optimization assisted by Gaussian random field metamodels MTM Emmerich, KC Giannakoglou, B Naujoks IEEE Transactions on Evolutionary Computation 10 (4), 421-439	578	2006

Indicator-Based Algorithm: **SMS-EMOA (HV-based Algorithm)**

Boris Naujoks

FOLLOW

Cologne University of Applied Sciences
Verified email at fh-koeln.de

Evolutionary Algorithms   Multiobjective Optimization

TITLE	CITED BY	YEAR
SMS-EMOA: Multiobjective selection based on dominated hypervolume N Beume, B Naujoks, M Emmerich European Journal of Operational Research 181 (3), 1653-1669	1710	2007
An EMO algorithm using the hypervolume measure as selection criterion M Emmerich, N Beume, B Naujoks International Conference on Evolutionary Multi-Criterion Optimization, 62-76	665	2005
Single-and multiobjective evolutionary optimization assisted by Gaussian random field metamodels MTM Emmerich, KC Giannakoglou, B Naujoks IEEE Transactions on Evolutionary Computation 10 (4), 421-439	578	2006

Indicator-Based Algorithm:

SMS-EMOA (HV-based Algorithm)

Goal: To maximize the HV value of the population.

Algorithm Framework: (+1) generation update.

One new solution is generated, and the worst solution is removed.

Fitness Evaluation: HV Contribution.

HVC($\boldsymbol{x}$) = HV(S)   HV(S {$\boldsymbol{x}$})

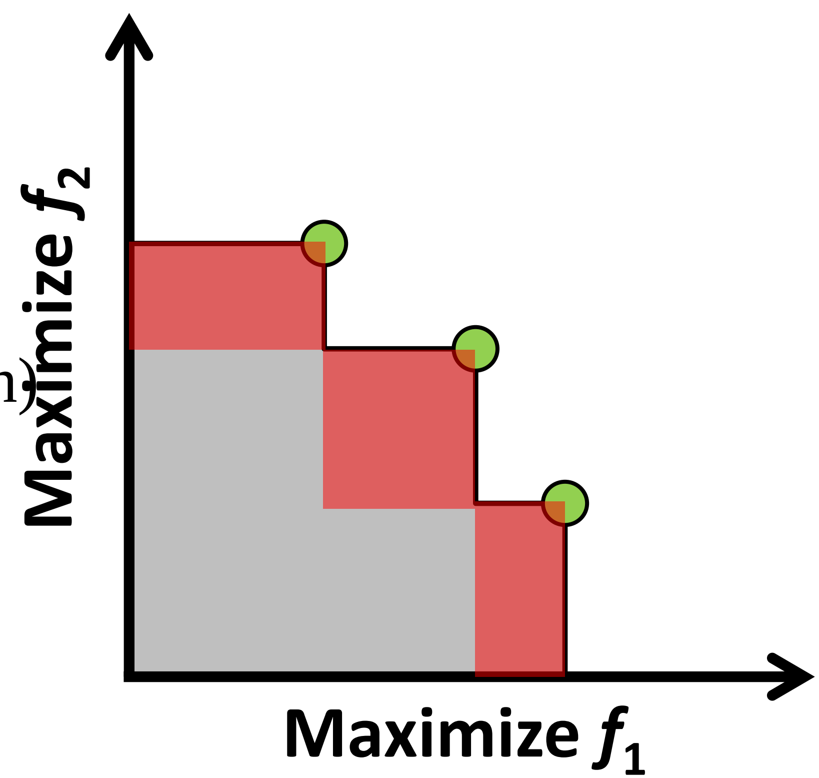
Worst solution in   +1 solutions

= Solution with the smallest HVC.

Reference point:

Current worst value   1 (maximization)

Current worst value + 1 (minimization)



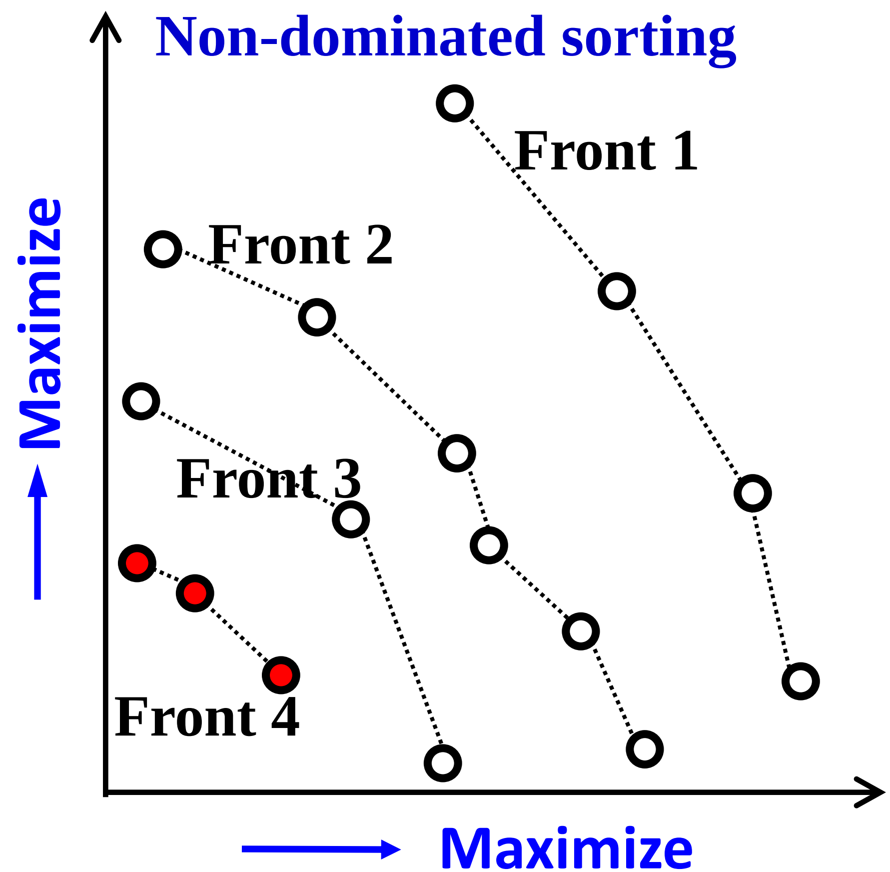
Algorithm Framework: (+1) generation update.

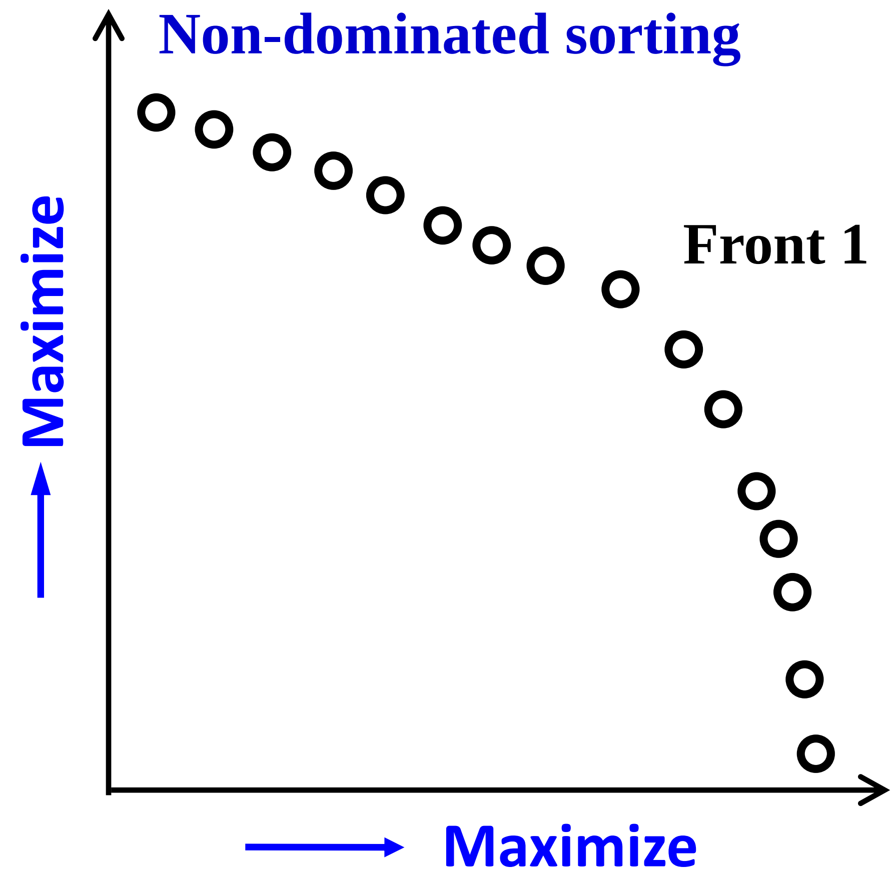
One new solution is generated, and the worst solution is removed.

Fitness Evaluation: HV Contribution.

HVC($\boldsymbol{x}$) = HV(S) HV(S {$\boldsymbol{x}$}) among the worst front solutions.

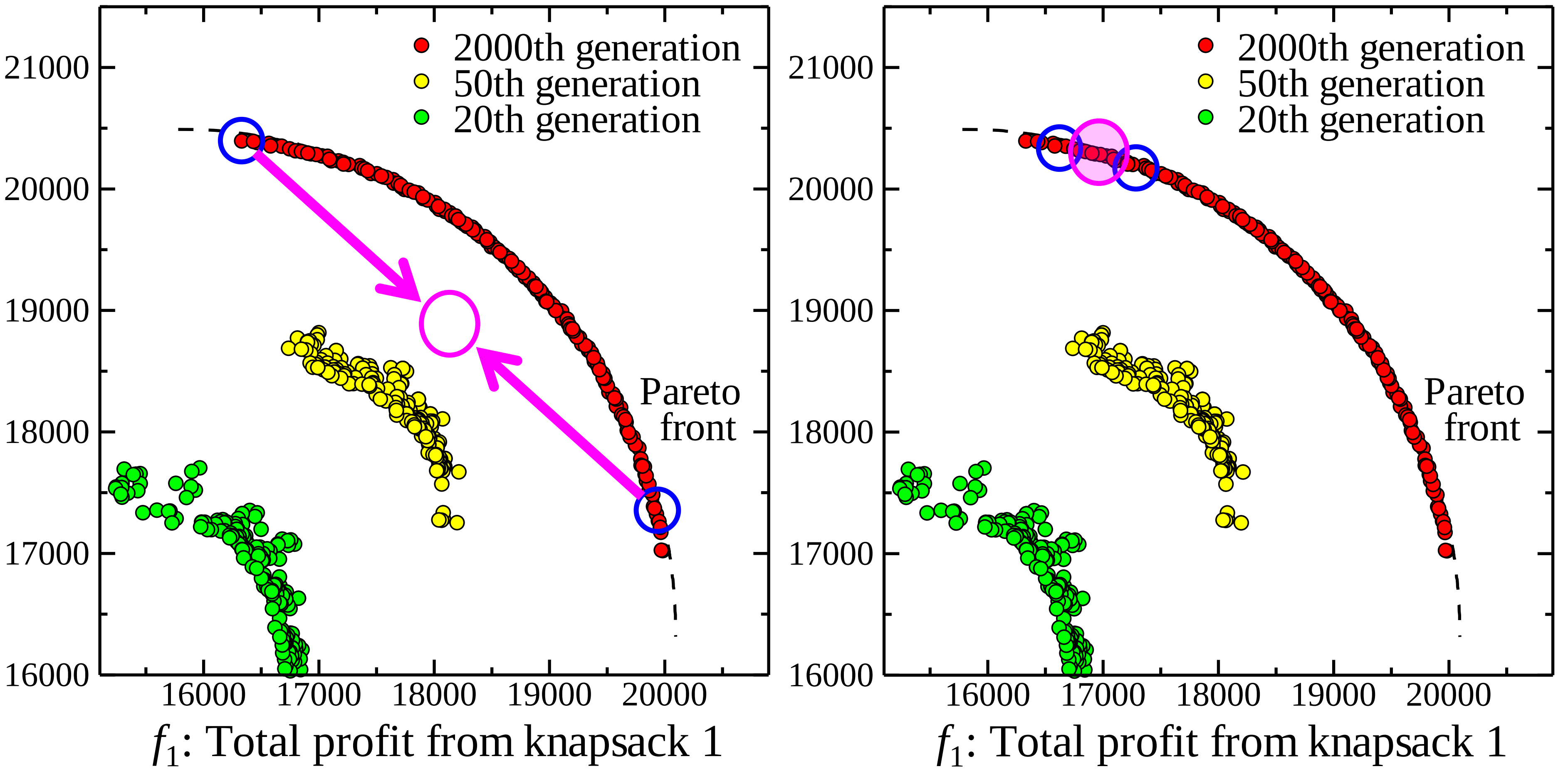
Worst solution in +1 solutions = Solution with the smallest HVC.





One Minor Difficulty of SMS-EMOA

Random Selection of Parents



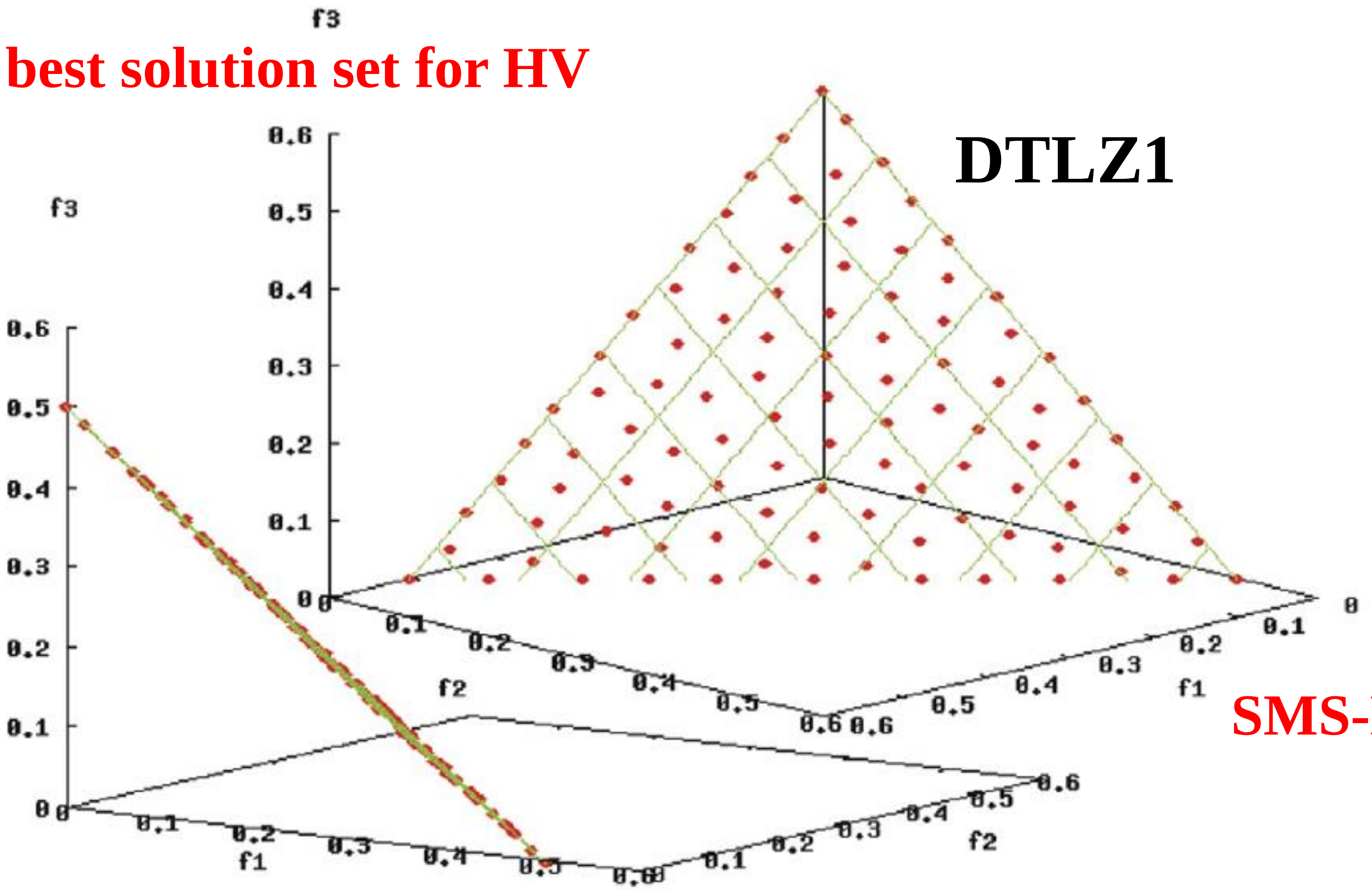
Different Parents **Similar Parents**

Three Major Difficulties of SMS-EMOA

1. **Uniformity of Solutions**
2. **Large Computation Load for HVC Calculation**
3. **Reference Point Specification**

The best solution set for HV

DTLZ1



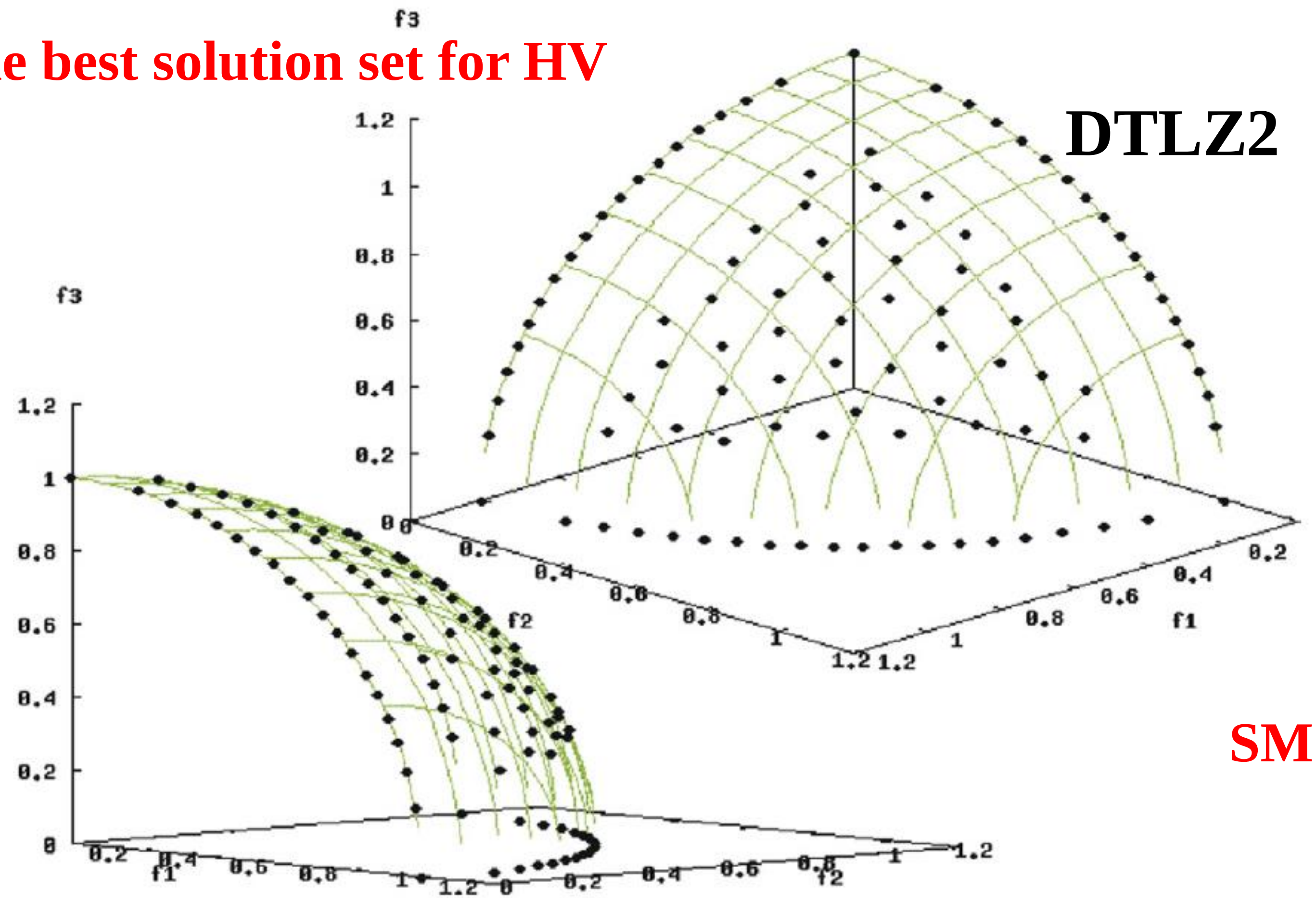
SMS-EMOA (2007)

Three Major Difficulties of SMS-EMOA

1. Uniformity of Solutions
2. Large Computation Load for HVC Calculation
3. Reference Point Specification

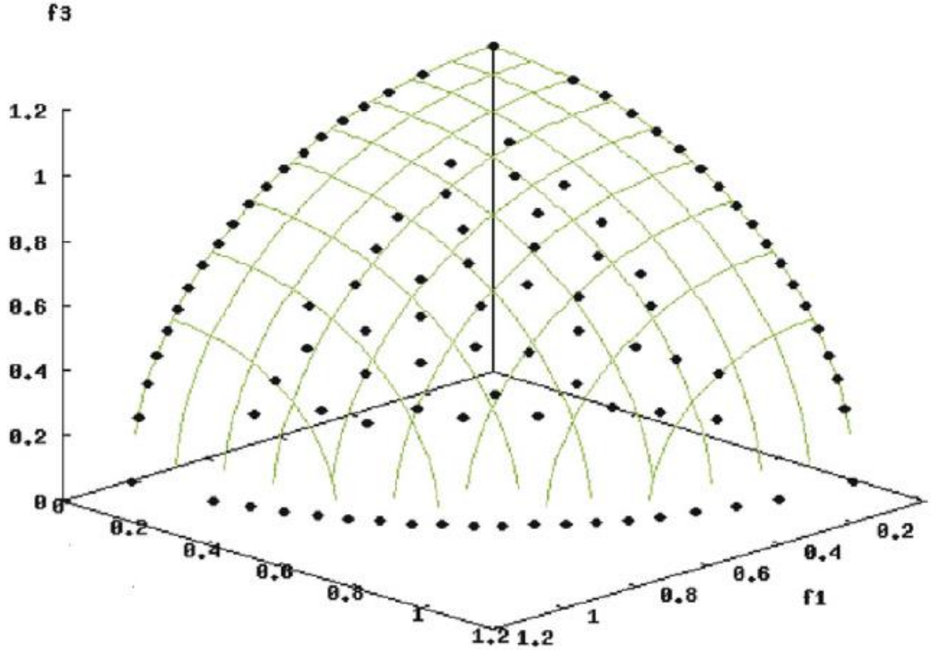
The best solution set for HV

DTLZ2

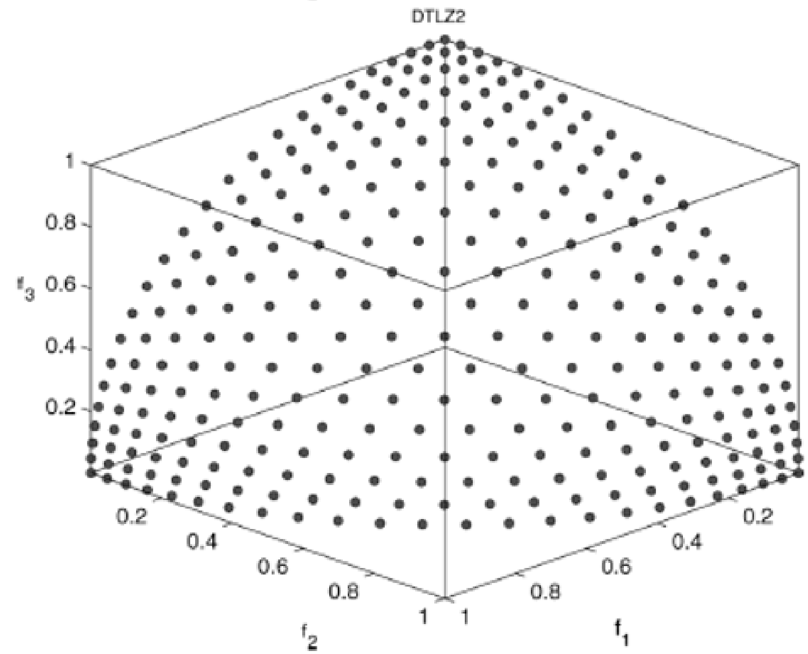


SMS-EMOA (2007)

Question: Which is better for DTLZ2?



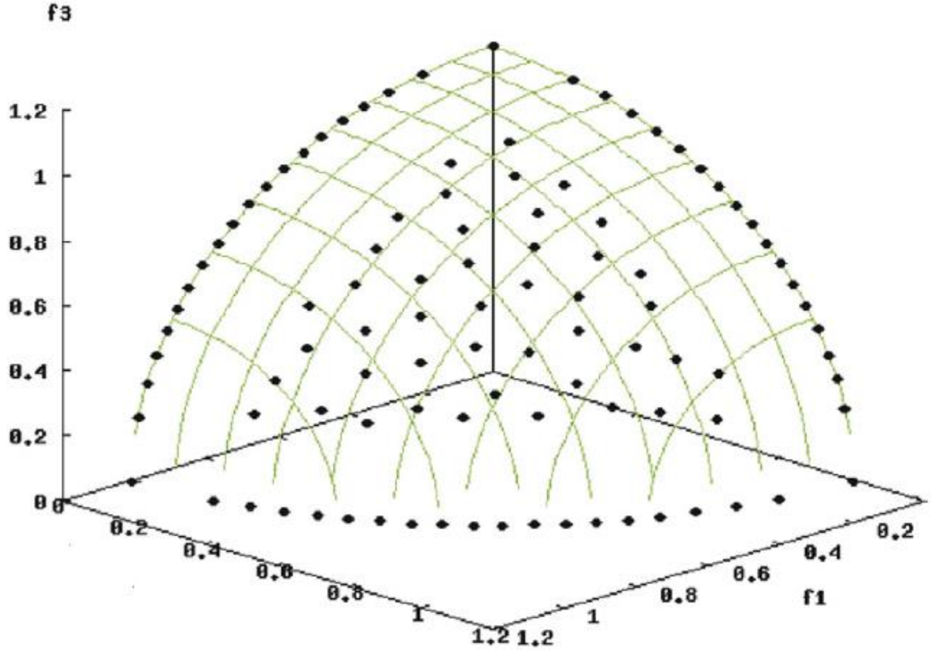
SMS-EMOA (2007)



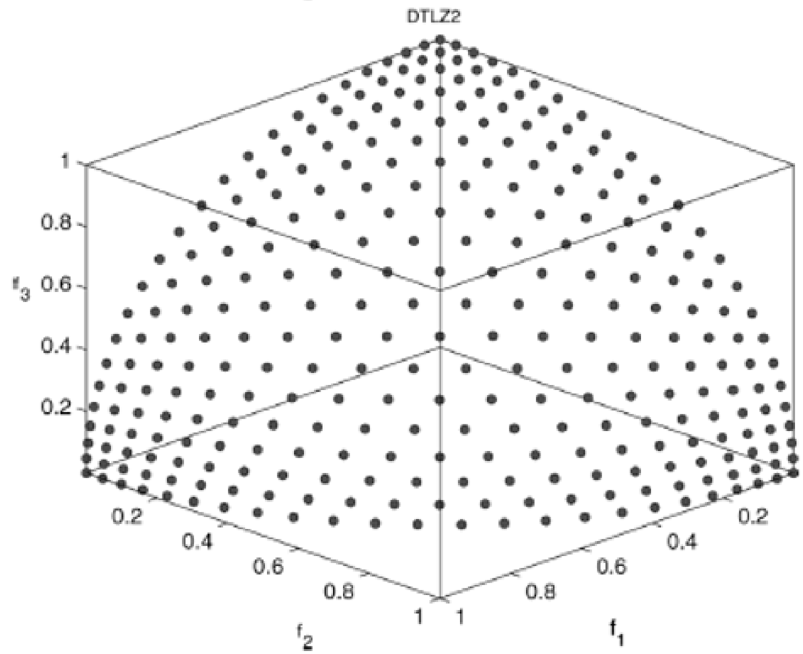
MOEA/D (2007)

Convergence: Almost the same (DTLZ2 is an easy test problem)
Maximum spread: The same (all the three extreme points are obtained.
Uniformity: The MOEA/D solution set looks clearly better.
HV-based comparison: The SMS-EMOA solution set is better.
IGD-based comparison: The MOEA/D solution set is better.

Question: Which is better for DTLZ2?



SMS-EMOA (2007)



MOEA/D (2007)

Convergence: Almost the same (DTLZ2 is an easy test problem)

Maximum spread: The same (all the three extreme points are obtained.

Uniformity: The MOEA/D solution set looks clearly better.

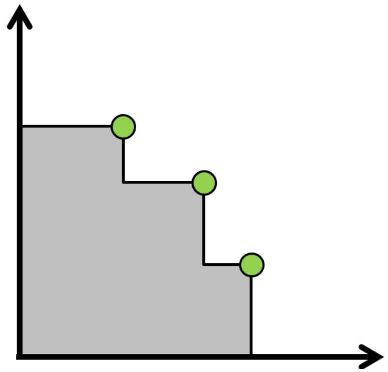
HV-based comparison: The SMS-EMOA solution set is better.

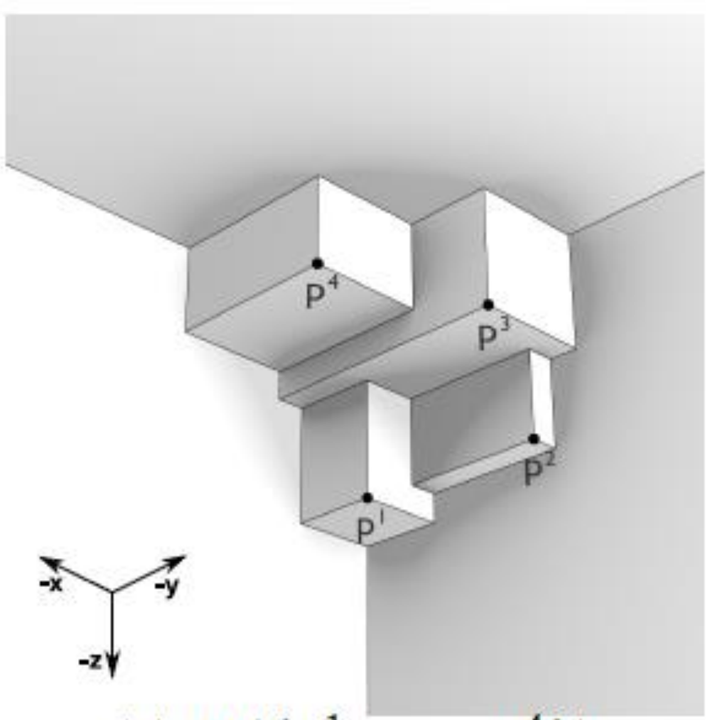
IGD-based comparison: The MOEA/D solution set is better.

Additional information may be needed (e.g., DM’s preference)

Three Major Difficulties of SMS-EMOA

1. Uniformity of Solutions
2. Large Computation Load for HVC Calculation
3. Reference Point Specification





(a) $H(\{p^1, \ldots, p^4\})$

Intuitive Visual Understanding

2-objective: Easy and intuitive.

3-objective: Somewhat Complicated but OK.

4-objective: Very difficult.

Calculation of HV and HVC Values

2-4 objectives: No problem

5-8 objectives: Increasing difficulty

9-10 objectives: Large sets are difficult.

More than 10 objectives: Difficult

Large sets with 10 or more objectives:

Exact calculation is almost impossible.
(e.g., 1000 solutions with 11 objectives).

Three Major Difficulties of SMS-EMOA

1. Uniformity of Solutions
2. Large Computation Load for HVC Calculation
3. Reference Point Specification

In SMS-EMOA Paper (EJOR 2007)

Test problems with 2 and 3 objectives

In MOEA/D Paper (IEEE TEVC 2007)

Test problems with 2, 3 and 4 objectives

In the SMS-EMOA paper, the large computation load was not a serious problem since SMS-EMOA was applied to multi-objective problems with 2 and 3 objectives.

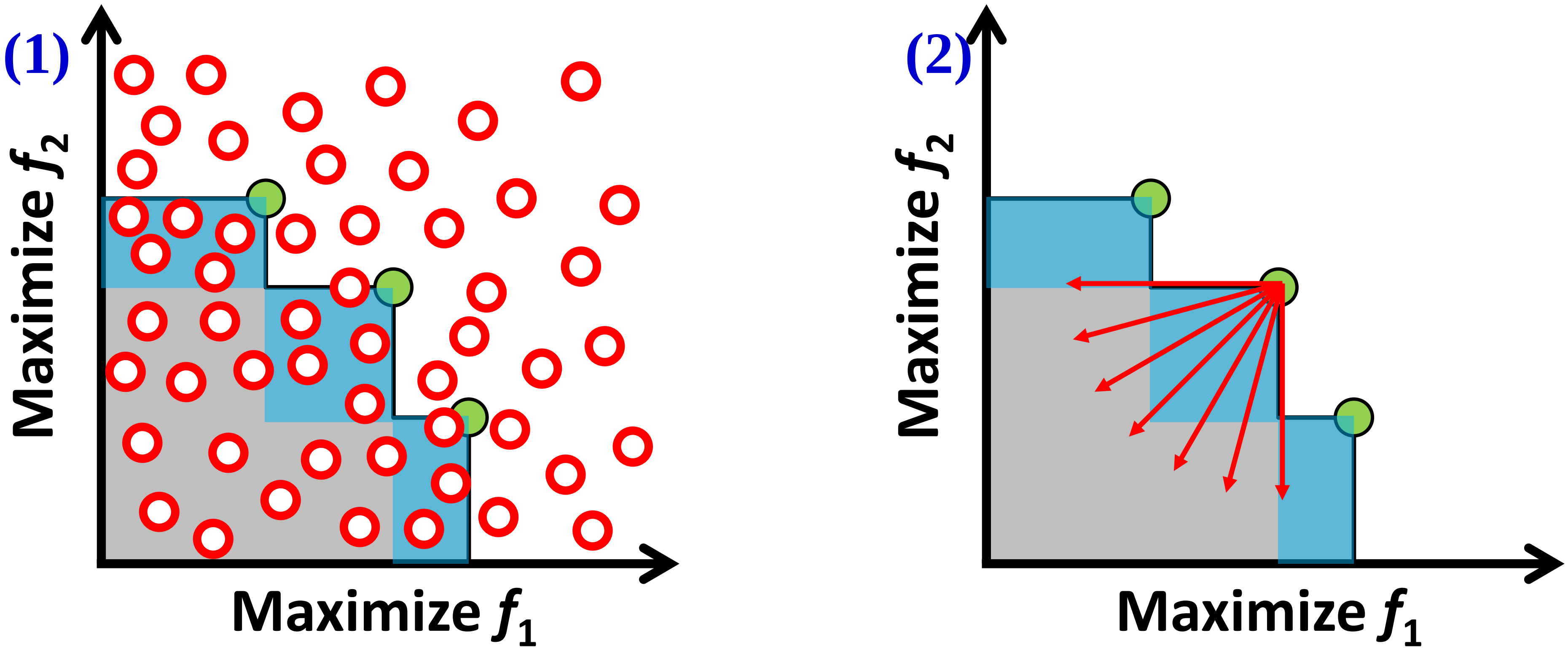
Comparison Results on Knapsack Problems (Ishibuchi et al., IEEE TEVC 2015)

TABLE II. RELATIVE AVERAGE HYPERVOLUME. THE REFERENCE POINT (0, 0, ..., 0) IS FAR FROM THE PARETO FRONT.

EMO Algorithm	2-500	4-500	6-500	8-500	10-500
NSGA-II	96.5	86.2	77.8	72.0	65.5
MOEA/D: WS	100.0	**100.0**	**100.0**	**100.0**	**100.0**
MOEA/D: Tchebycheff	100.7	99.7	94.0	90.1	87.7
MOEA/D: PBI (0.01)	99.9	99.7	99.8	**100.0**	99.7
MOEA/D: PBI (0.05)	99.4	98.6	98.5	98.4	98.5
MOEA/D: PBI (0.1)	98.8	96.9	96.6	96.0	95.9
MOEA/D: PBI (0.5)	92.7	82.7	79.6	77.0	74.3
MOEA/D: PBI (1.0)	96.1	78.1	68.0	66.0	63.0
MOEA/D: PBI (5)	100.9	89.3	73.8	67.4	61.9
SMS-EMOA	95.2	91.8	88.4	-	-
HypE	97.2	94.1	92.8 1)	93.9 2)	92.5
F100-NSGA-II	100.0	94.7	87.2	79.9	68.2
F90-NSGA-II	**101.0**	98.2	92.4	86.5	82.2

1) 50 runs, 2) 30 runs

2. Large Computation Load for HVC Calculation



Instead of exactly calculating the HV and HVC values, they are approximated by randomly generated points in (1) and uniformly generated lines in (2).

[1] J. Bader and E. Zitzler, "HypE: An algorithm for fast hypervolume-based many-objective optimization," *Evolutionary Computation*, 2011.

[2] K. Shang and H. Ishibuchi, "A new hypervolume-based evolutionary algorithm for many-objective optimization," IEEE Trans. on Evolutionary Computation (Early Access).

Comparison Results on Knapsack Problems (Ishibuchi et al., IEEE TEVC 2015)

TABLE II. RELATIVE AVERAGE HYPERVOLUME. THE REFERENCE POINT (0, 0, ..., 0) IS FAR FROM THE PARETO FRONT.

EMO Algorithm	2-500	4-500	6-500	8-500	10-500
NSGA-II	96.5	86.2	77.8	72.0	65.5
MOEA/D: WS	100.0	**100.0**	**100.0**	**100.0**	**100.0**
MOEA/D: Tchebycheff	100.7	99.7	94.0	90.1	87.7
MOEA/D: PBI (0.01)	99.9	99.7	99.8	**100.0**	99.7
MOEA/D: PBI (0.05)	99.4	98.6	98.5	98.4	98.5
MOEA/D: PBI (0.1)	98.8	96.9	96.6	96.0	95.9
MOEA/D: PBI (0.5)	92.7	82.7	79.6	77.0	74.3
MOEA/D: PBI (1.0)	96.1	78.1	68.0	66.0	63.0
MOEA/D: PBI (5)	100.9	89.3	73.8	67.4	61.9
SMS-EMOA	95.2	91.8	88.4	-	-
HypE	97.2	94.1	92.8 1)	93.9 2)	92.5
F100-NSGA-II	100.0	94.7	87.2	79.9	68.2
F90-NSGA-II	**101.0**	98.2	92.4	86.5	82.2

1) 50 runs, 2) 30 runs

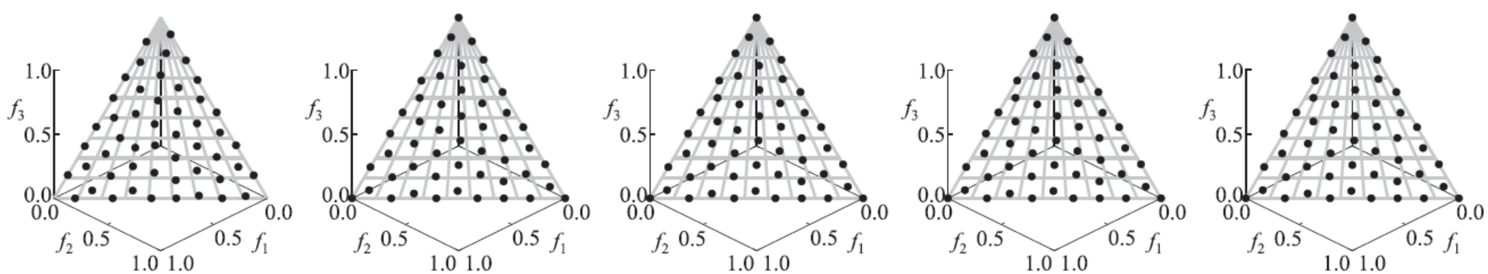
Three Major Difficulties of SMS-EMOA

3. Reference Point Specification

Current worst value + 1 (minimization)

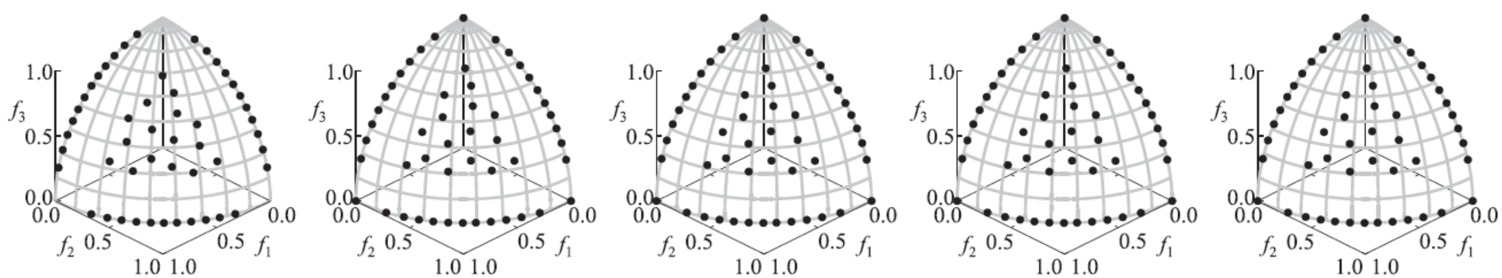
In the normalized objective space, DTLZ1: r = 3.0, DTLZ2: r = 2.0

Reference point specification is not an important problem since almost the same results are obtained from r = 1.1, 1.5, 2.0 and 10.



(a) $r = 1.0$. (b) $r = 1.1$. (c) $r = 1.5$. (d) $r = 2.0$. (e) $r = 10$.

Figure 1: Obtained solution sets for the three-objective normalized DTLZ1.



(a) $r = 1.0$. (b) $r = 1.1$. (c) $r = 1.5$. (d) $r = 2.0$. (e) $r = 10$.

Figure 8: Obtained solution sets for the three-objective DTLZ2 problem.

Ishibuchi et al., Evolutionary Computation, 2018

Three Major Difficulties of SMS-EMOA

3. Reference Point Specification

Current worst value + 1 (minimization)

This specification strongly depends on the scale of each objective.

(a) If the Pareto front is within $[0.00, 0.01]^m$, the reference point is far away from the Pareto front.

(b) If the Pareto front is within $[0, 10000]^m$, the reference point is very close to the Pareto front (almost the same as the nadir point).

DTLZ

(1) Pareto front range: $[0.0, 0.5]^m$ in DTLZ1, $[0, 1]^m$ in DTLZ2-4.

"Current worst value + 1" is a good specification.

(2) Pareto front shape: Triangular shape.

Reference point specification has almost no effect.

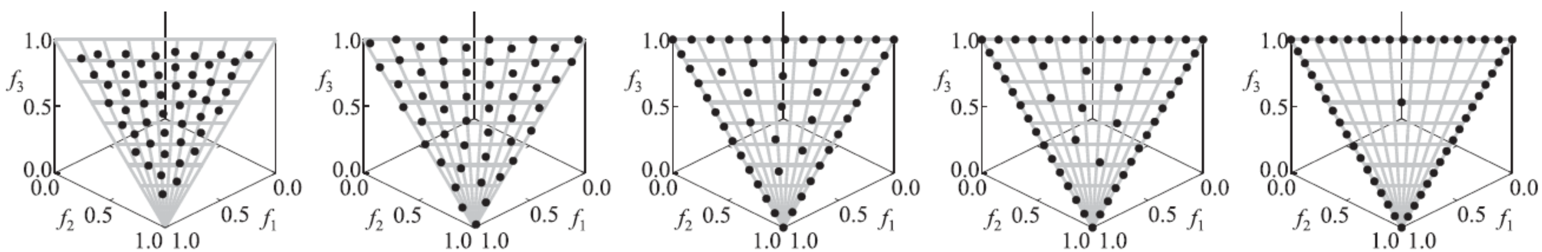
Three Major Difficulties of SMS-EMOA

3. Reference Point Specification

Current worst value + 1 (minimization)

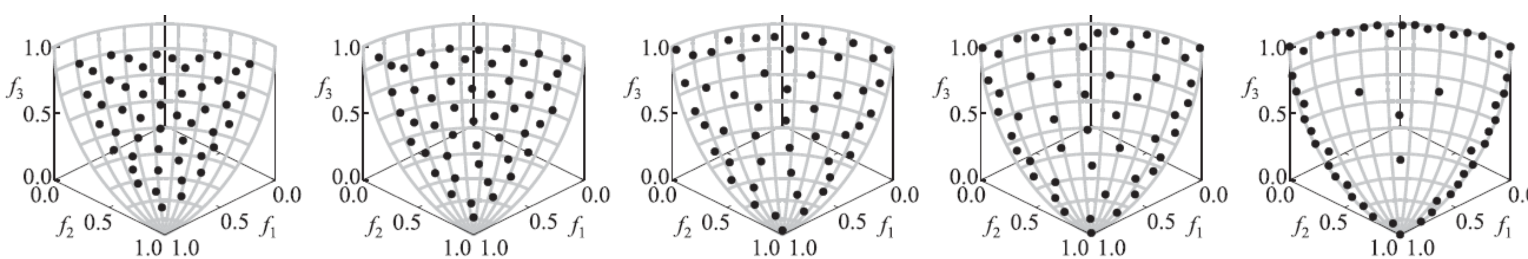
Minus-DTLZ1: r = almost 1.0 (1.0017)

Reference point specification has a large effect on the final results when the Pareto front is inverted triangular.



(a) $r = 1.0$. (b) $r = 1.1$. (c) $r = 1.5$. (d) $r = 2.0$. (e) $r = 10$.

Figure 2: Obtained solution sets for the three-objective normalized Minus-DTLZ1.



(a) $r = 1.0$. (b) $r = 1.1$. (c) $r = 1.5$. (d) $r = 2.0$. (e) $r = 10$.

Figure 9: Obtained solution sets for the three-objective Minus-DTLZ2 problem.

Ishibuchi et al., Evolutionary Computation, 2018

Most of them are not severe for the three-objective DTLZ

Most of them are not severe in their original papers in 2007.

Difficulties in MOEA/D

1. A new solution can be better than all neighbors.
2. A new solution can be far away from the neighborhood.
3. Uniform weight vectors are not always appropriate.
4. No inside weight vectors for many-objective problems.

Difficulties in SMS-EMOA

1. Uniform solutions are not always obtained.
2. Large Computation Load for HVC Calculation is needed.
3. Reference Point Specification is difficult.

They are difficult for other problems (not in the 2007 papers)

Difficulties in MOEA/D

1. A new solution can be better than all neighbors.
2. A new solution can be far away from the neighborhood.
3. **Uniform weight vectors are not always appropriate.**
   WFG (normalization), Problems with concave Pareto fonts,
   Problems with inverted triangular Pareto fronts.
4. **No inside weight vectors for many-objective problems.**
   Many-objective problems with more than 5 or 6 objectives.

Difficulties in SMS-EMOA

1. **Uniform solutions are not always obtained.**
   DTLZ2-4 and WFG with nonlinear Pareto fronts.
2. **Large computation load for HVC calculation is needed.**
   Many-objective problems with more than 7 or 8 objectives.
3. **Reference point specification is difficult.**
   Problems with inverted triangular Pareto fronts.

Difficulties in MOEA/D due to Many Objectives

1. A new solution can be better than all neighbors.
2. A new solution can be far away from the neighborhood.
3. Uniform weight vectors are not always appropriate.
4. **No inside weight vectors for many-objective problems.**
   ==> Two layer method in NSGA-III (2014)

Difficulties in SMS-EMOA due to Many-Objectives

1. Uniform solutions are not always obtained.
2. **Large computation load for HVC calculation is needed.**
   ==> Approximate HV calculation in HypE (2011)
3. Reference point specification is difficult.

These difficulties have become clear by increasing the number of objectives in DTLZ and WFG (i.e., scalable test problems).

Relation between Test Problems and EMO Algorithms

Around 1995
Simple two-objective problems (no need of strong convergence)
Non-elitist Pareto dominance-based algorithms: MOGA, NPGA, NSGA
These algorithms work well on simple two-objective problems

Around 2000
Difficult two-objective problems: ZDT
Elitist Pareto dominance-based algorithms: SPEA, SPEA2, NSGA-II
These algorithms work well on two-objective problems

2000s
Easy scalable test problems (the number of objectives can be arbitrarily specified: *m*-objective problems): **DTLZ and WFG**
Decomposition-based algorithms: **MOEA/D**
Indicator-based algorithms: **SMS-EMOA**
These algorithms work well on problems with 3-6 objectives
They have difficulties for many-objective problems (e.g., 10 objectives)

2010s
Many-objective algorithms: HypE, NSGA-III

Highly-Cited Papers published in 2010s

Many-Objective EMO Algorithms

HypE J. Bader and E. Zitzler, "**HypE: An algorithm for fast hypervolume-based many-objective optimization**," *Evolutionary Computation*, vol. 19 no. 1, pp. 45-76, Spring 2011.

NSGA-III K. Deb and H. Jain, "**An evolutionary many-objective optimization algorithm using reference-point-based non-dominated sorting approach, Part I: Solving problems with box constraints,**" *IEEE Trans. on Evolutionary Computation*, vol. 18, no. 4, pp. 577-601, August 2014.

Many-Objective Optimization

Single-Objective Optimization: Maximize $f(\boldsymbol{x})$

Multi-Objective Optimization: Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ \ldots,\ f_k(\boldsymbol{x})$

- Multi-Objective Optimization:

Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x})$

Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x})$

- Many-Objective Optimization:

Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x}),\ f_4(\boldsymbol{x})$

Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x}),\ f_4(\boldsymbol{x}),\ f_5(\boldsymbol{x})$

Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x}),\ f_4(\boldsymbol{x}),\ f_5(\boldsymbol{x}),\ f_6(\boldsymbol{x})$

Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x}),\ f_4(\boldsymbol{x}),\ f_5(\boldsymbol{x}),\ f_6(\boldsymbol{x}),\ f_7(\boldsymbol{x})$

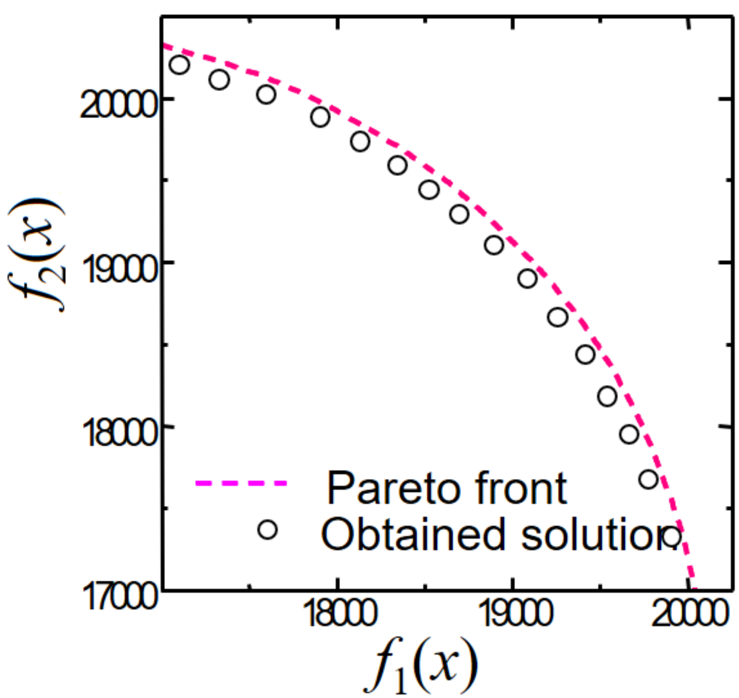
Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x}),\ f_4(\boldsymbol{x}),\ f_5(\boldsymbol{x}),\ f_6(\boldsymbol{x}),\ f_7(\boldsymbol{x}),\ f_8(\boldsymbol{x})$

Q. What is the difference between three-objective and four-objective problems?

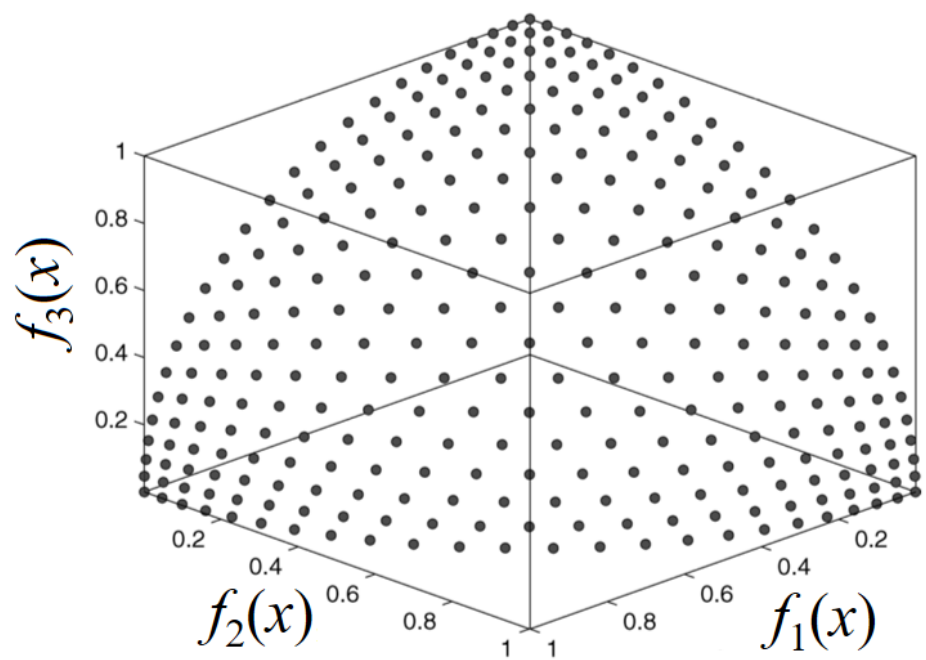
Multi-Objective Problem: Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x})$

Many-Objective Problem: Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x}),\ f_4(\boldsymbol{x})$

One clear difference is the difficulty of the visualization:

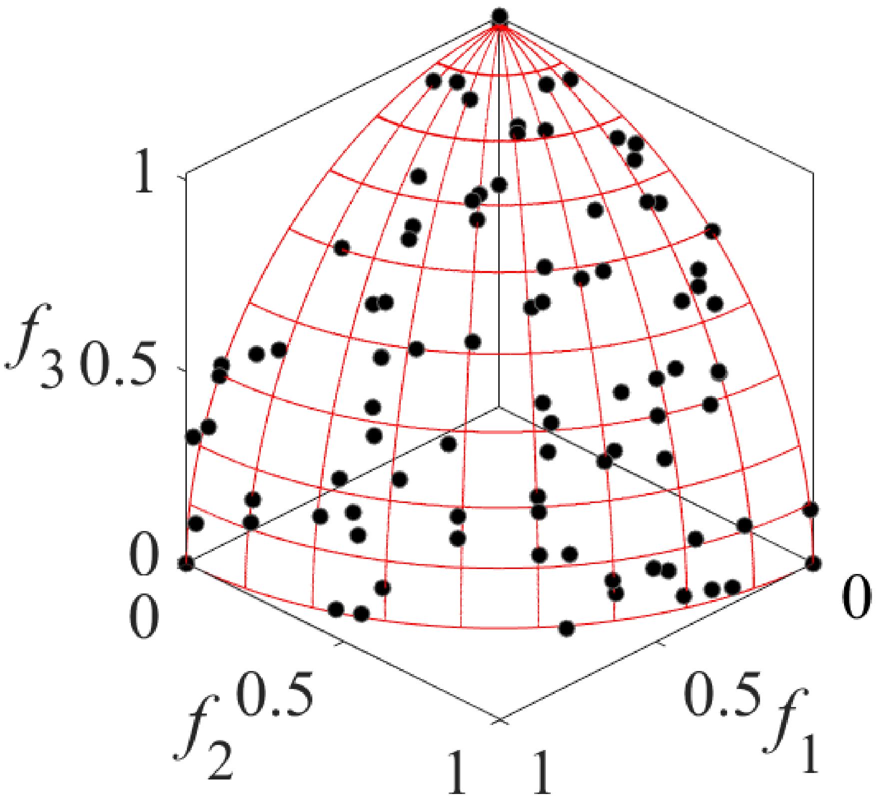


Two-Objective Problem

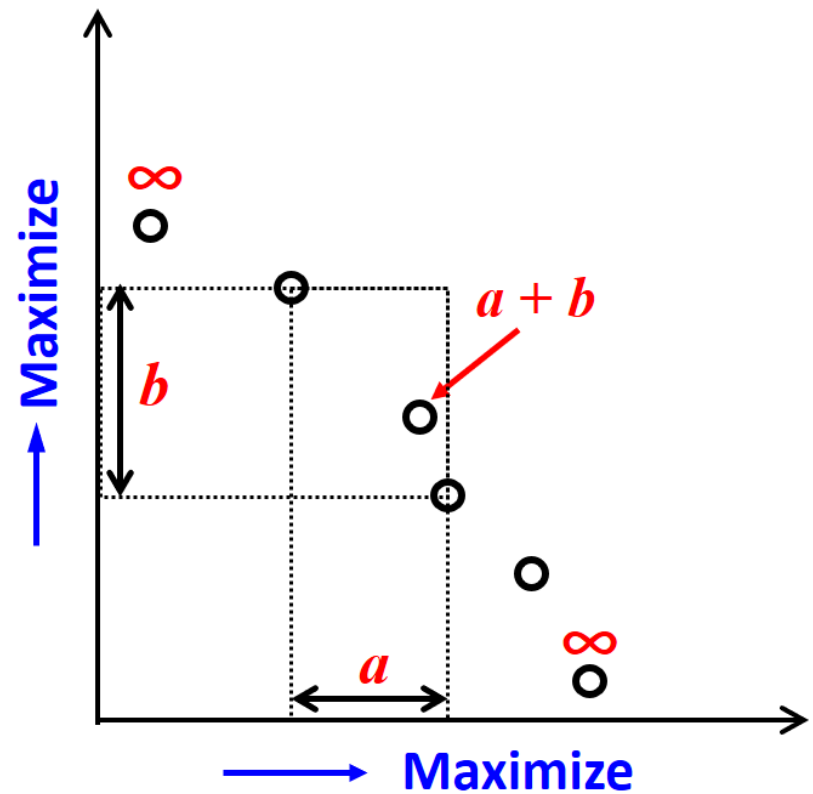


Three-Objective Problem

For NSGA-II, there is a large difference between two-objective and three-objective problems



Final population (obtained solutions).



Crowding distance calculation.

It is difficult for NSGA-II to find a well-distributed solution set for a three-objective problem.

Relation between Test Problems and EMO Algorithms

2000s

Easy scalable test problems (the number of objectives can be arbitrarily specified: *m*-objective problems): DTLZ and WFG
Decomposition-based algorithms: MOEA/D
Indicator-based algorithms: SMS-EMOA
These algorithms work well on problems with 3-6 objectives
They have difficulties for many-objective problems (e.g., 10 objectives)

2010s

Many-objective algorithms: HypE, NSGA-III
These algorithms work well on easy many-objective problems: DTLZ and WFG

Late 2010s

Scalable test problems with inverted triangular Pareto front: Minus-DTLZ and Minus-WFG (2017)
Adaptable decomposition-based algorithms
Reference point specification in HV-base algorithms

2020s

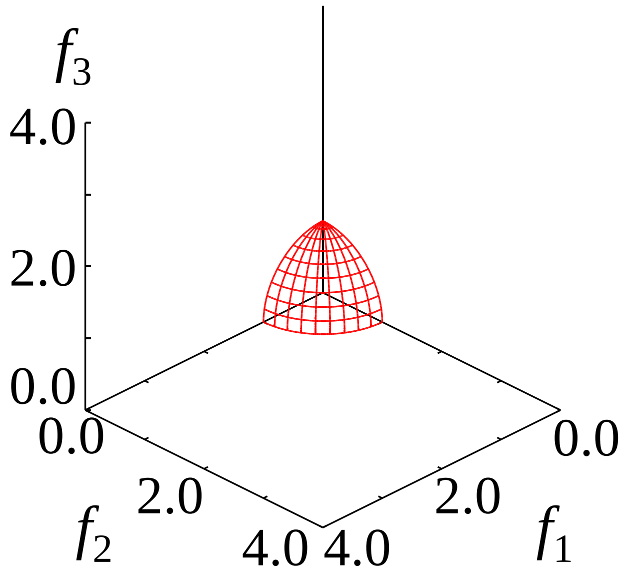
Difficult (and realistic) many-objective test problems
New algorithms for difficult (and realistic) many-objective problems

Minus-DTLZ and Minus-WFG

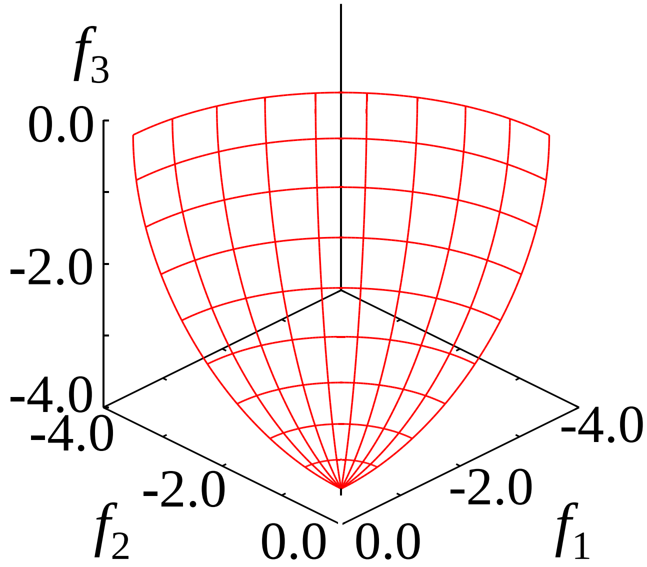
Ishibuchi et al. IEEE TEVC (2017)

(- 1) x DTLZ and (- 1) x WFG Test Problems:

All objectives are multiplied by “   1”, which has the same effect as changing from “minimization” to “maximization”.



Pareto front of DTLZ2

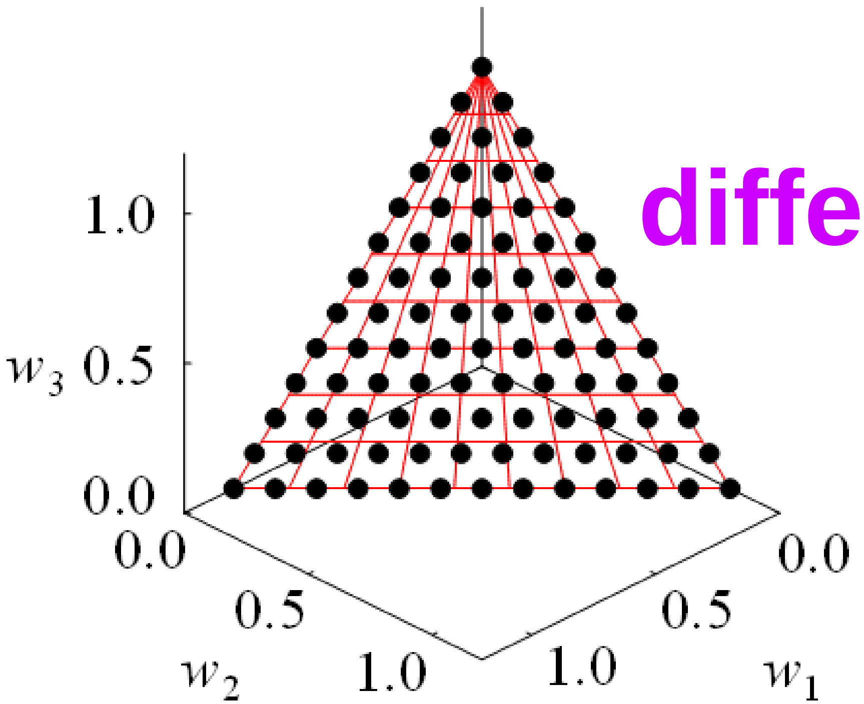


Pareto front of (-1) x DTLZ2: Minus-DTLZ2

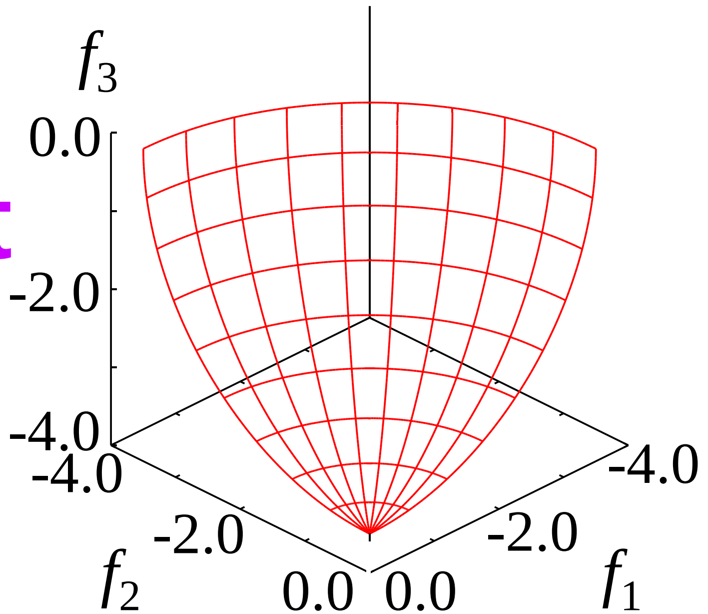
Formulation of Minus-DTLZ and Minus-WFG

Original Aim of the Formulations:

To examine the performance of MOEA/D when the Pareto front shape is totally different from the weight vector distribution.



different



Weight Vectors

Pareto front (Minus-DTLZ 2)

Relation between Test Problems and EMO Algorithms

2010s

Many-objective algorithms: HypE, NSGA-III

These algorithms work well on easy many-objective problems:

DTLZ and WFG

Late 2010s

Scalable test problems with inverted Pareto front:

Minus-DTLZ and Minus-WFG (2017)

A number of difficulties have become clear in MOEA/D, SMS-EMOA, NSGA-III and other many-objective algorithms as well as test problems.

==> Adaptable decomposition-based algorithms

==> Reference point specification in HV-base algorithms

==> New test problem generation

2020s

Difficult (and realistic) many-objective test problems

New algorithms for difficult (and realistic) many-objective problems

Current Hot Topics in EMO

(1) Evolutionary Many-Objective Optimization

Multi-Objective Problems: Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ ...,\ f_m(\boldsymbol{x})$

Multi-Objective Problems:

$$\text{Maximize } f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x})$$

$$\text{Maximize } f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x})$$

Many-Objective Problems:

$$\text{Maximize } f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x}),\ f_4(\boldsymbol{x})$$

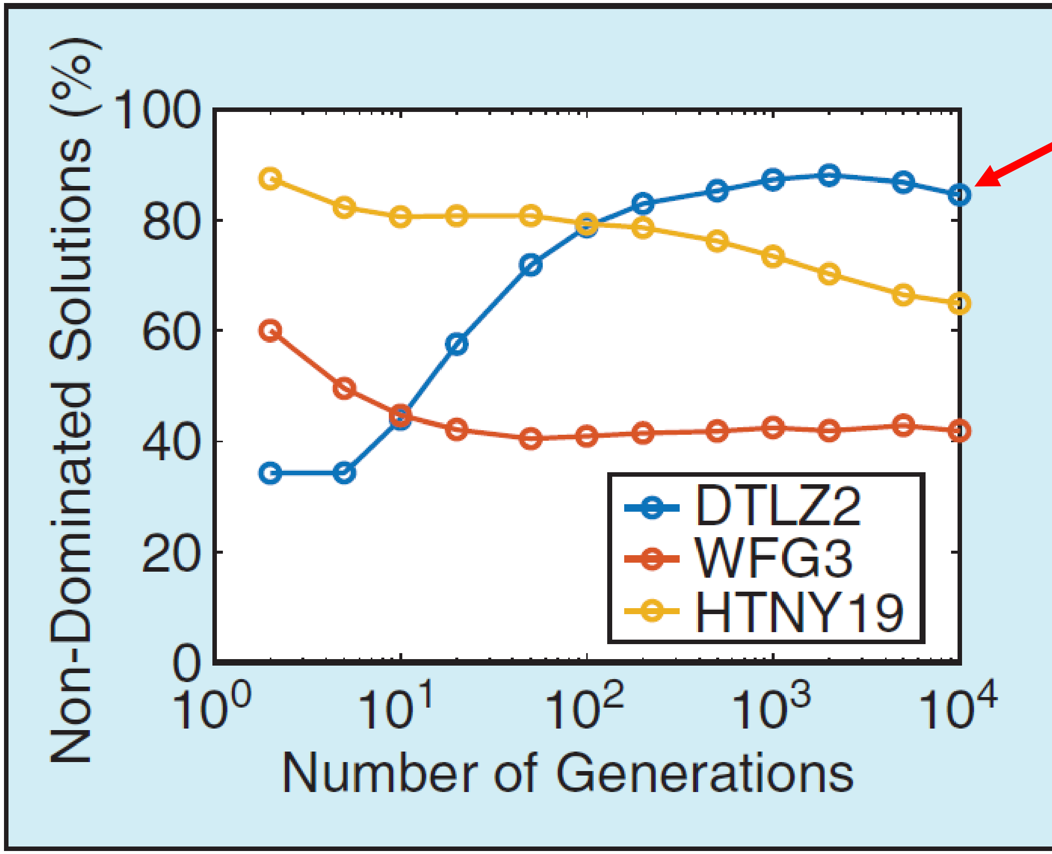
$$\text{Maximize } f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x}),\ f_4(\boldsymbol{x}),\ f_5(\boldsymbol{x})$$

$$\text{Maximize } f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ f_3(\boldsymbol{x}),\ f_4(\boldsymbol{x}),\ f_5(\boldsymbol{x}),\ f_6(\boldsymbol{x}),\ ....$$

When an optimization problem has many objectives (e.g., 10 objectives), almost all solutions in the current population are non-dominated.

==> We cannot use Pareto dominance as a solution evaluation criterion.

==> We have a huge number of good solutions.



Among 2,750,000 solutions, 84.6% are non-dominated.

Percentage of non-dominated solutions among examined solutions by MOEA/D-Tch on each 10-objective test problem (Population size: 275). Fig. 14 of Ishibuchi et al.: “Difficulties in Fair Performance Comparison of Multi-Objective Evolutionary Algorithms”, IEEE CIM Feb 2022 Issue.

Current Hot Topics in EMO

(2) Large-Scale Multi- and Many-Objective Optimization

$$\text{Maximize} \;\; f_1(\boldsymbol{x}),\; f_2(\boldsymbol{x}),\; ...,\; f_m(\boldsymbol{x})$$

The number of decision variables is huge (e.g., 1000, 10000, ...)
==> Multi-objective search in a high-dimensional search space.

- Standard Test Problems have 10-20 decision variables.
- Standard EMO algorithms were designed for those test problems.

Thus, some special techniques are needed for efficient search to handle large-scale multi- and many-objective optimization.

Current Hot Topics in EMO

(3) Expensive Multi- and Many-Objective Optimization

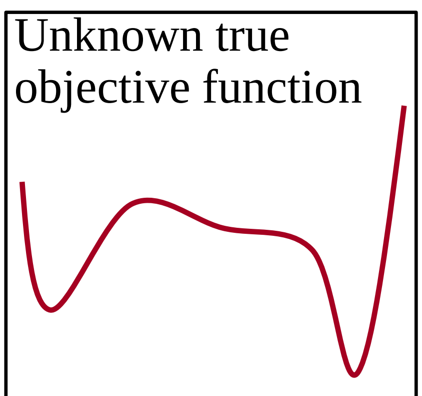
$$\text{Maximize } f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ ...,\ f_m(\boldsymbol{x})$$

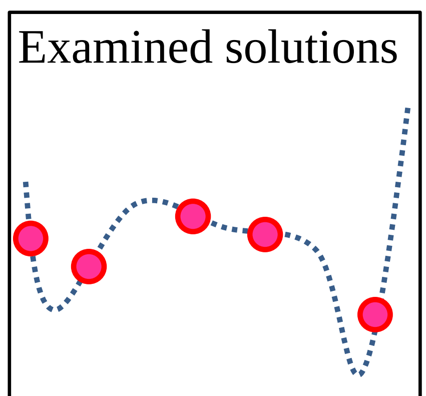
Solution evaluation is very expensive and/or time consuming.

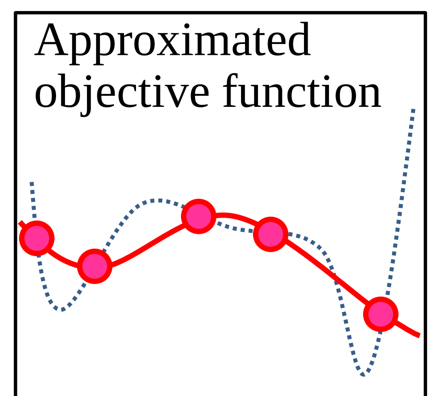
==> We can examine only a small number of solutions (e.g., 500)

Surrogate-Based Optimization:

To approximate each objective function by a machine learning technique (e.g., neural network) using a small number of examined solutions, and to apply an EMO algorithm to a multi-objective problem with the approximated objectives.







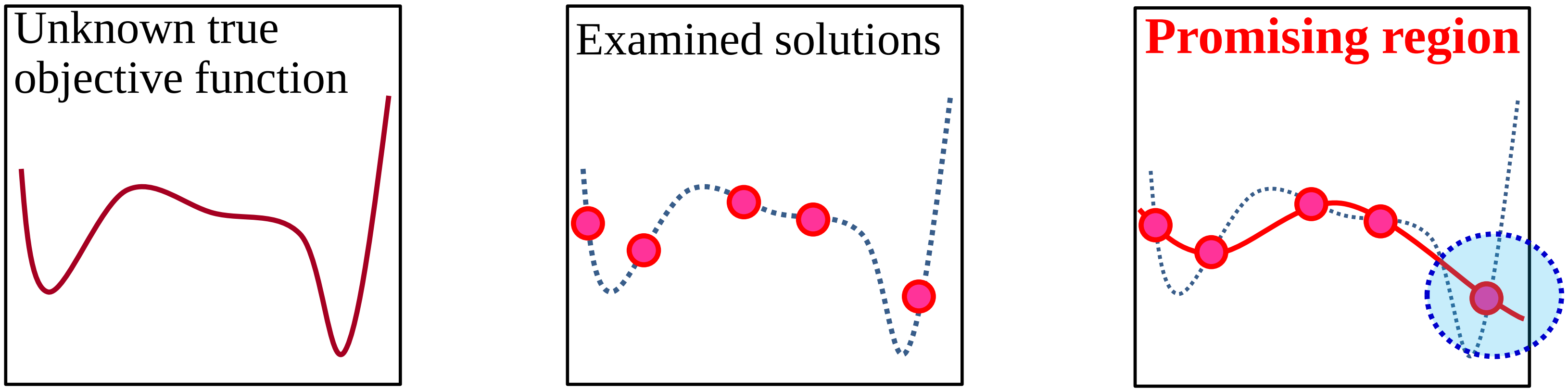
Current Hot Topics in EMO

(3) Expensive Multi- and Many-Objective Optimization

Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ ...,\ f_m(\boldsymbol{x})$

Solution evaluation is very expensive and/or time consuming.
==> We can examine only a small number of solutions (e.g., 500)

Surrogate-Based Optimization:
To approximate each objective function by a machine learning technique (e.g., neural network) using a small number of examined solutions, and to apply an EMO algorithm to a multi-objective problem with the approximated objectives.



Current Hot Topics in EMO

(3) Expensive Multi- and Many-Objective Optimization

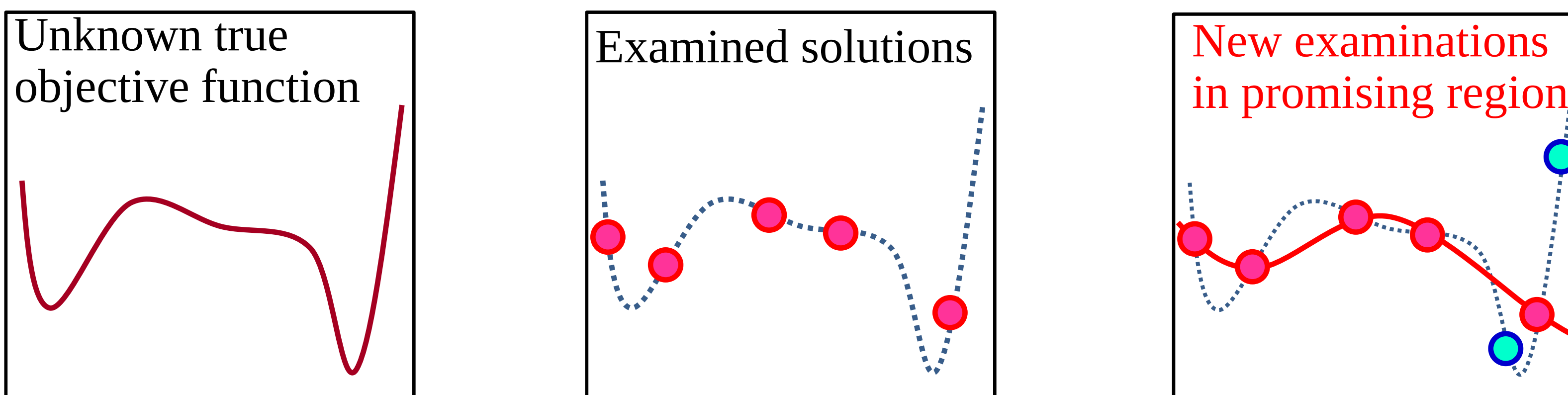
Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ ...,\ f_m(\boldsymbol{x})$

Solution evaluation is very expensive and/or time consuming.

==> We can examine only a small number of solutions (e.g., 500)

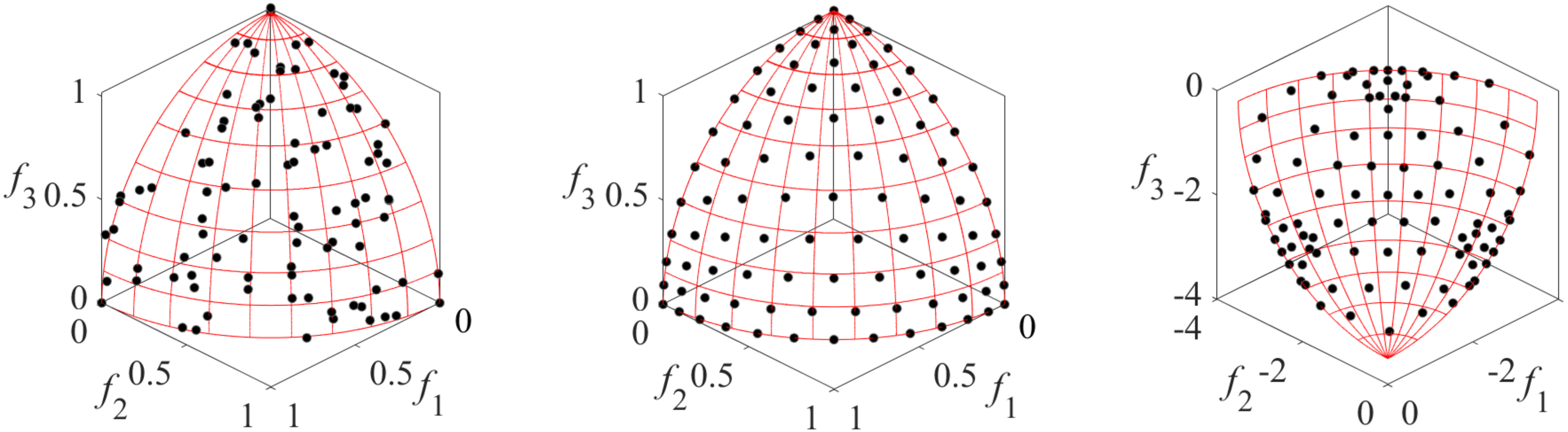
Surrogate-Based Optimization:

To approximate each objective function by a machine learning technique (e.g., neural network) using a small number of examined solutions, and to apply an EMO algorithm to a multi-objective problem with the approximated objectives.



Current Hot Topics in EMO

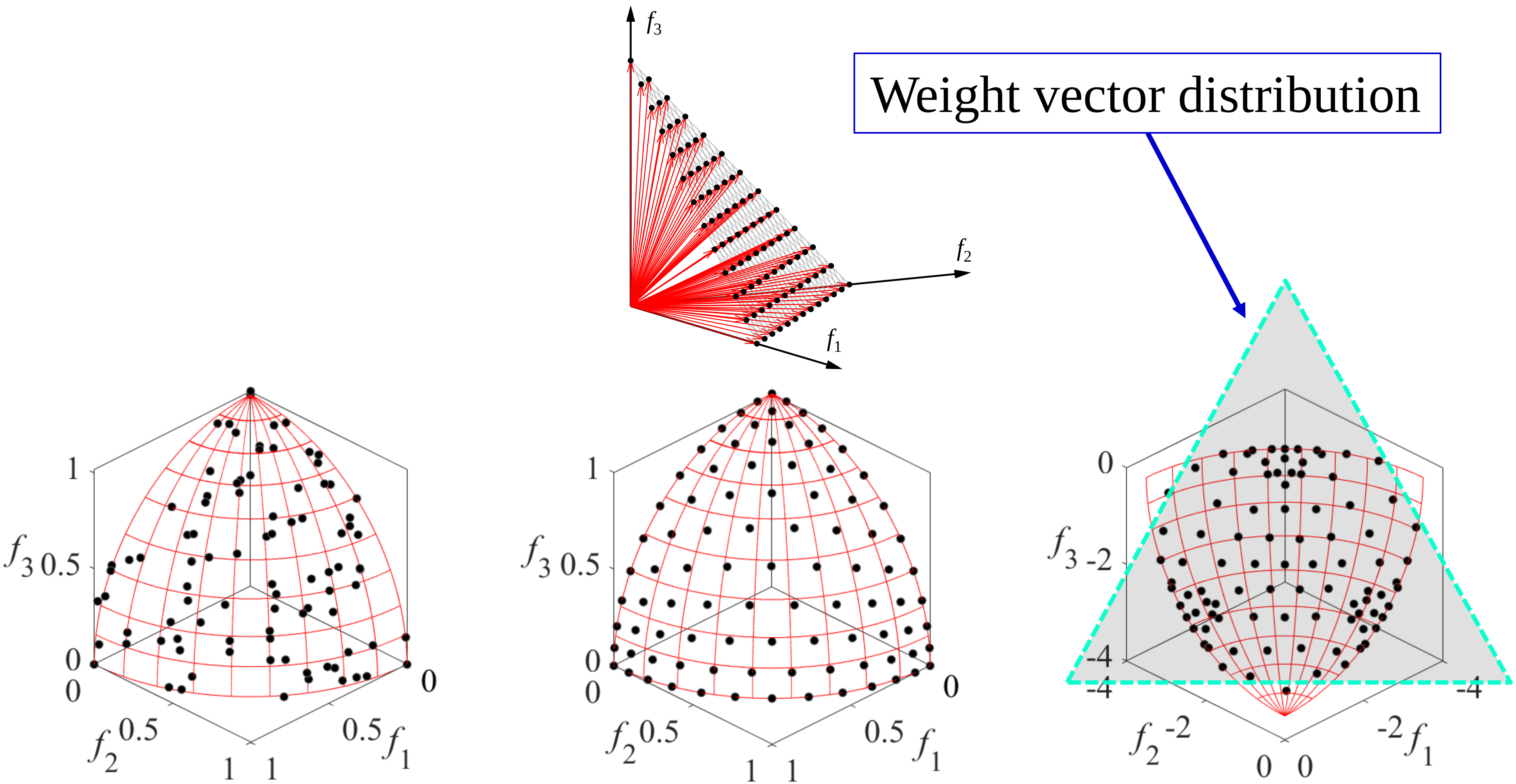
(4) Weight vector (reference vector) adaptation in MOEA/D



NSGA-II on DTLZ2 **MOEA/D on DTLZ2** **MOEA/D on Minus-DTLZ2**

Current Hot Topics in EMO

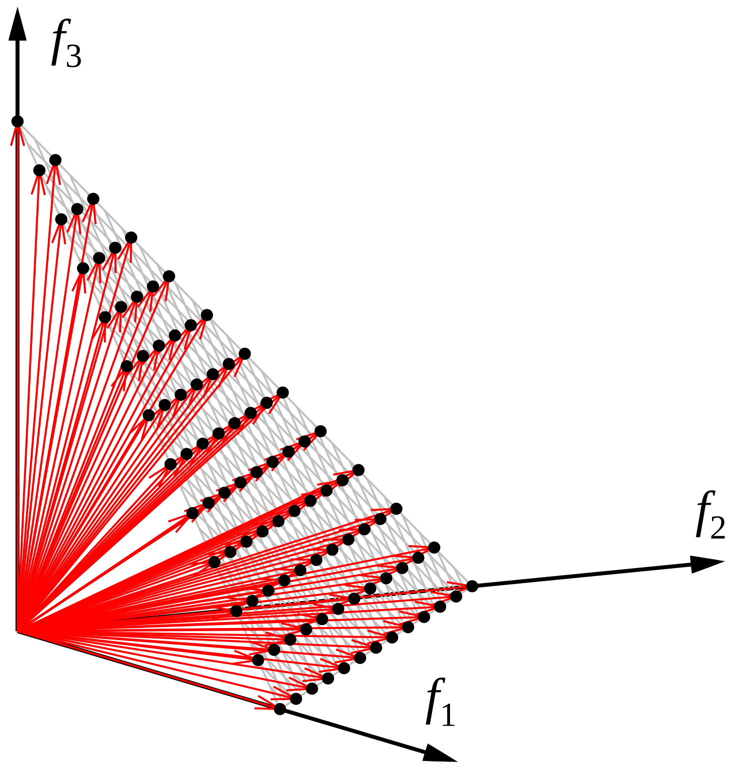
(4) Weight vector (reference vector) adaptation in MOEA/D



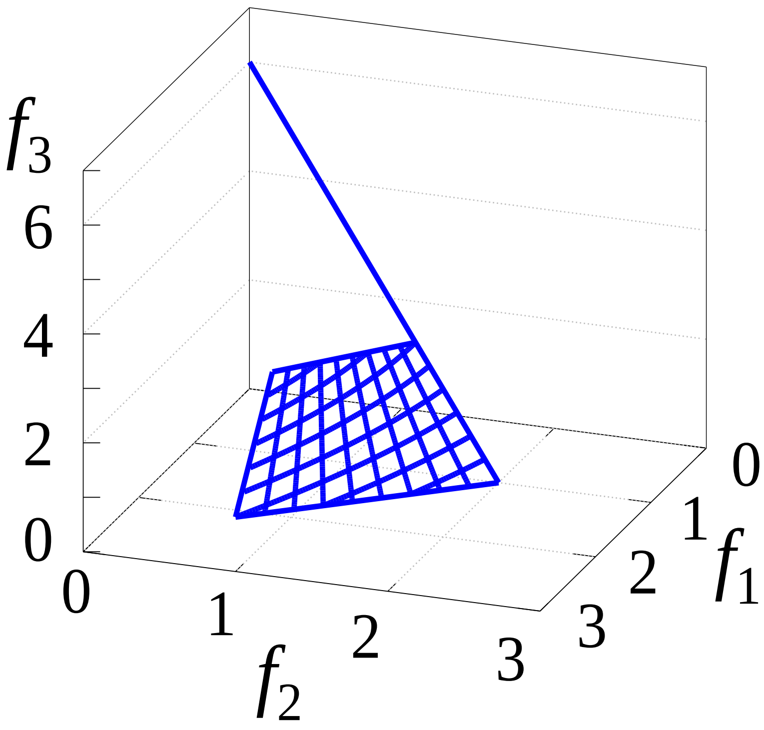
NSGA-II on DTLZ2   MOEA/D on DTLZ2   MOEA/D on Minus-DTLZ2

Current Hot Topics in EMO

(4) Weight vector (reference vector) adaptation in MOEA/D



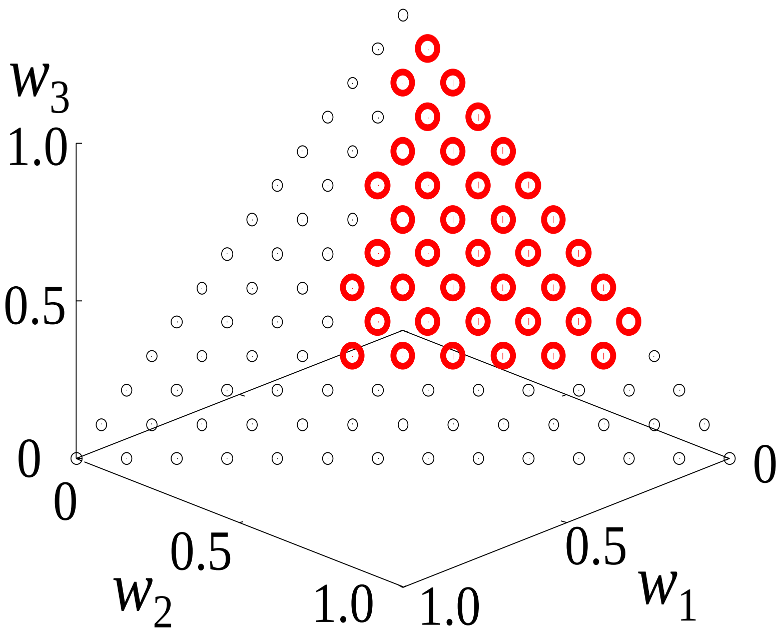
Weight Vector Distribution



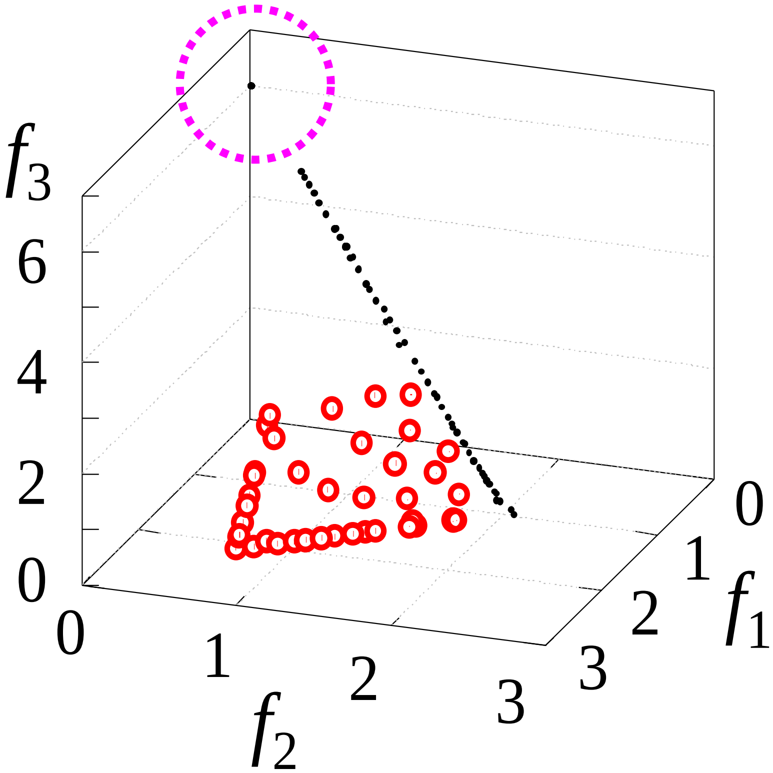
Pareto front of WFG3

Current Hot Topics in EMO

(4) Weight vector (reference vector) adaptation in MOEA/D



Weight Vector Distribution



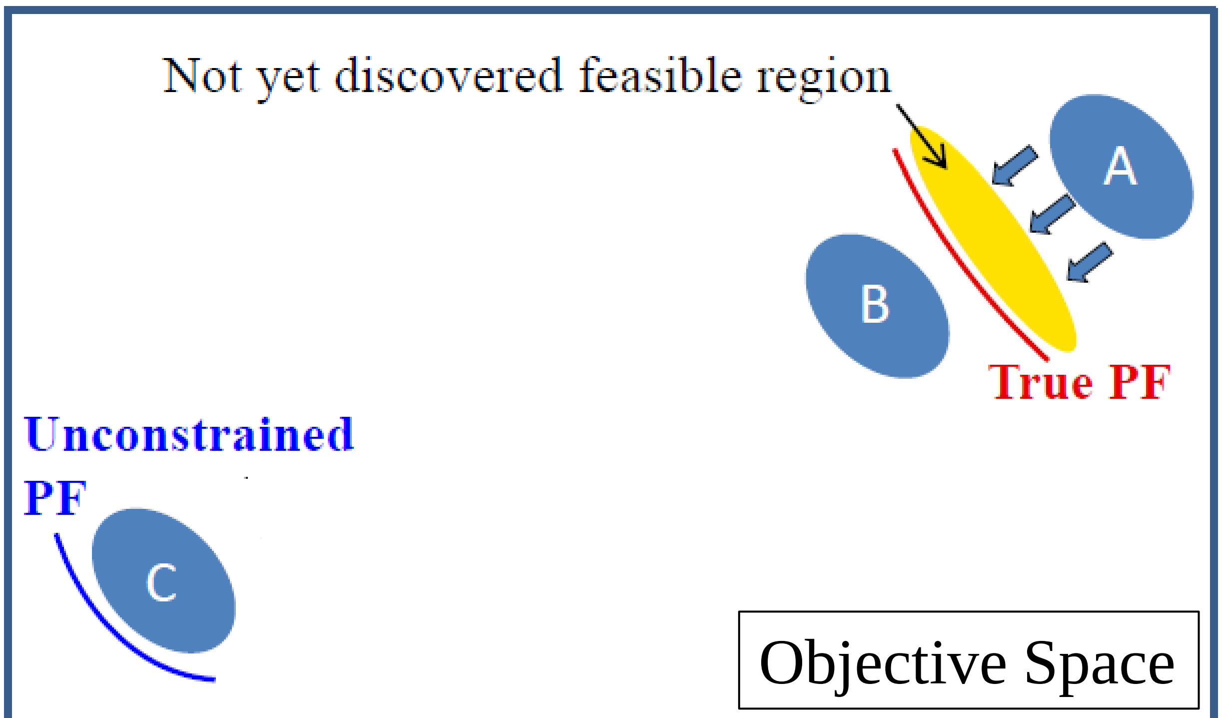
MOEA/D on WFG3

Current Hot Topics in EMO

(5) Constrained Multi- and Many-Objective Optimization

$$\text{Maximize } f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ \ldots,\ f_m(\boldsymbol{x})$$
$$\text{subject to } g_i(\boldsymbol{x}) \geq 0,\ i = 1, 2, \ldots, K$$

An interesting issue: How to utilize infeasible solutions for efficient search
(How to utilize populations B and C to pull Population A towards the PF)

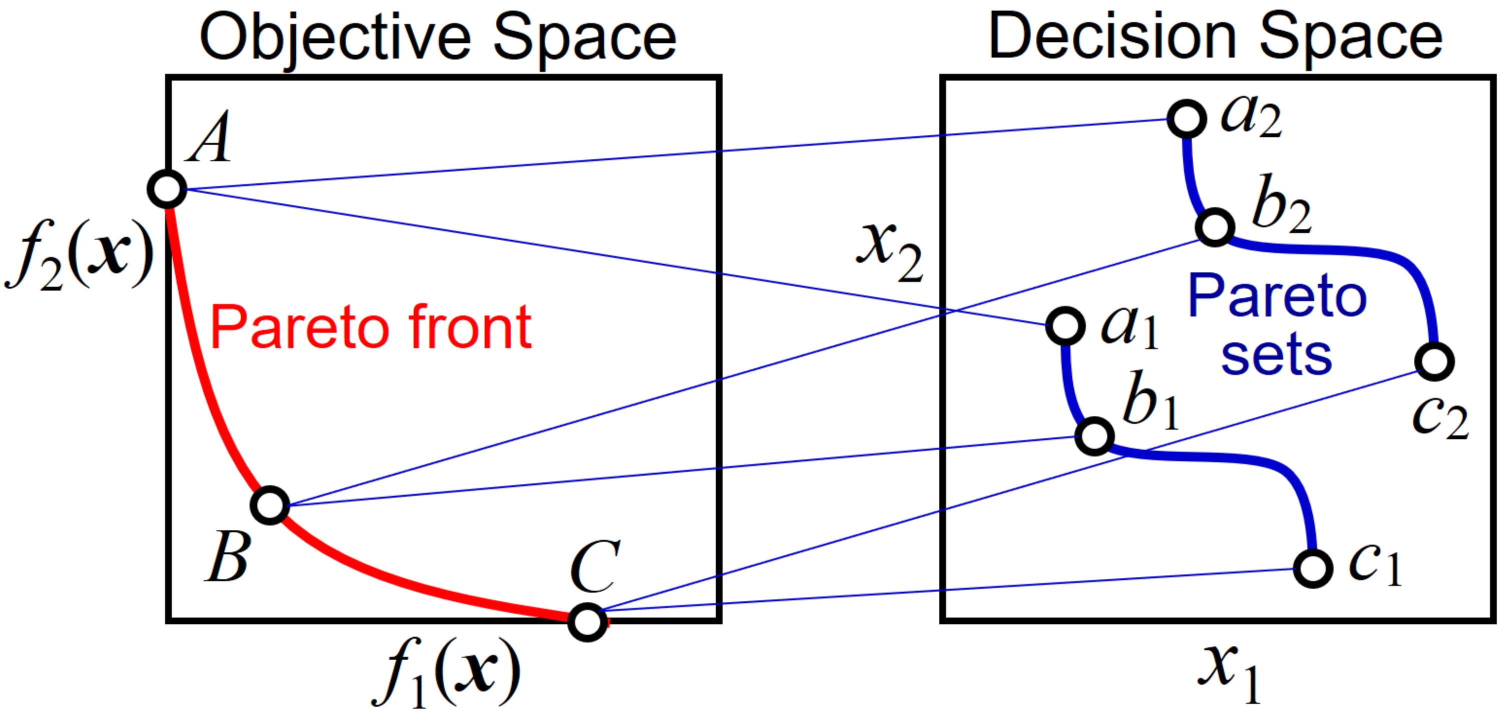


Current Hot Topics in EMO

(6) Multi-Modal Multi- and Many-Objective Optimization

Problem Definition: Each point on the Pareto front corresponds to different Pareto optimal solutions.

Search Goal: To find well-distributed points on the Pareto front and to find all the corresponding Pareto optimal solutions.



Current Hot Topics in EMO

(7) Utilization of Stored Solutions in Unbounded Archive

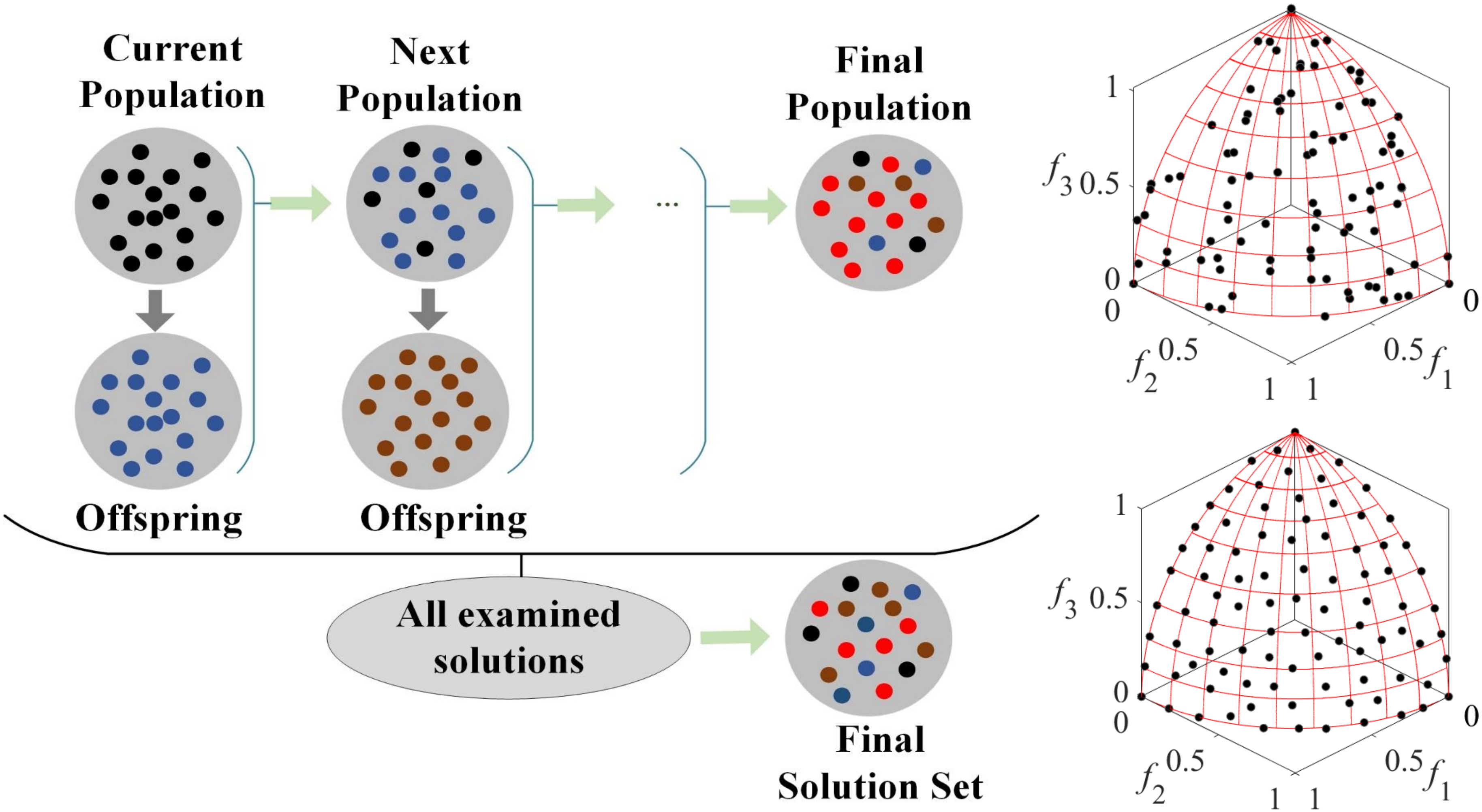
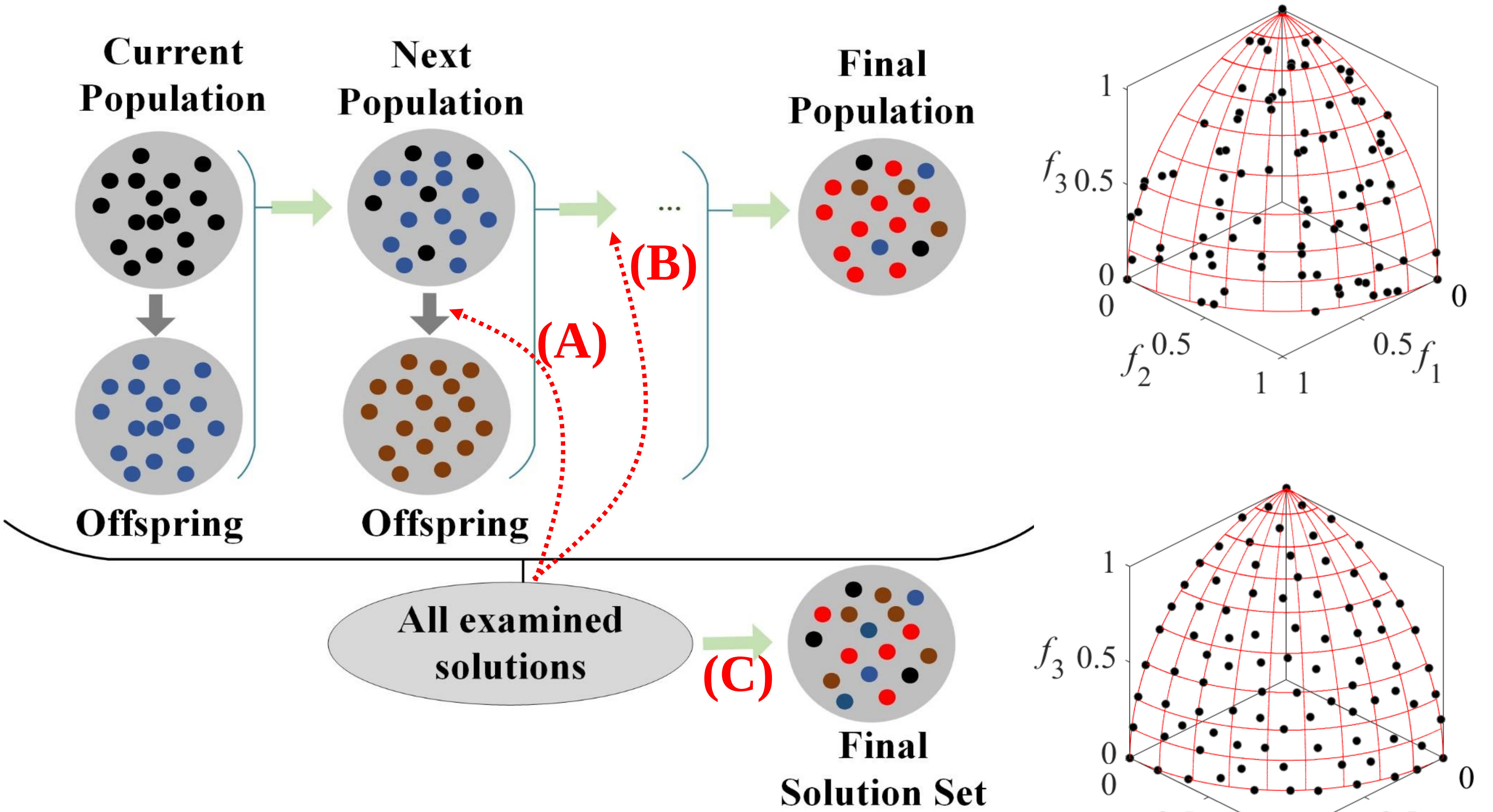


Fig. 4 of Ishibuchi et al., "A New Framework of Evolutionary Multi-Objective Algorithms with an Unbounded External Archive", Proc. of ECAI 2020.

New EMO Framework



Interesting Research Topics:

(A) How to utilize the stored solutions to generate new solutions.

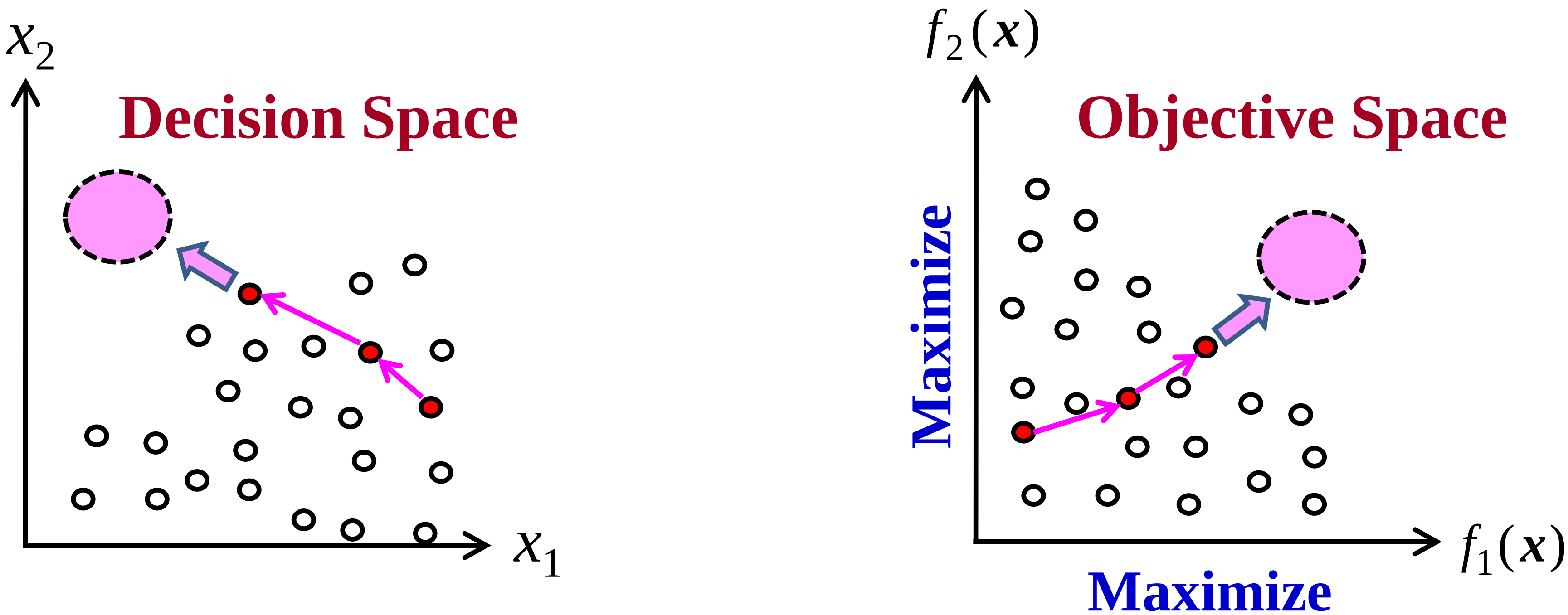
(B) How to utilize the stored solutions to choose the next population.

(C) How to choose the final solution set from the stored solutions.

(A) How to utilize the stored solutions to generate new solutions.

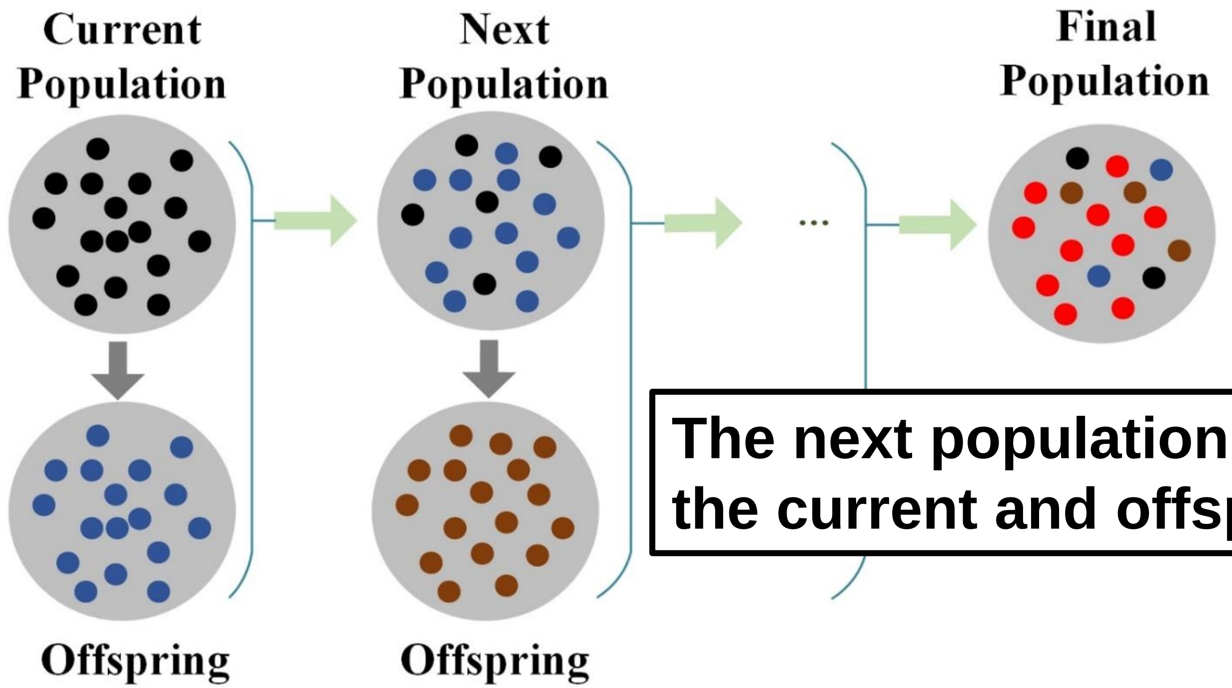
A simple idea:

Use of good moves in the previous generations to create new solutions.



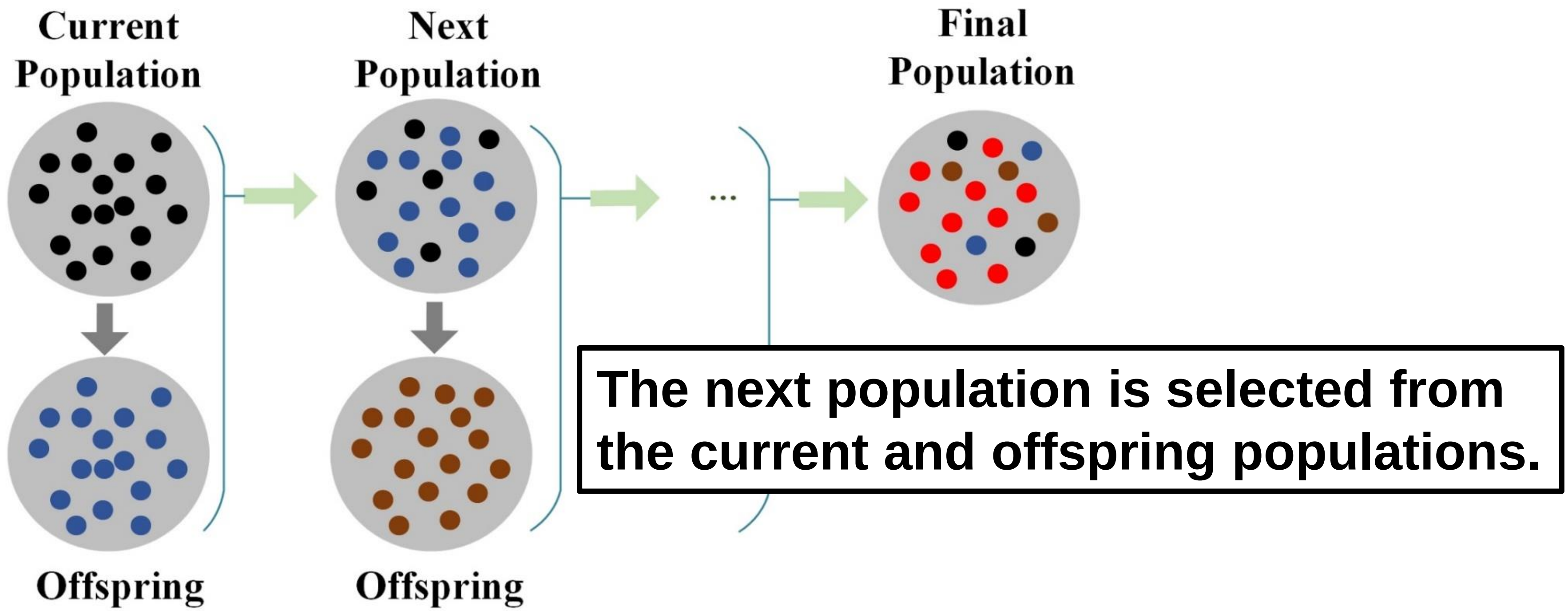
[1] S Mittal, DK Saxena, K Deb, ED Goodman: A Learning-based Innovized Progress Operator for Faster Convergence in Evolutionary Multi-objective Optimization, *ACM Transactions on Evolutionary Learning and Optimization* (2022).

[2] L. Chen, L. M. Pang, and H. Ishibuchi, "New Solution Creation Operator in MOEA/D for Faster Convergence," *Proc. of PPSN 2022* (2022)



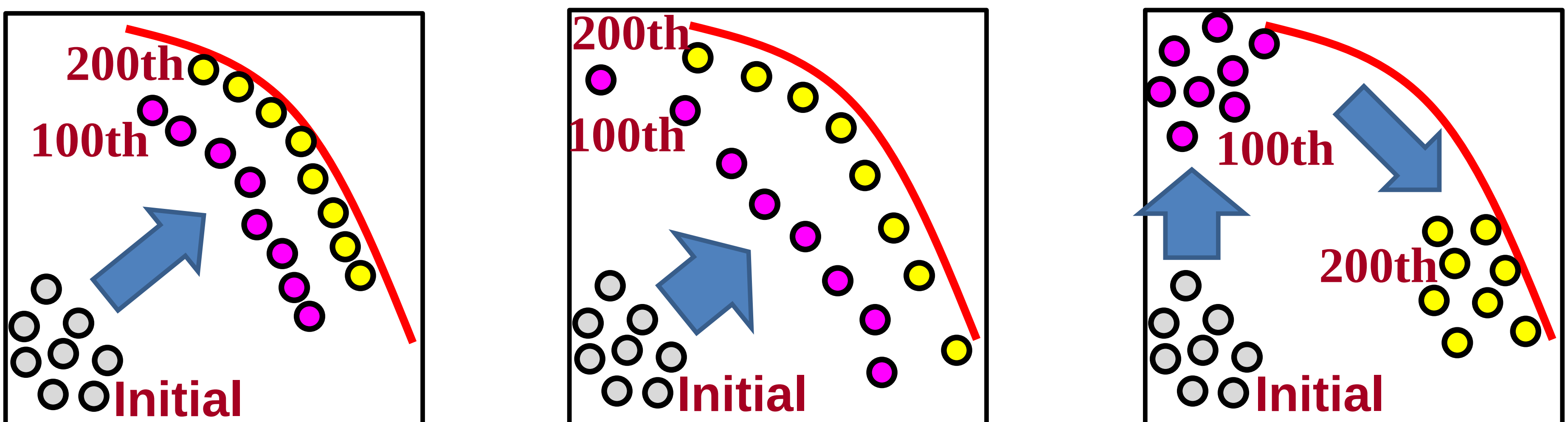
The next population is selected from the current and offspring populations.

Difficulty (2) in this framework:
The current population should be always a good solution set (because an EMO algorithm is used under various termination conditions).

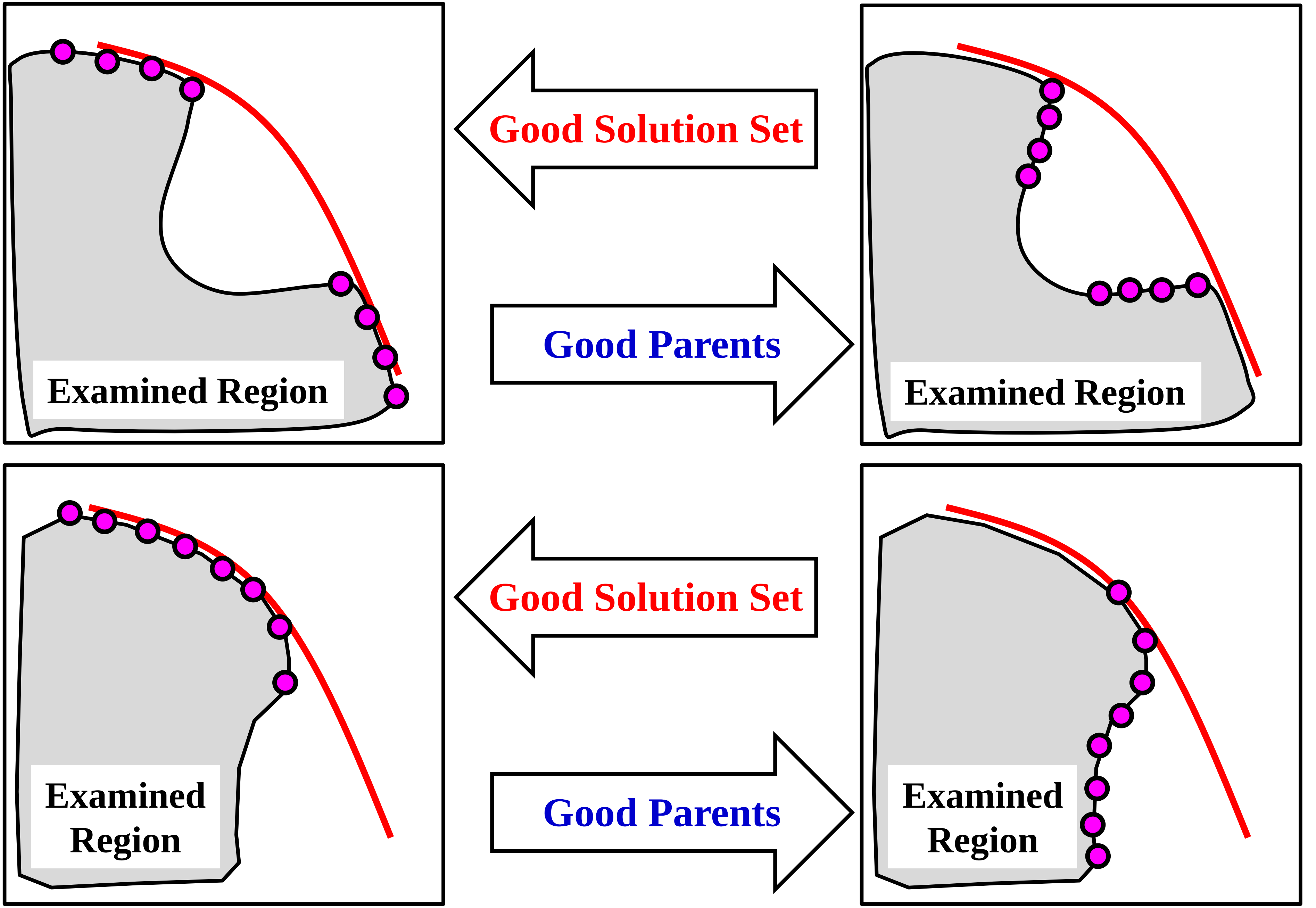


Difficulty in this framework:

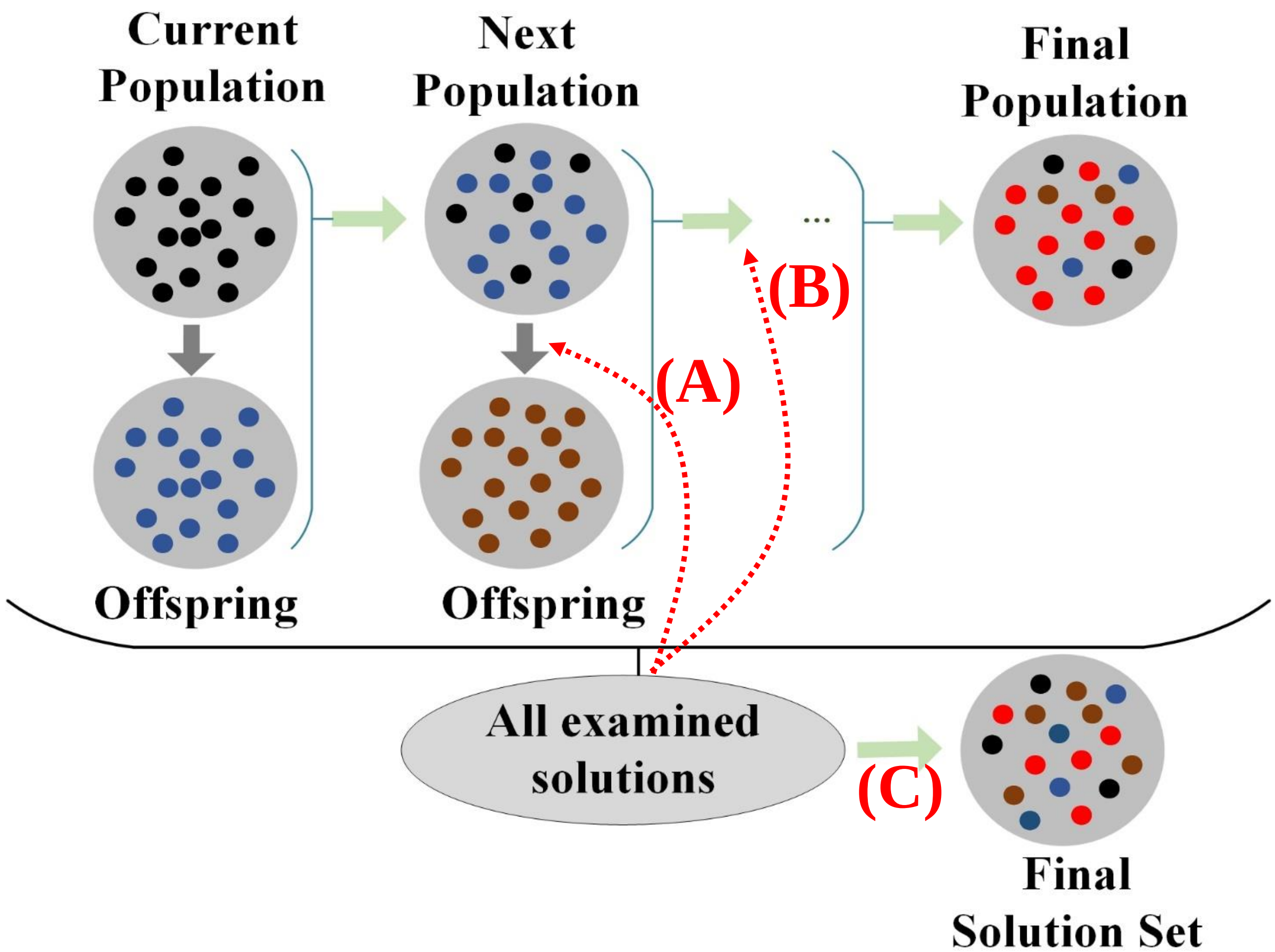
The current population should be always a good solution set (because an EMO algorithm is used under various termination conditions).

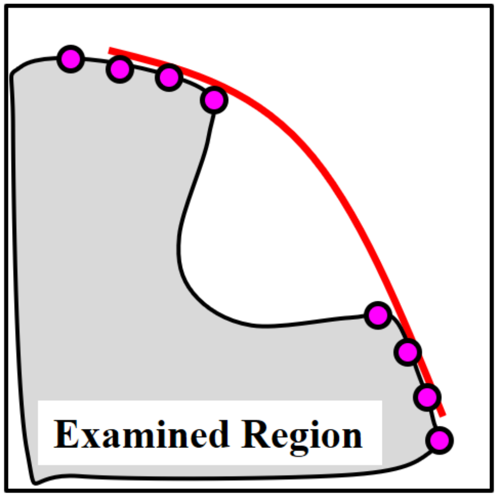


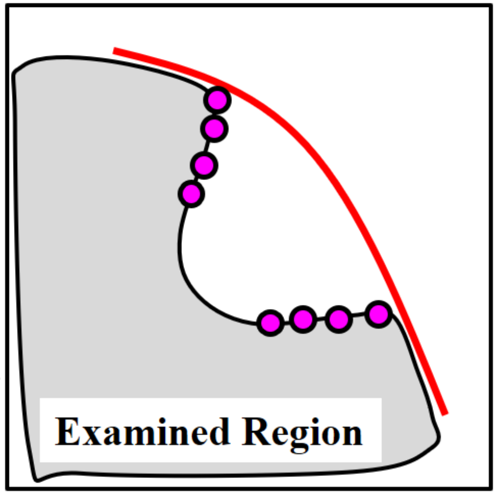
If some regions are very difficult to search:



New EMO Framework (Not new: This was used in some studies)





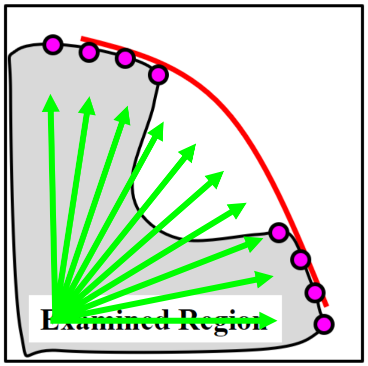


Interesting Research Topics:

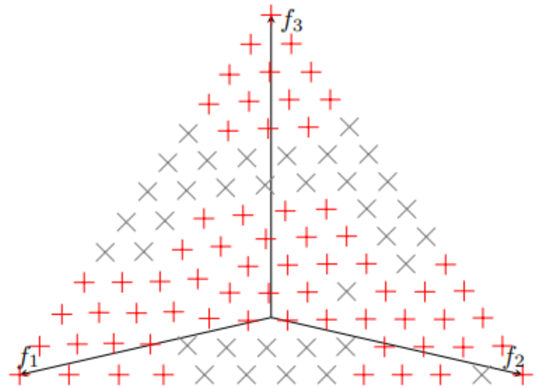
(A) How to utilize the stored solutions to generate new solutions.
(B) How to utilize the stored solutions to choose the next population.
(C) How to choose the final solution set from the stored solutions.

K. Deb et al., “Identifying Pareto Fronts Reliably Using a Multi-Stage Reference-vector-based Framework,” *IEEE TEVC* (2023).

First stage: Standard search.
Some parts are not obtained.

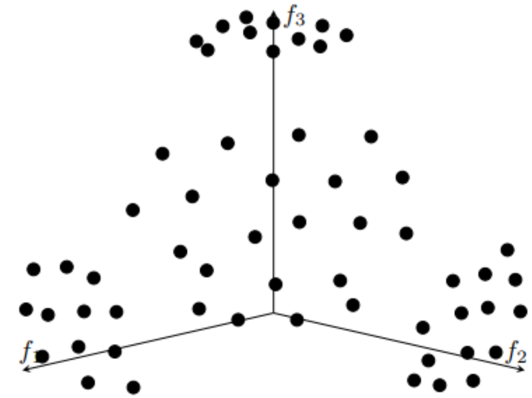


(a) RV set R_1 for Stage 1



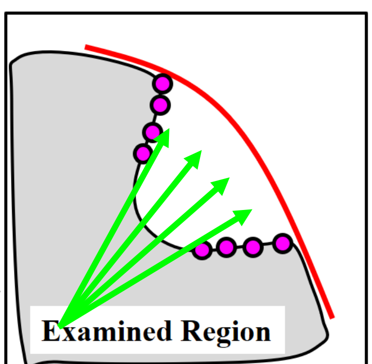
$+ R_{1c}$: ARVs $\times R_{1\bar{c}}$: IRVs

(b) Stage 1 ND set

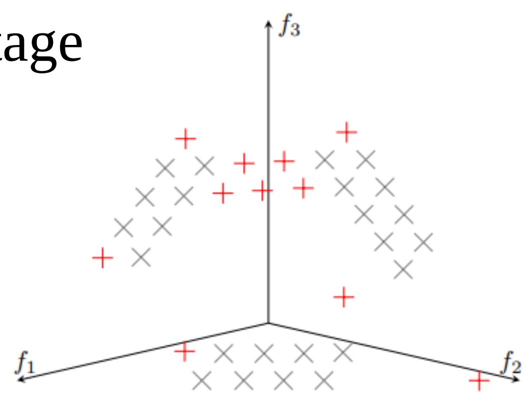


$\bullet\, S_{1c}^{T1}$ solution set

Second stage: Focused search.
Reference vectors in the regions with no solutions in the first stage are used.

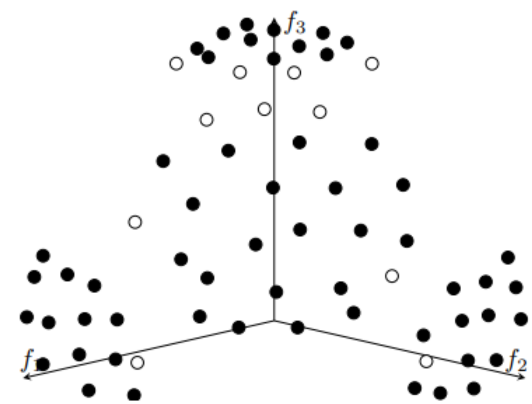


(a) RV set R_1 for Stage 2

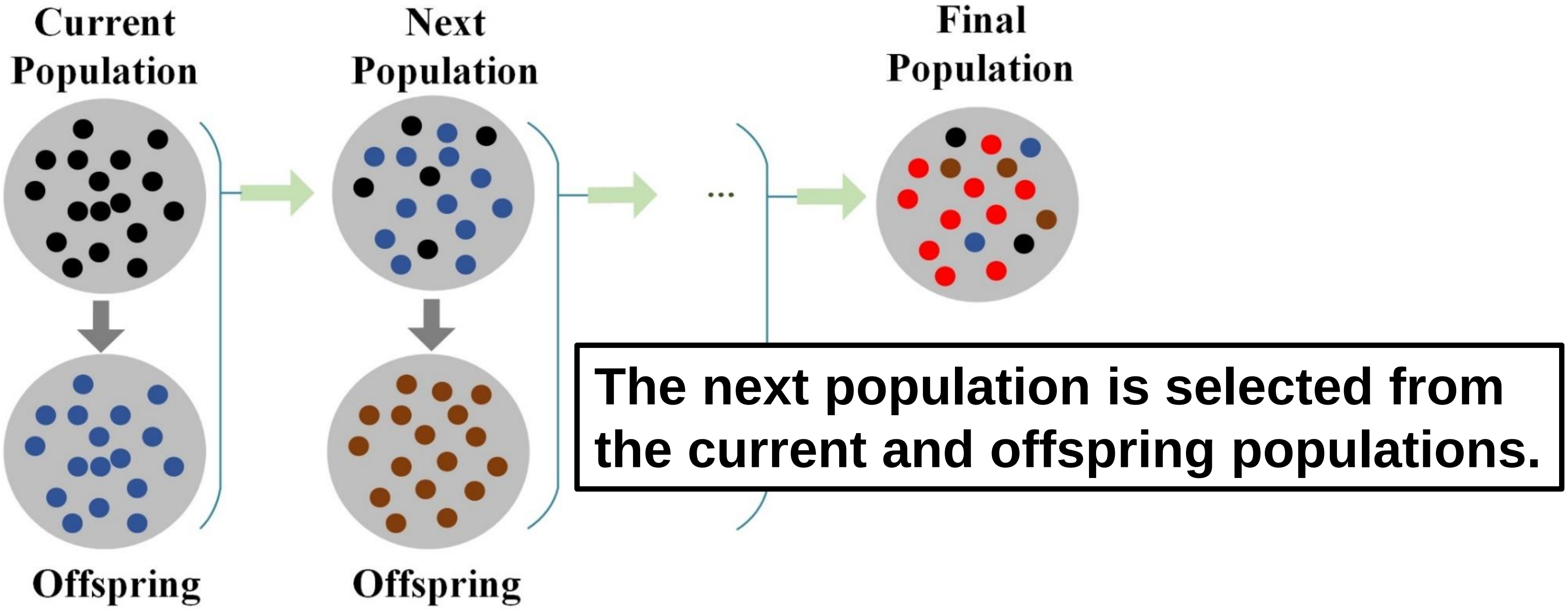


$+ R_{1\bar{c}}$: ARVs $\times R_{1\bar{c}}$: IRVs

(b) Stage 2 ND set

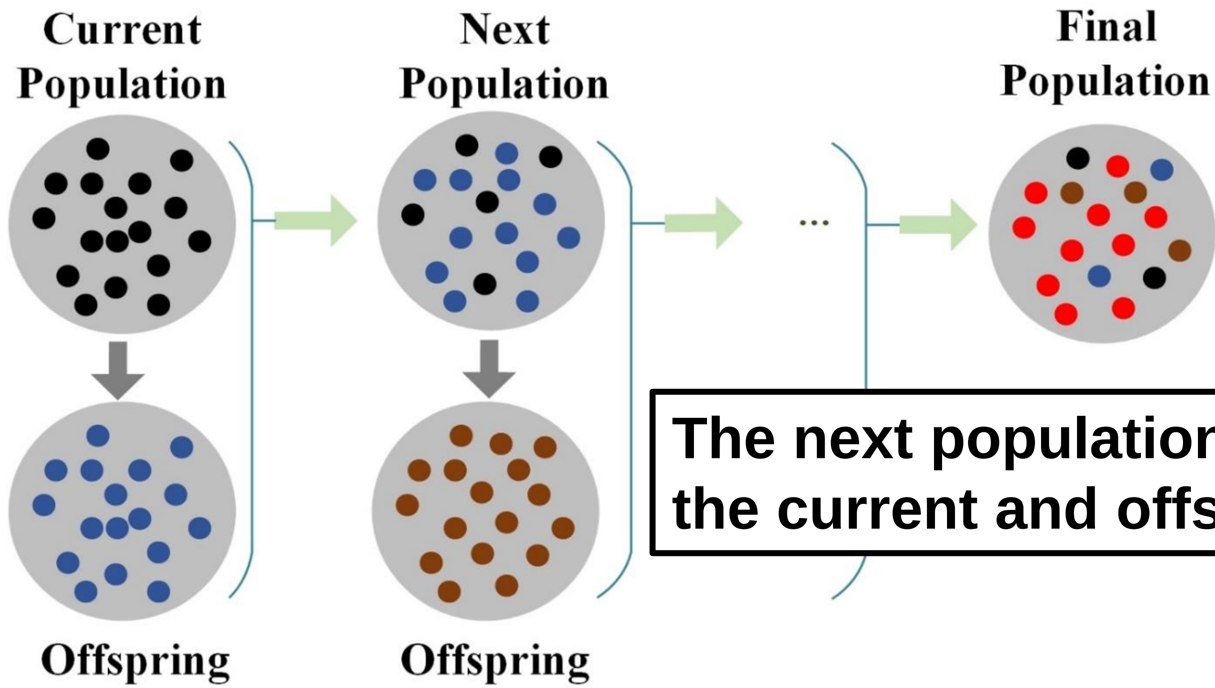


$\bullet\, S_{1c}^{T1}$ solution set $\circ\, S_2$ solution set



Difficulty in this framework:

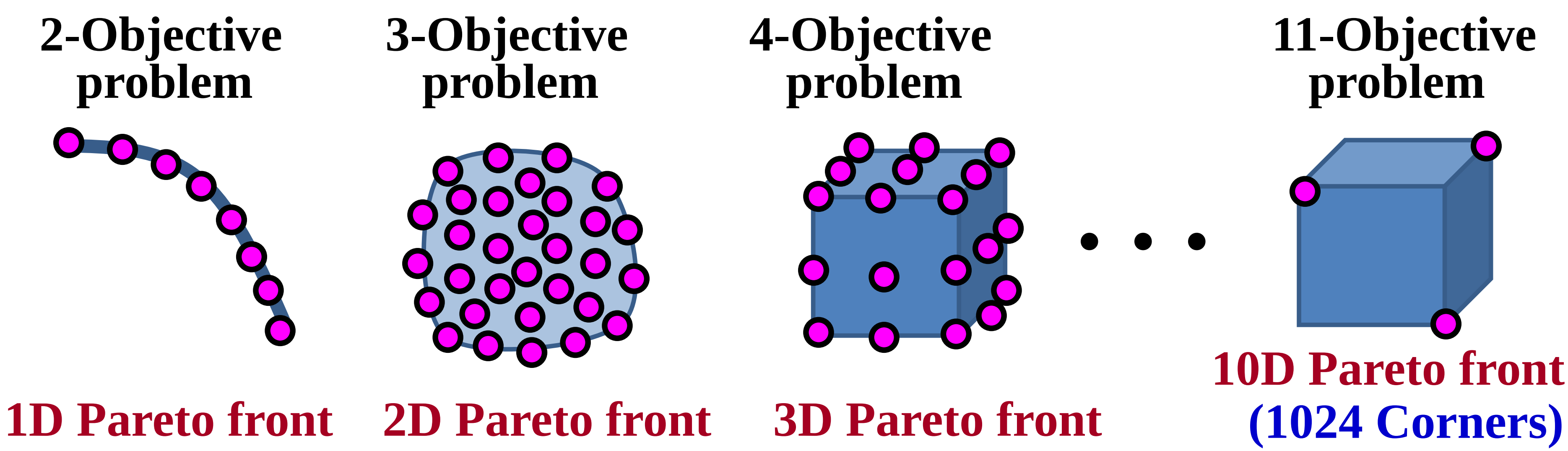
The final population size (i.e., population size) is limited (e.g., 200, 500)



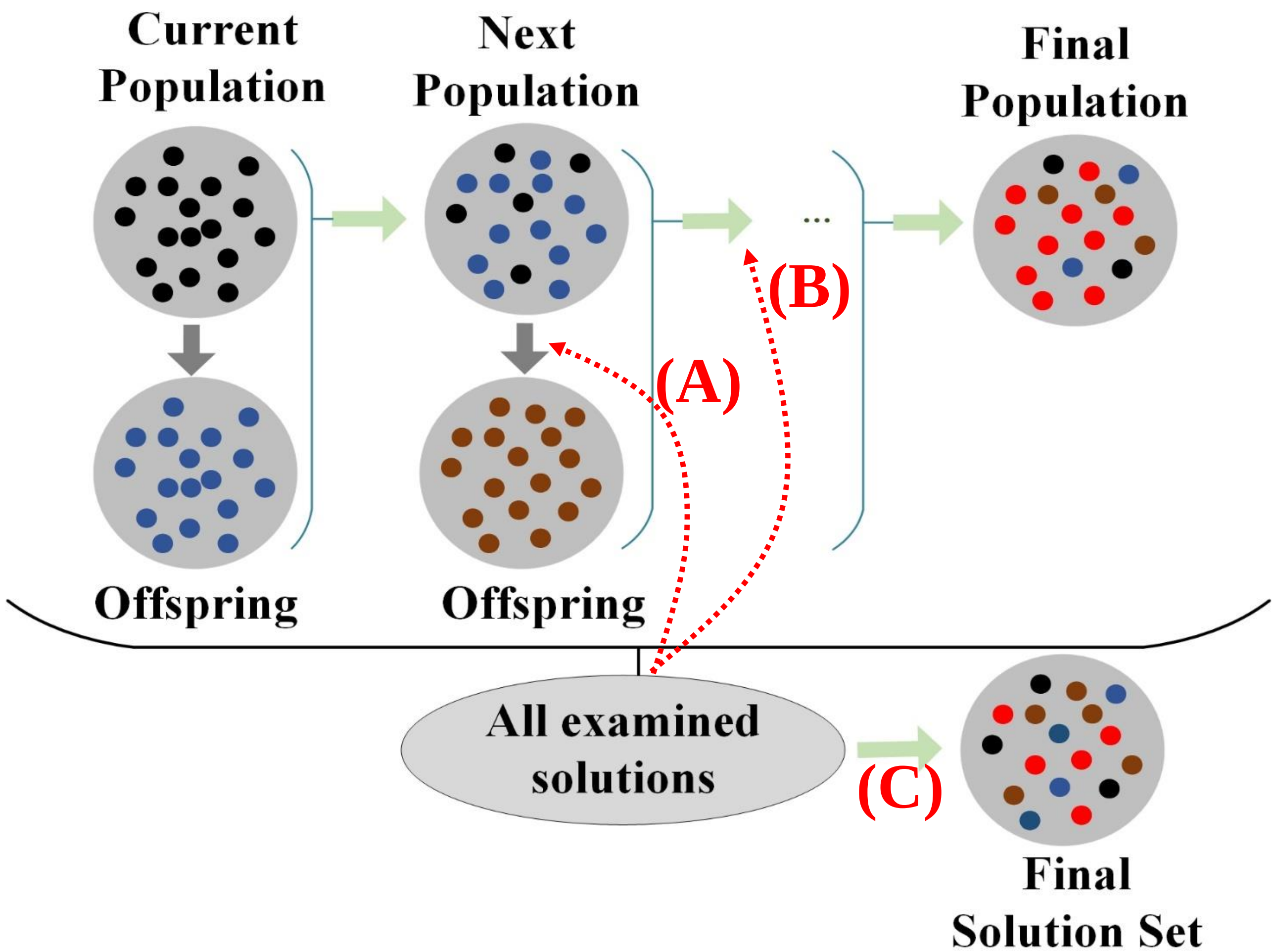
The next population is selected from the current and offspring populations.

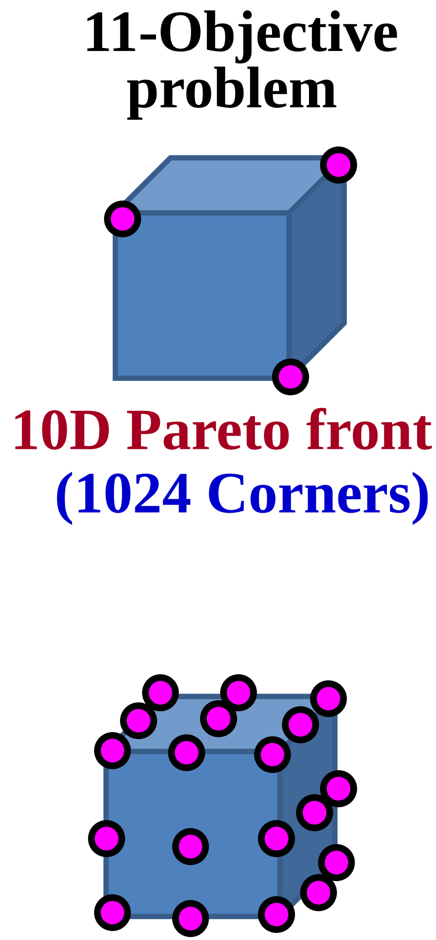
Difficulty in this framework:

The final population size (i.e., population size) is limited (e.g., 200, 500)



New EMO Framework (Not new: This was used in some studies)

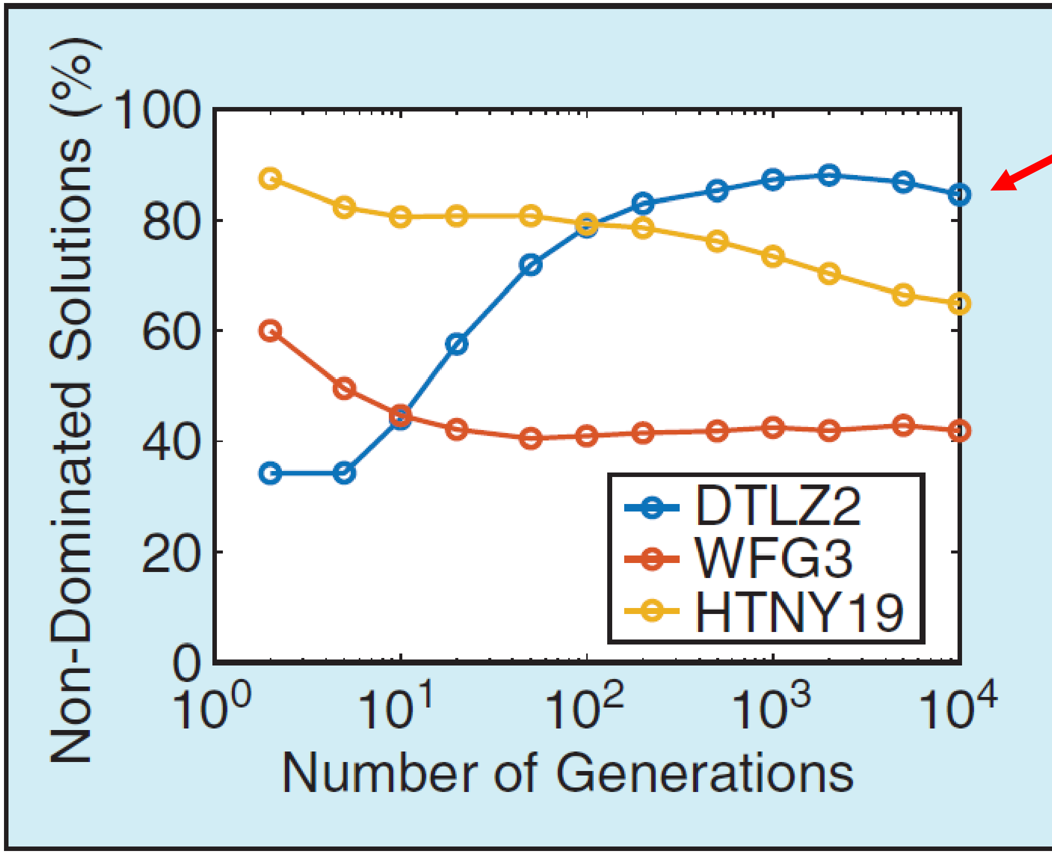




Interesting Research Topics:

(A) How to utilize the stored solutions to generate new solutions.
(B) How to utilize the stored solutions to choose the next population.
(C) How to choose the final solution set from the stored solutions.

Number of non-dominated solutions among examined solutions



Among 2,750,000 solutions,
84.6% are non-dominated.
(More than 2 million)

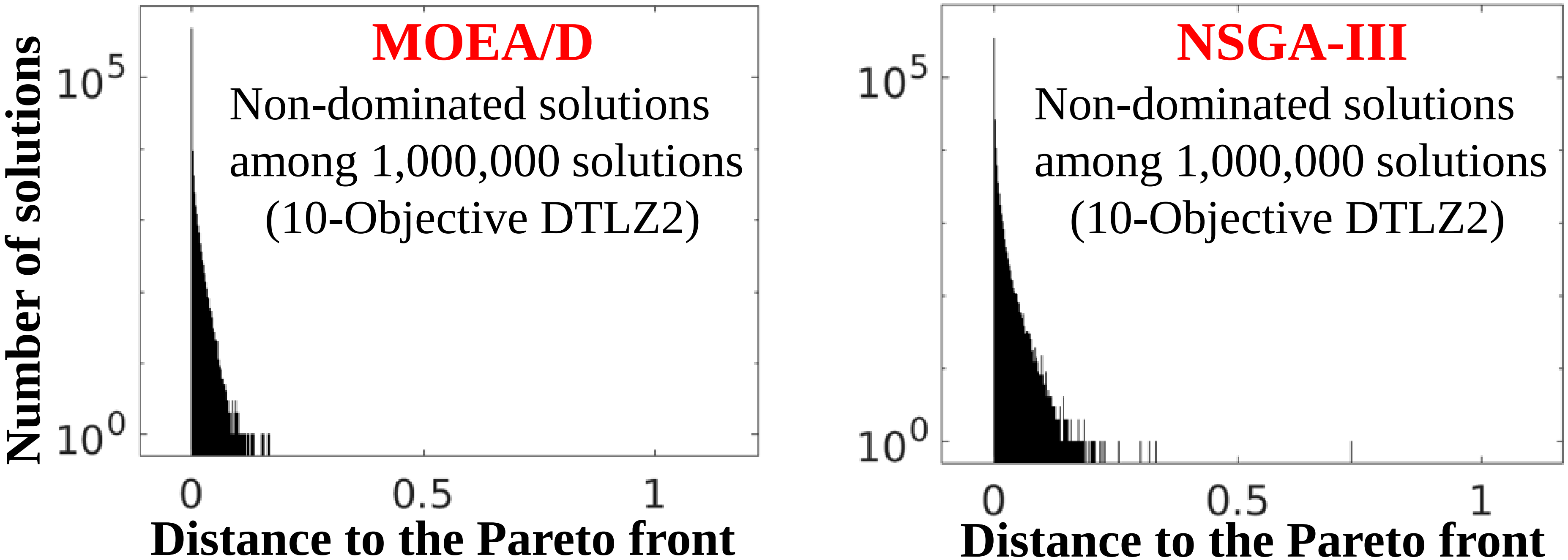
Percentage of non-dominated solutions among examined solutions by MOEA/D-Tch on each 10-objective test problem (Population size: 275).
Fig. 14 of Ishibuchi et al.: "Difficulties in Fair Performance Comparison of Multi-Objective Evolutionary Algorithms", IEEE CIM Feb 2022 Issue.

Q. Are you sure that all of them are good solutions?

Distribution of non-dominated solutions among the examined solutions

Distance from the Pareto front: Pareto front is in $[0, 1]^m$

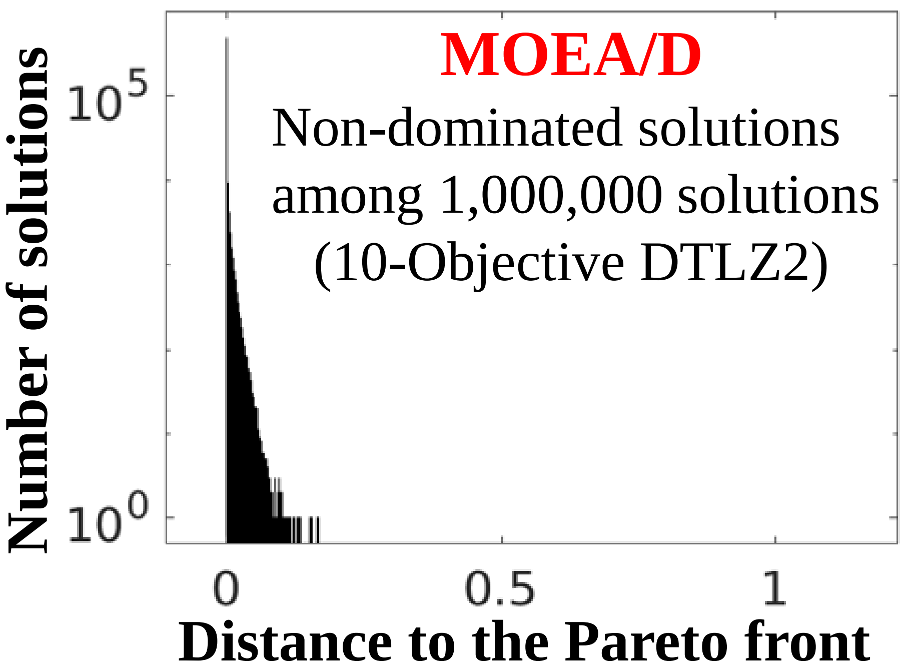
After 1,000,000 solutions are examined for the 10-objective DTLZ2

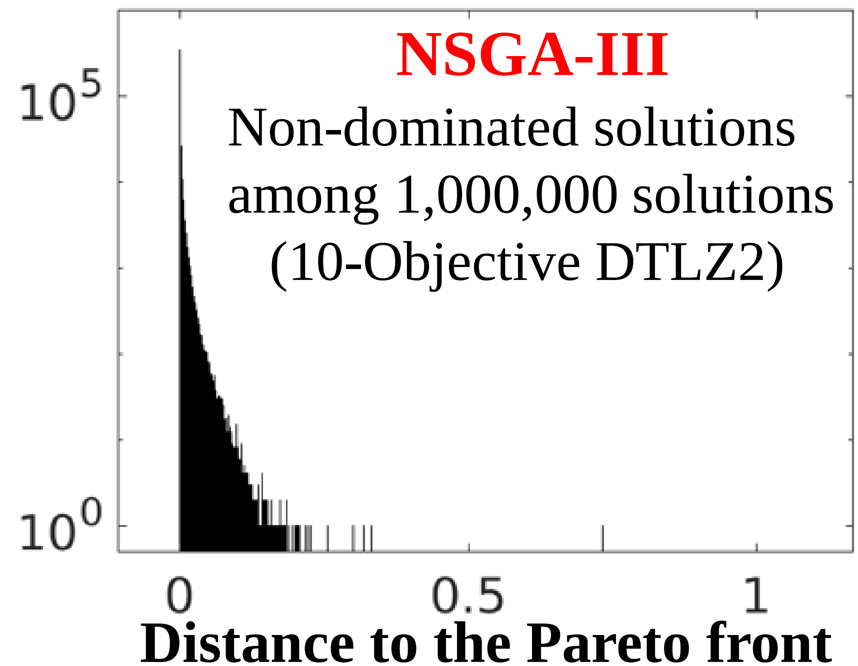


Distribution of non-dominated solutions among the examined solutions

Distance from the Pareto front: Pareto front is in $[0, 1]^m$

After 1,000,000 solutions are examined for the 10-objective DTLZ2

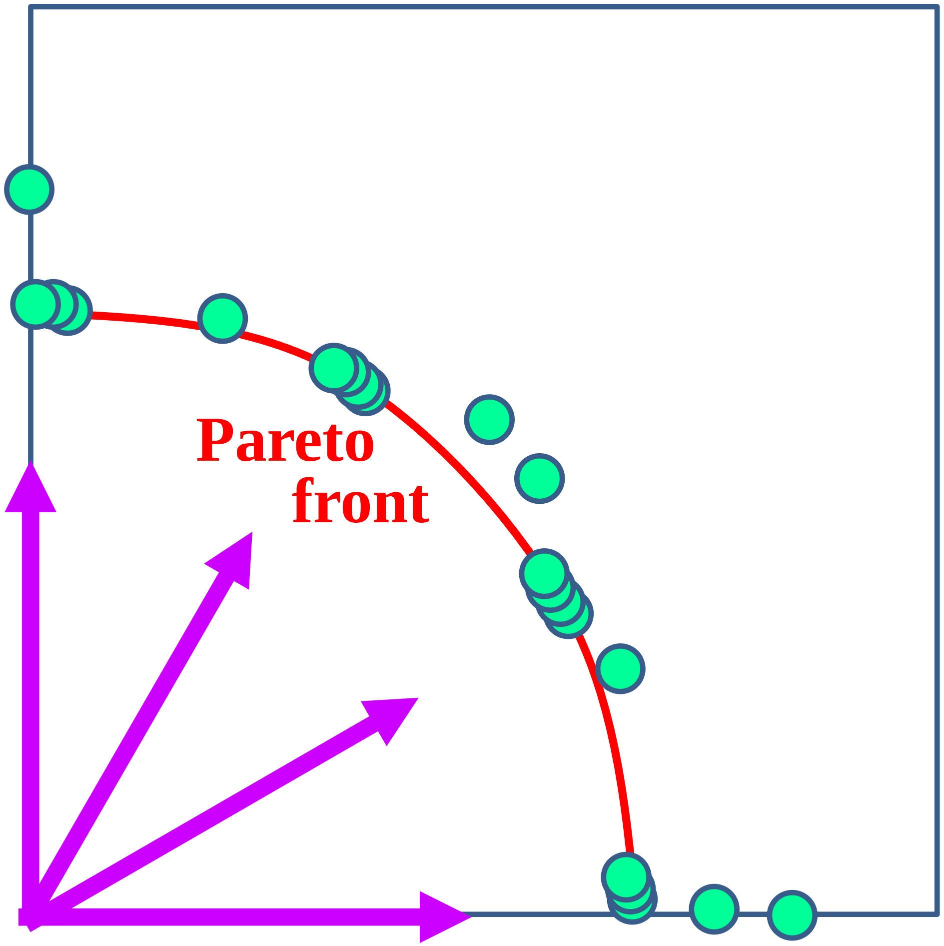


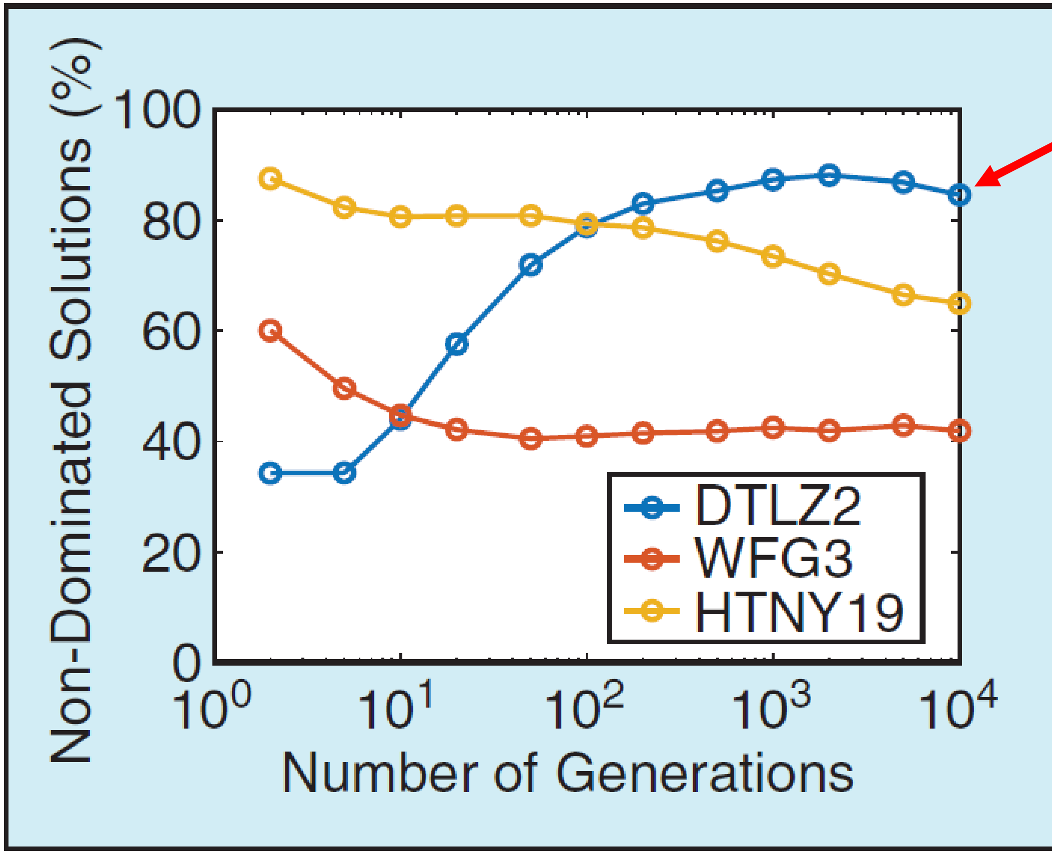


Observations:

- Many non-dominated solutions are very close to the Pareto front.
- Other non-dominated solutions are not very close to the Pareto front.

Possible distribution of non-dominated solutions in the 10-D Space



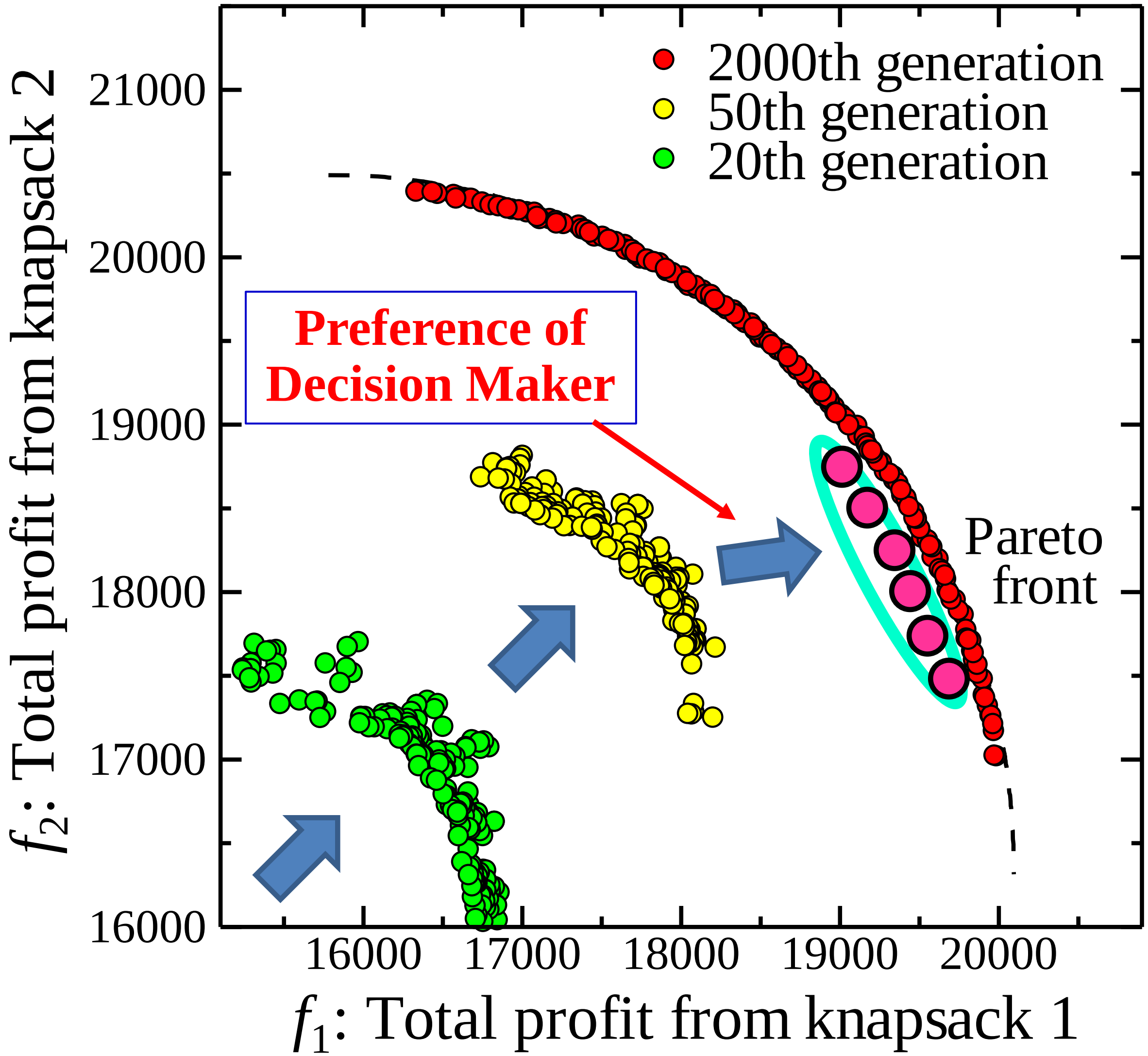


Among 2,750,000 solutions, 84.6% are non-dominated.

Percentage of non-dominated solutions among examined solutions by MOEA/D-Tch on each 10-objective test problem (Population size: 275). Fig. 14 of Ishibuchi et al.: “Difficulties in Fair Performance Comparison of Multi-Objective Evolutionary Algorithms”, IEEE CIM Feb 2022 Issue.

Current Hot Topics in EMO

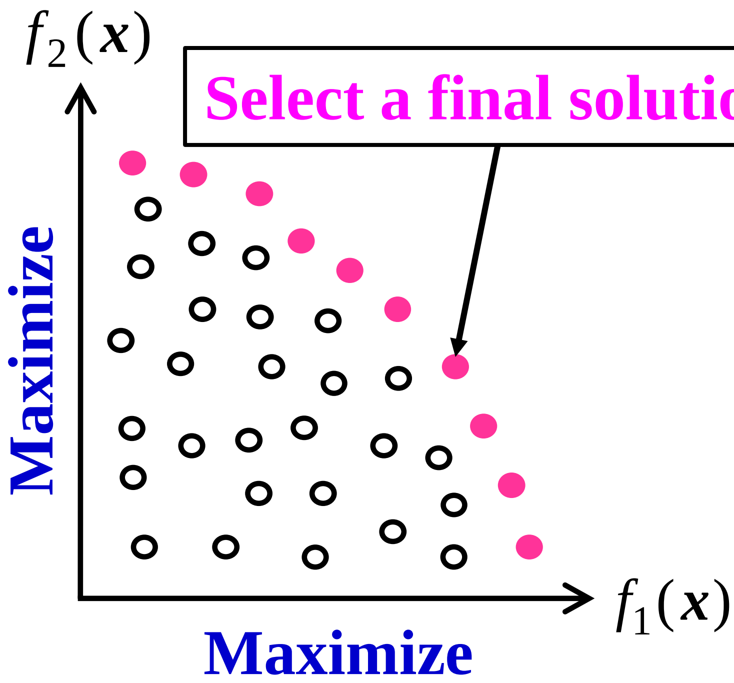
(8) Preference incorporation in EMO algorithms



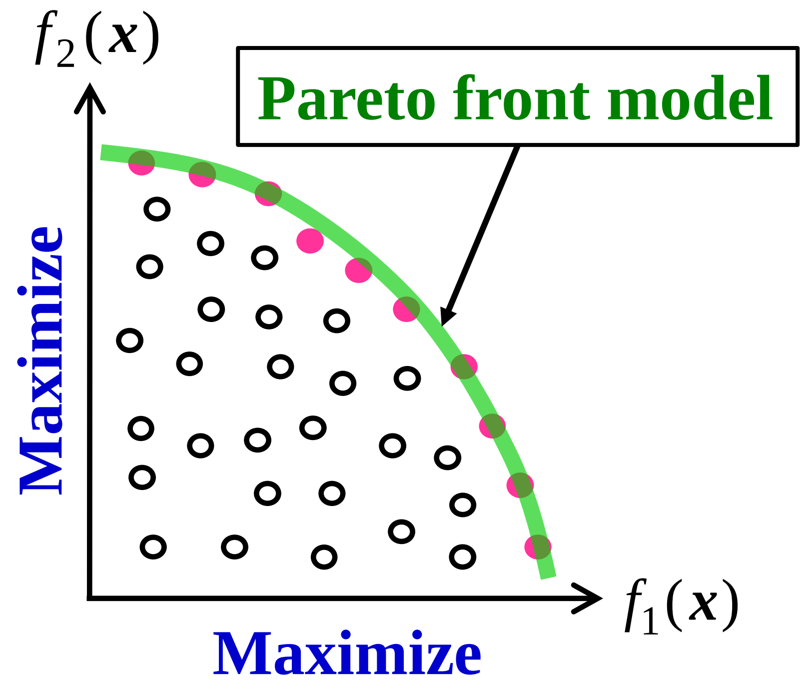
Current Hot Topics in EMO

(9) Pareto front modeling

Standard EMO Approach



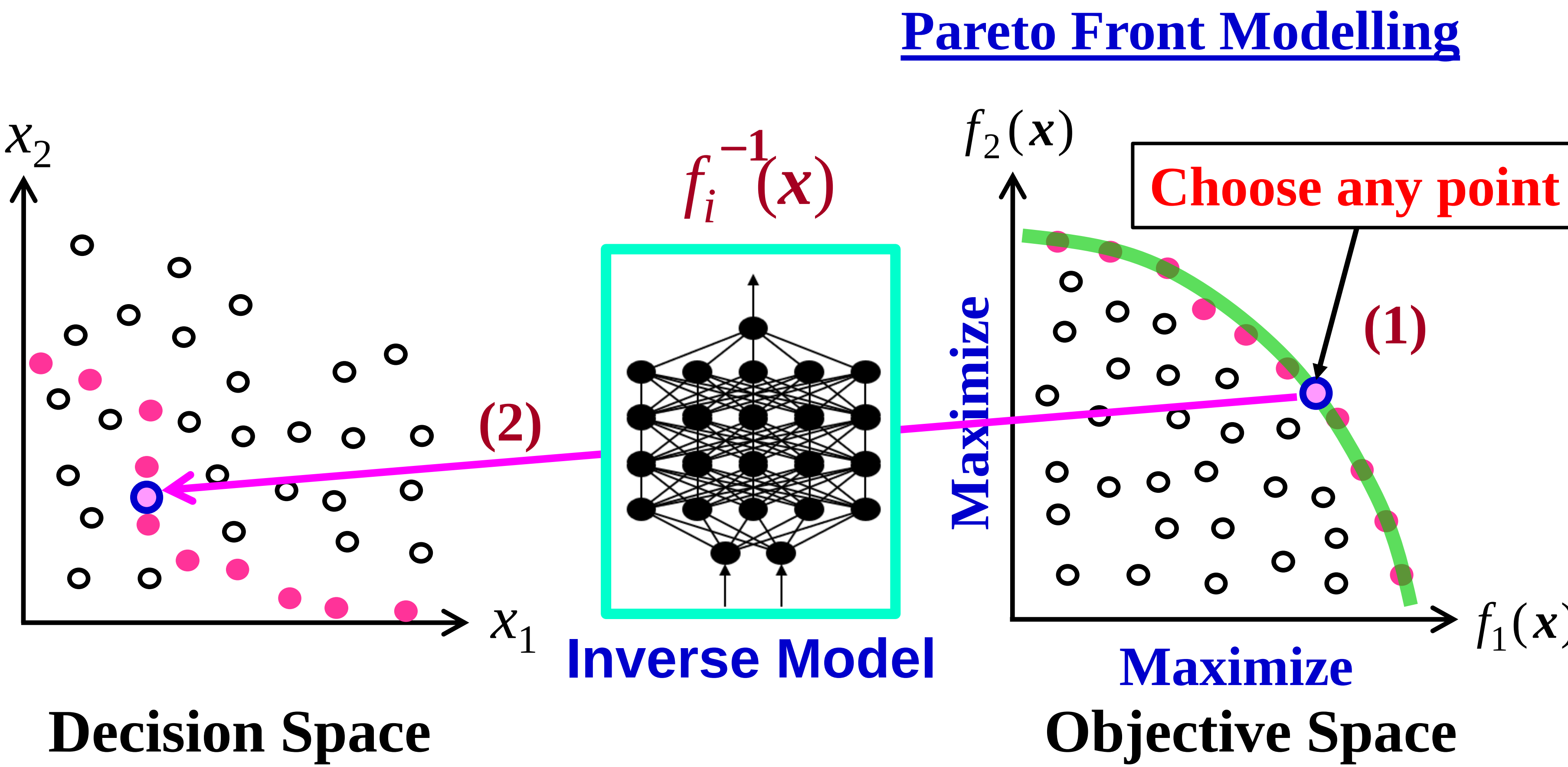
Pareto Front Modelling



[1] X Lin, Z Yang, Q Zhang, Pareto Set Learning for Neural Multi-objective Combinatorial Optimization (ICLR 2022)

[2] X Lin, Z Yang, X Zhang, Q Zhang, Pareto Set Learning for Expensive Multi-Objective Optimization (NeurIPS 2022)

New Idea of EMO-based Decision Making



Standard EMO Approach:

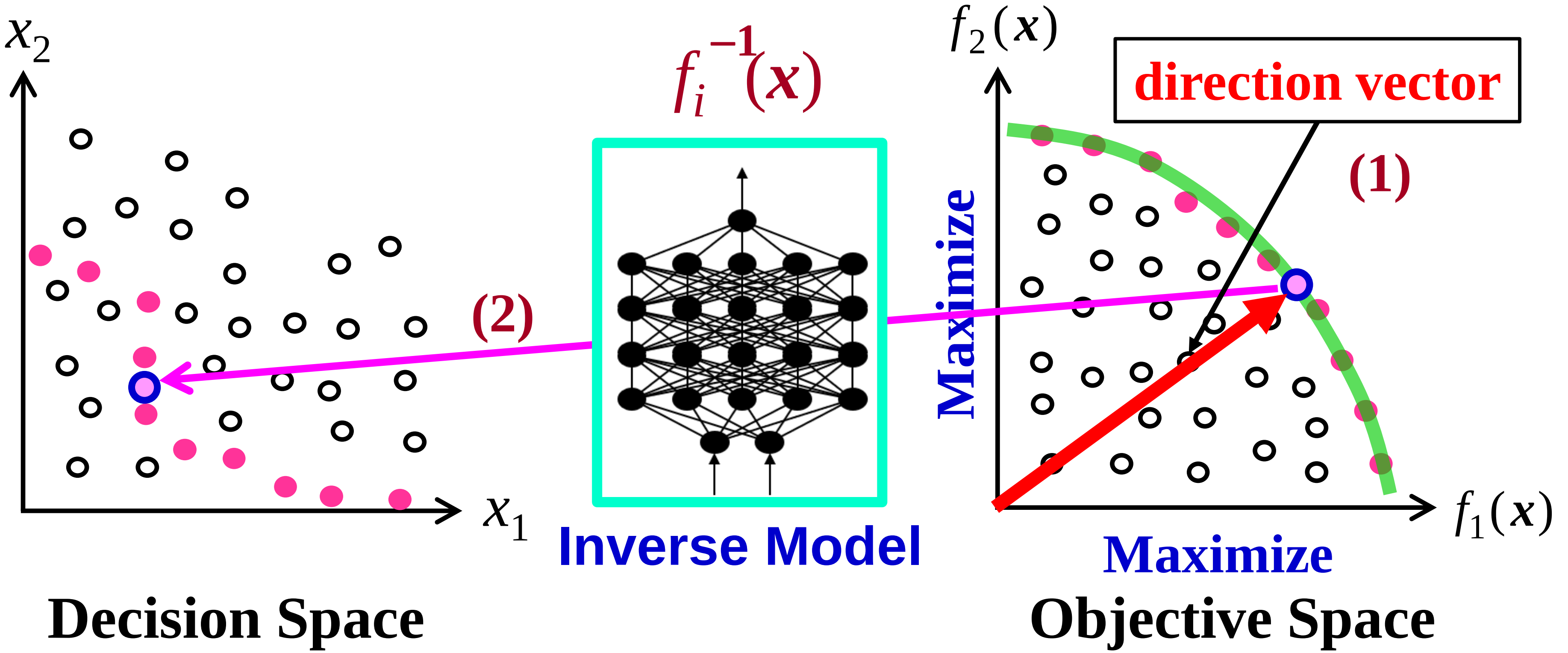
- One of the obtained solutions is selected by the decision maker.

Pareto front modeling:

- Any point on the Pareto front can be selected.

New Idea of EMO-based Decision Making

The decision maker can specify any direction vector. Then the corresponding solution can be obtained from the inverse model.



Standard EMO Approach:

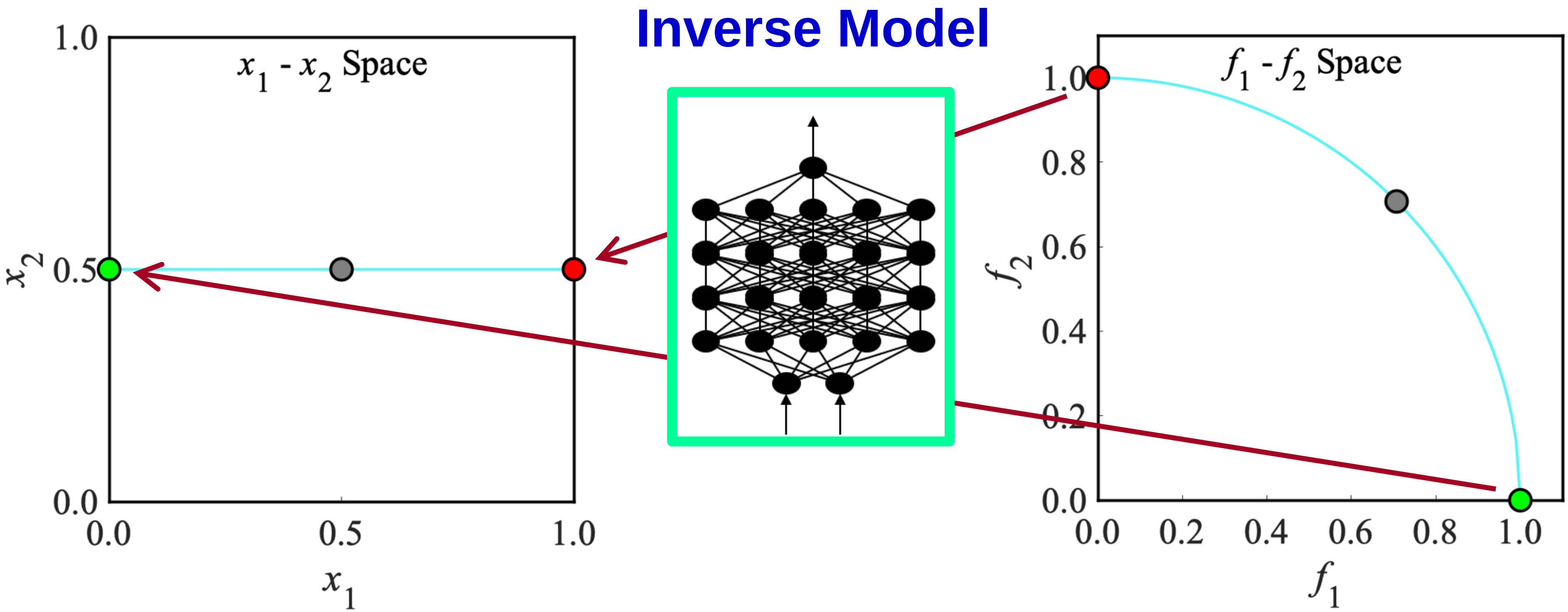
- One of the obtained solutions is selected by the decision maker.

Pareto front modeling:

- Any point on the Pareto front can be selected.

Surprisingly good results have been reported in the literature.

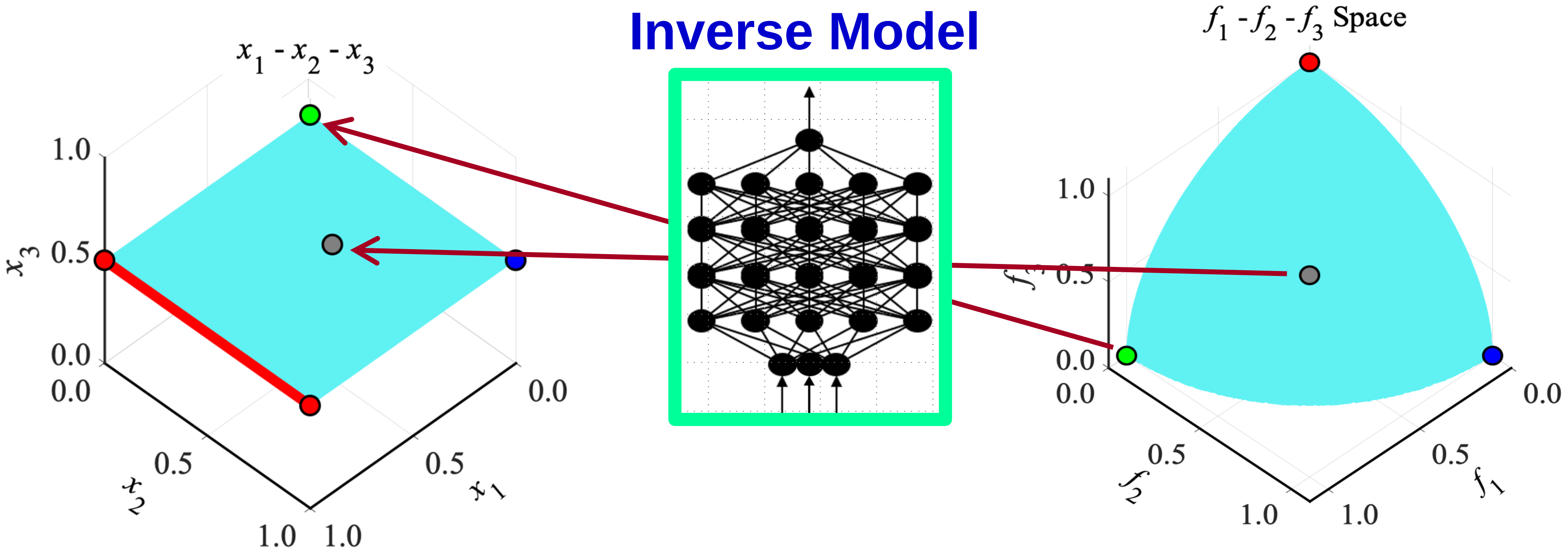
One possibility: Very good results are obtained based on special characteristic features of test problems (i.e., very simple relations between the decision space and the objective space).



Two-Objective DTLZ2 Problem

Surprisingly good results have been reported in the literature.

One possibility: Very good results are obtained based on special characteristic features of test problems (i.e., very simple relations between the decision space and the objective space).



Three-Objective DTLZ2 Problem

Current Hot Topics in EMO

(10) New EMO Algorithm Design for Real-World Problems

Some examples of Pareto Fronts from [2].

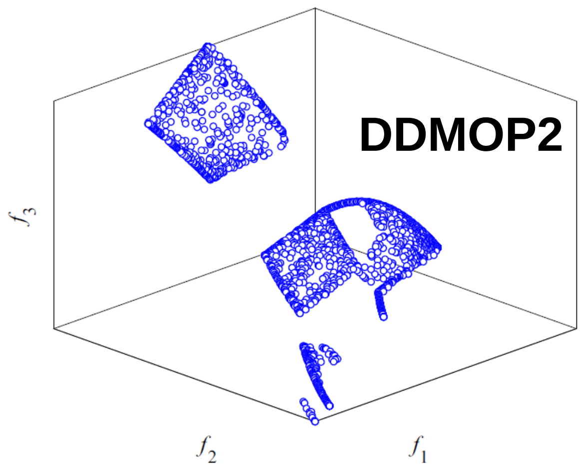


Fig. 3 The approximate POF of DDMOP2

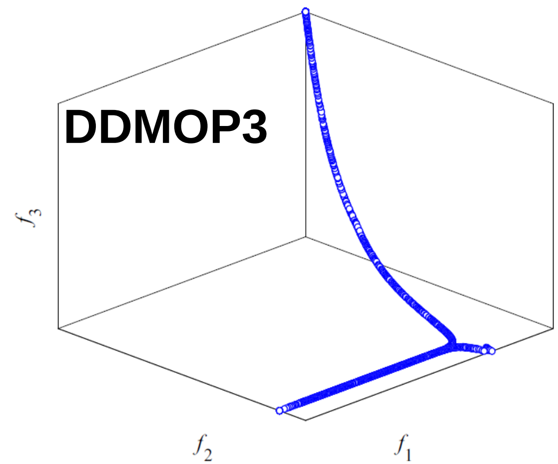


Fig. 4 The approximate POF of DDMOP3

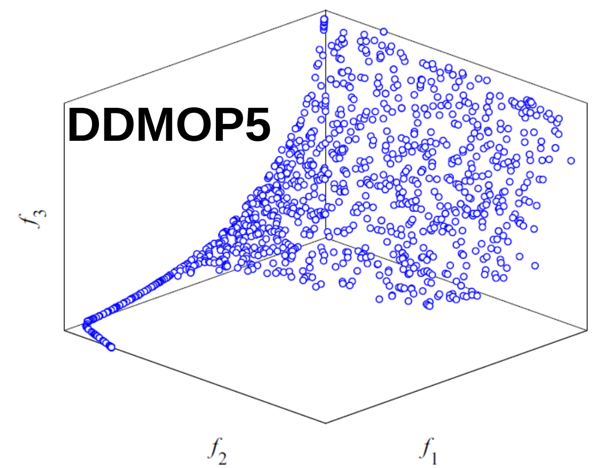


Fig. 5 The approximate POF of DDMOP5

[1] **16 RE problems**: Tanabe, R., Ishibuchi, H.: **An easy-to-use real-world multi-objective optimization problem suite.** *Applied Soft Computing*, 106078 (2020).

[2] **7 DDMOP problems**: He, C., Tian, Y., Wang, H., Jin, Y.: **A repository of real-world datasets for data-driven evolutionary multiobjective optimization.** *Complex & Intelligent Systems* 6, 189-197 (2020).

[3] **50 RCM problems**: Kumar, A., Wu, G., Ali, M. Z., Luo, Q., Mallipeddi, R., Suganthan, P. N., Das, S.: **A benchmark-suite of real-world constrained multi-objective optimization problems and some baseline results.** *Swarm and Evolutionary Computation* 67, 100961 (2021).

Current Hot Topics in EMO

(10) New EMO Algorithm Design for Real-World Problems

Some examples of Pareto Fronts from [2].

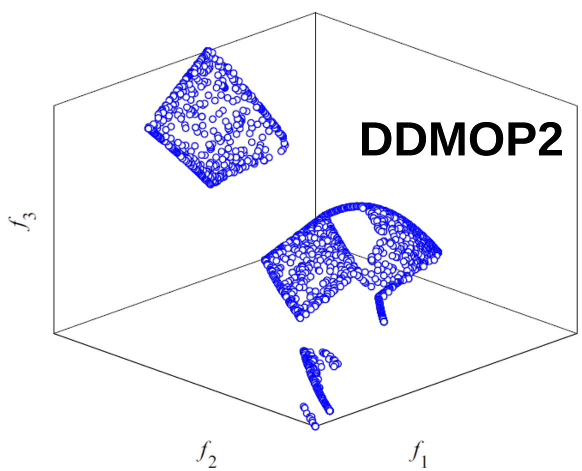


Fig. 3 The approximate POF of DDMOP2

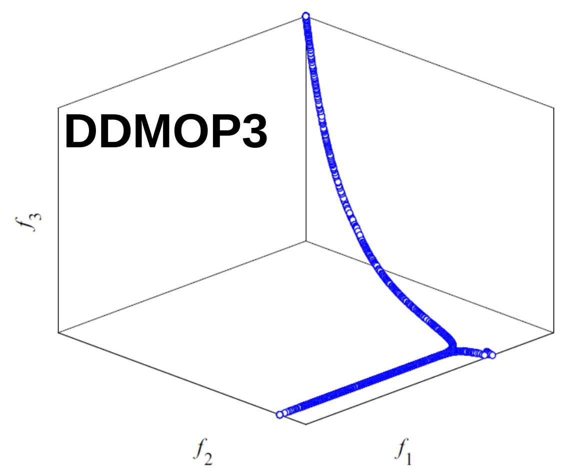
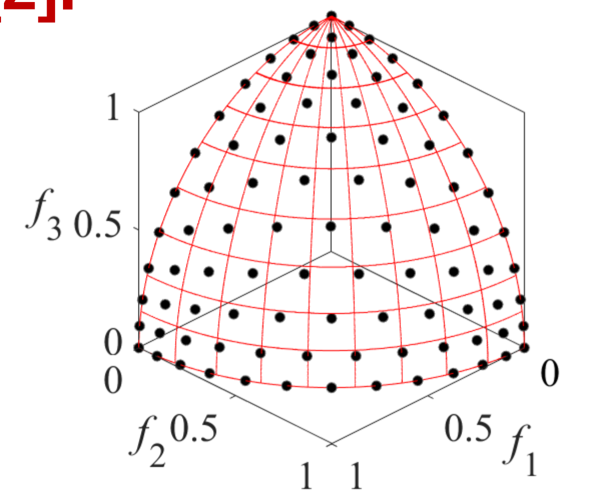


Fig. 4 The approximate POF of DDMOP3



MOEA/D on DTLZ2

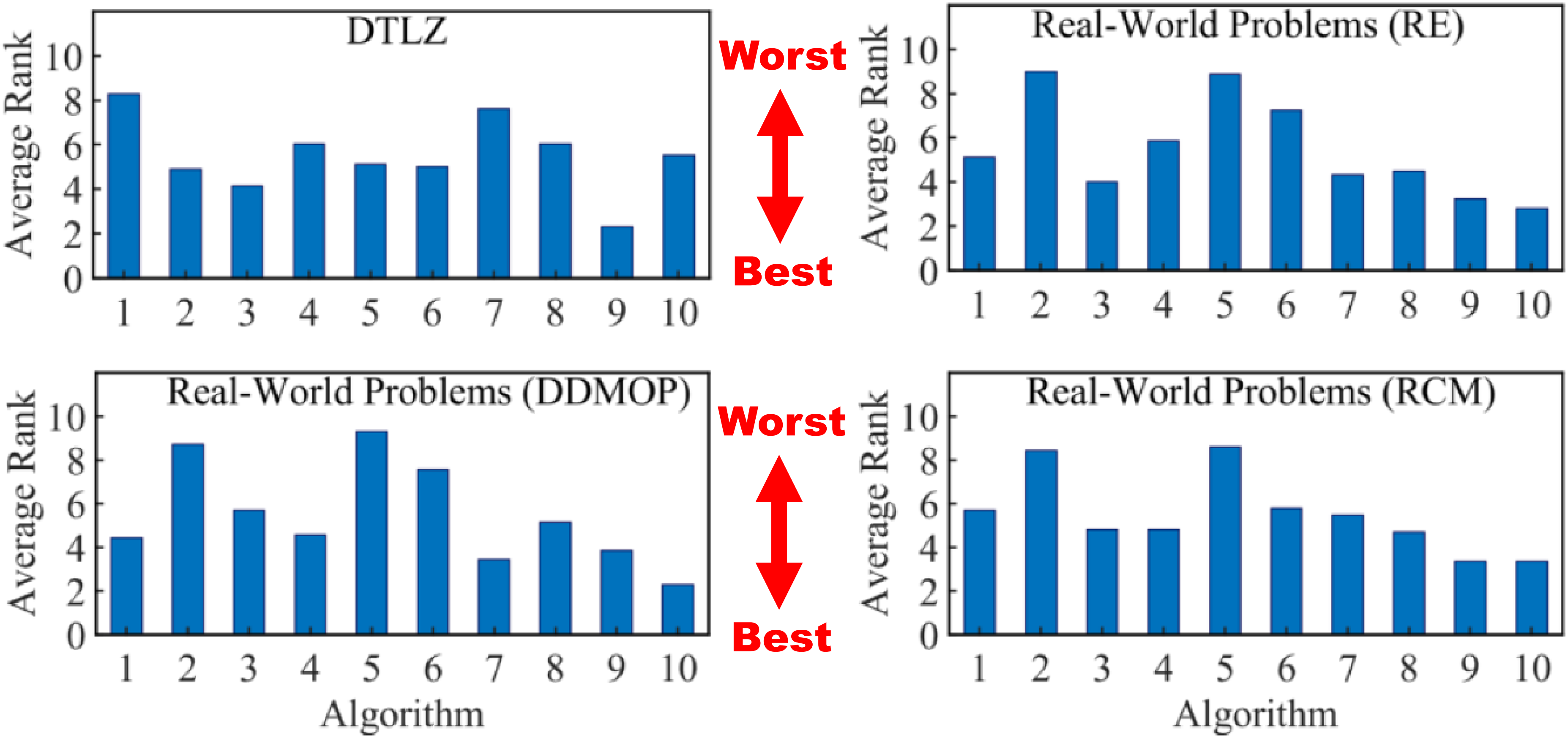
[1] **16 RE problems**: Tanabe, R., Ishibuchi, H.: **An easy-to-use real-world multi-objective optimization problem suite.** *Applied Soft Computing*, 106078 (2020).

[2] **7 DDMOP problems**: He, C., Tian, Y., Wang, H., Jin, Y.: **A repository of real-world datasets for data-driven evolutionary multiobjective optimization.** *Complex & Intelligent Systems* 6, 189-197 (2020).

[3] **50 RCM problems**: Kumar, A., Wu, G., Ali, M. Z., Luo, Q., Mallipeddi, R., Suganthan, P. N., Das, S.: **A benchmark-suite of real-world constrained multi-objective optimization problems and some baseline results.** *Swarm and Evolutionary Computation* 67, 100961 (2021).

Comparison Results on DTLZ and Real-world Problems

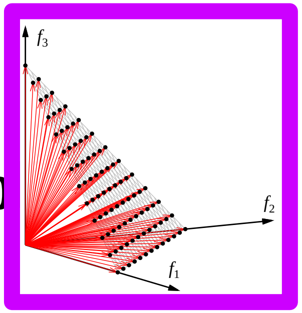
1: NSGA-II	3: SMS-EMOA & HypE	5: MOEA/DD	7: SparseEA	9: R2HCA-EMOA
2: MOEA/D-PBI	4: NSGA-III	6: RVEA	8: DEA-GNG	10: PREA



Totally different comparison results of ten EMO algorithms between the test problem DTLZ and the real-world problems

H. Ishibuchi, Y. Nan, and L. M. Pang, "Performance evaluation of multi-objective evolutionary algorithms using artificial and real-world problems," *Proc. EMO 2023*.

Current Hot Topics in EMO



(10) New EMO Algorithm Design for Real-World Pro

Some examples of Pareto Fronts from [2].

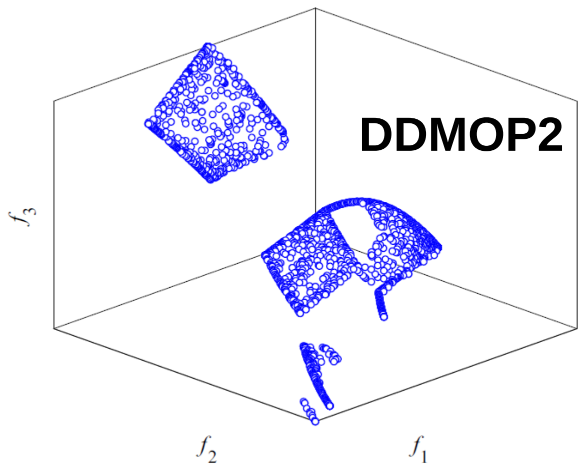


Fig. 3 The approximate POF of DDMOP2

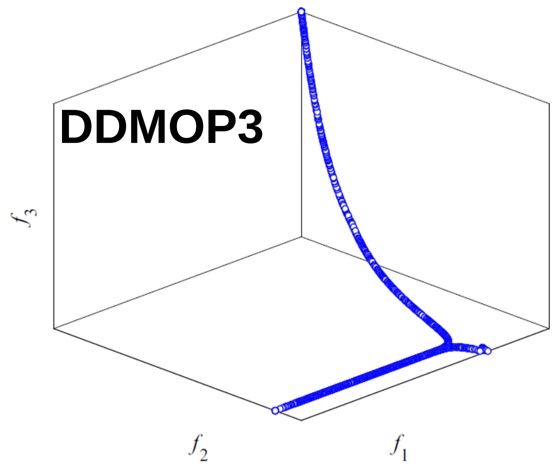
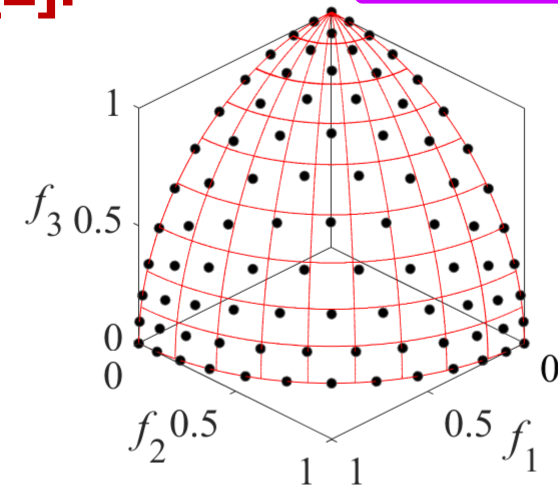


Fig. 4 The approximate POF of DDMOP3



MOEA/D on DTLZ2

[1] **16 RE problems**: Tanabe, R., Ishibuchi, H.: **An easy-to-use real-world multi-objective optimization problem suite.** *Applied Soft Computing*, 106078 (2020).

[2] **7 DDMOP problems**: He, C., Tian, Y., Wang, H., Jin, Y.: **A repository of real-world datasets for data-driven evolutionary multiobjective optimization.** *Complex & Intelligent Systems* 6, 189-197 (2020).

[3] **50 RCM problems**: Kumar, A., Wu, G., Ali, M. Z., Luo, Q., Mallipeddi, R., Suganthan, P. N., Das, S.: **A benchmark-suite of real-world constrained multi-objective optimization problems and some baseline results.** *Swarm and Evolutionary Computation* 67, 100961 (2021).

Thirty-one logicians came from different countries to participate in the Annual International Conference on Logic. After greeting all 31 participants, the main organizer remarked that it would be necessary to run a special test to check whether all participants were indeed logicians as they claimed to be. He explained kindly that in the past there had been cases where some non-logicians tried to get into the conference, and he would not allow that to happen again. He further explained the basis of the test: He said that each participant would get a dot of some color that he would place on each participant's forehead. Each participant would be allowed to look around (thus everyone would see the dots of all other participants except his own), but no communication of any sort would be allowed. After a while, the organizer would ring a bell and if any participant had deduced the color of his or her dot, they should leave the room. The organizer would ring a bell as many times as necessary. As the organizer knows the colors of all of the dots, he also knows when each participant should leave the room (if the participant is a logician). This was the essence of the test. At this stage, the organizer asked the participants whether there were any questions. One participant raised his hand and asked whether it was possible to pass the test, i.e., to correctly guess the color of his dot. The organizer replied that he had selected the colors of all the dots in such a way that every participant should be able to deduce the color of his/her dot. As this was the only question from the crowd, the test started. The organizer placed the color dots on the foreheads of all the participants and waited for a while, so that everyone had a chance to look around. After a few minutes, he rang the bell for the first time. At this moment, four participants left the room. When he rang the bell for the second time, all the participants with red dots left the room. When he rang the bell for the third time, no one moved. When he rang the bell for the fourth time, at least one participant left. Soon afterwards, when he rang the bell again, the participant who asked the only question before the commencement of the test left together with his sister and some other participants. He and his sister had dots of different color. At this stage there were still some participants left in the room.

Assuming that all the participants were true logicians (so everyone was leaving the room at the right time), how many times did the organizer ring the bell? (Hint: If one participant has a different color from all the others, he/she cannot deduce his/her color. Thus, he/she can assume that at least one participant has the same color as his/her color since all of them were true logicians.)

Lab Session Task 5 (Optional):

Three-Objective Problem: Minimize $f_1(\boldsymbol{x})$, $f_2(\boldsymbol{x})$, and $f_3(\boldsymbol{x})$

Let us consider the following two Pareto fronts in $[0, 1]^3$:

(i) $f_1(\boldsymbol{x}) + f_2(\boldsymbol{x}) + f_3(\boldsymbol{x}) = 1$. (ii) $f_1(\boldsymbol{x}) + f_2(\boldsymbol{x}) + f_3(\boldsymbol{x}) = 2$.

The following distribution is to maximize the hypervolume value for each Pareto front when the reference point is (1000, 1000):

(i) Solutions are well distributed over the entire Pareto front.

(ii) All (almost all) solutions are on the boundary of the Pareto front.

Please explain why these totally different solution sets maximize the hypervolume value for the two Pareto fronts.