GECCO 2019 Tutorial

Evolutionary Many-Objective Optimization

Hisao Ishibuchi
Southern University of Science Technology (SUSTech)

Hiroyuki Sato
The University of Electro-Communications

GECCO '19, *July 13 - 17, 2019, Prague, Czech Republic*

ACM ISBN 978-1-4503-6748-6/19/07… $15.00
https://doi.org/10.1145/3319619.3323377

Plan of This Tutorial

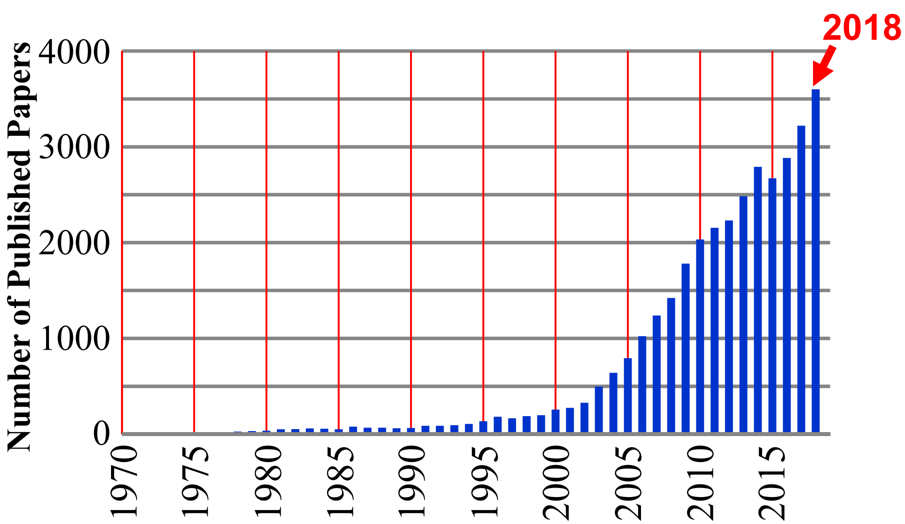
We explain why many-objective optimization is difficult for EMO (Evolutionary Multiobjective Optimization) algorithms, and how those difficulties can be handled. We also suggest some promising future research directions.

Part 1 (Hisao Ishibuchi): Difficulties

Part 2 (Hiroyuki Sato): Approaches

Part 3 (Hisao Ishibuchi): Future Directions

Number of Papers with "Multi-objective" or "Multiobjective" in the Paper Titles (Source: Scopus Database)



Many-objective Optimization

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

Multi-Objective Optimization:

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

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

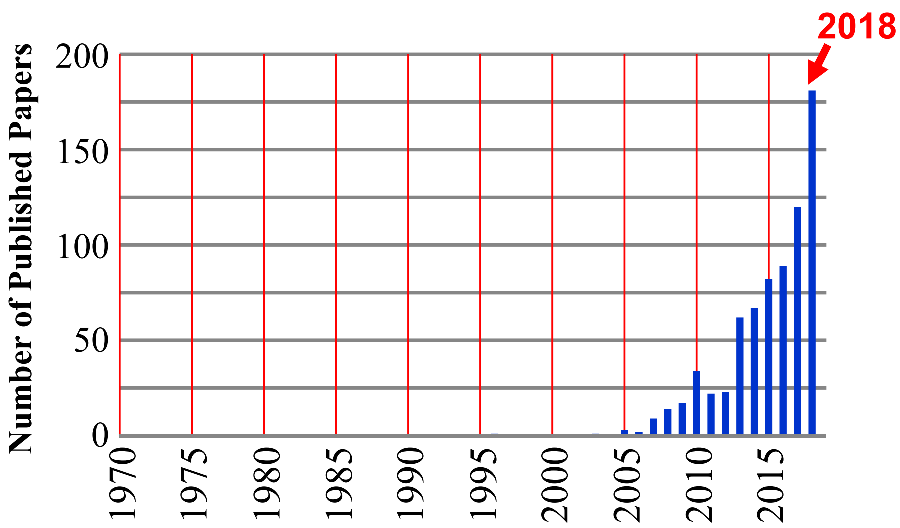
Many-Objective Optimization:

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

Number of Papers with "Many-Objective" in the Paper Titles (Source: Scopus Database)



Today's Plan (Part 1)

Difficulties in Evolutionary Many-Objective Optimization Studies

1. Difficulties related to many-objective search
2. Difficulties related to test problems
3. Difficulties related to performance evaluation

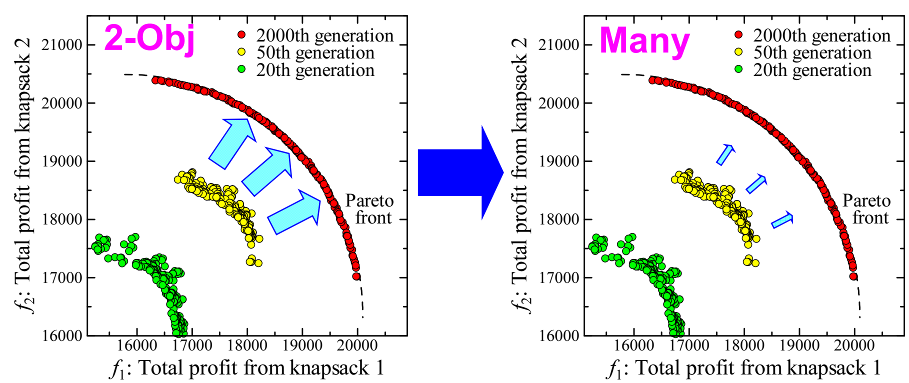
Many-Objective Optimization
Frequently Discussed Difficulties

1. **Search for Pareto Optimal Solutions**
   Pareto dominance does not work well.
2. **Approximation of the Entire Pareto Front**
   A huge number of solutions are needed.
3. **Presentation of Obtained Solutions to DM**
   Visualization of high-dimensional solutions is difficult.
4. **Selection of a Single Final Solution**
   Choice of a single final solution is difficult for DM.
5. **Examination of Search Behavior**
   Visual observation of many-objective search is difficult.
6. **Large Diversity of Solutions in a Population**
   Usefulness of crossover may be degraded.

1. Search for Pareto Optimal Solutions
Pareto dominance does not work well

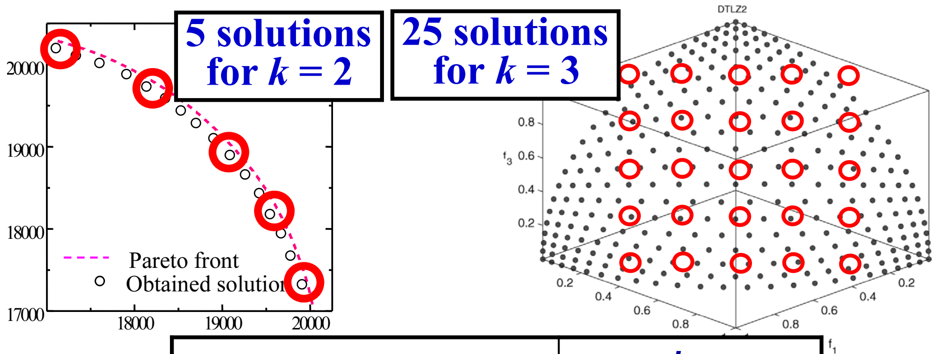
Q. Why are many-objective problems difficult for EMO ?

A. Solutions with many objectives are usually non-dominated with each other. Thus no strong selection pressure towards the Pareto front can be generated by Pareto dominance.



2. Approximation of the Entire Pareto Front

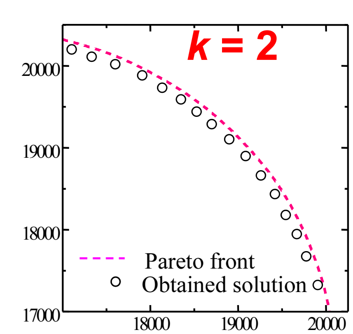
A huge number of solutions are needed

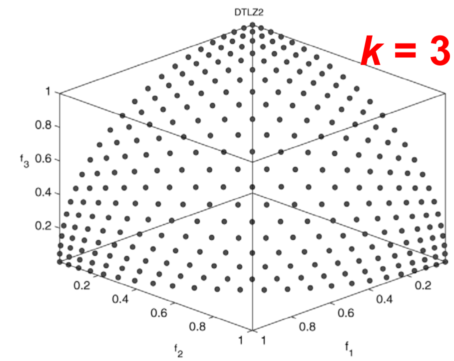


k-Objective Problem	$5^{(k-1)}$
2-Objective Problem	5
3-Objective Problem	25
10-Objective Problem	10 million

3. Presentation of Obtained Solutions to DM

Visualization of high-dimensional solutions is difficult





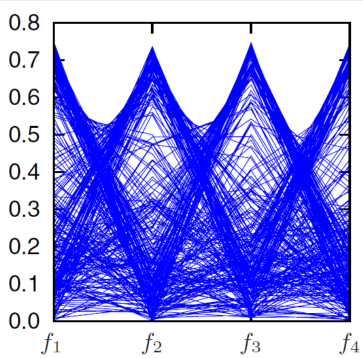
$k = 4$

How can we show a number of four-dimensional vectors to the decision maker?

3. Presentation of Obtained Solutions to DM

Visualization of high-dimensional solutions is difficult

$k = 4$

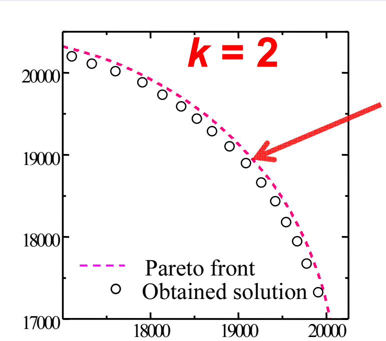


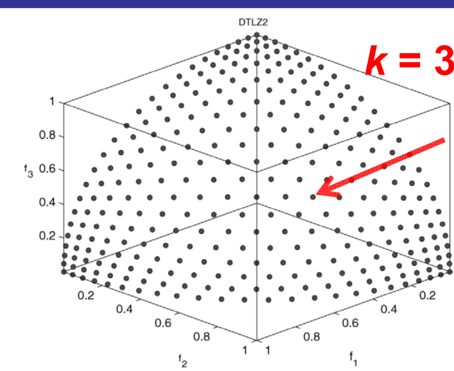
Obtained Solutions for a Four-Objective Problem

We can see that a wide variety of solutions are obtained. But, it is difficult to examine each solution.

4. Selection of a Single Final Solution

Choice of a single final solution is difficult for DM



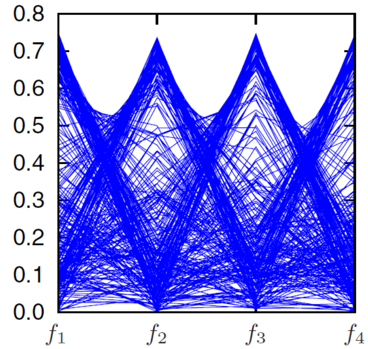


$k = 4$

How can we choose a single final solution from a large number of four-dimensional vectors?

4. Selection of a Single Final Solution
Choice of a single final solution is difficult for DM

k = 4

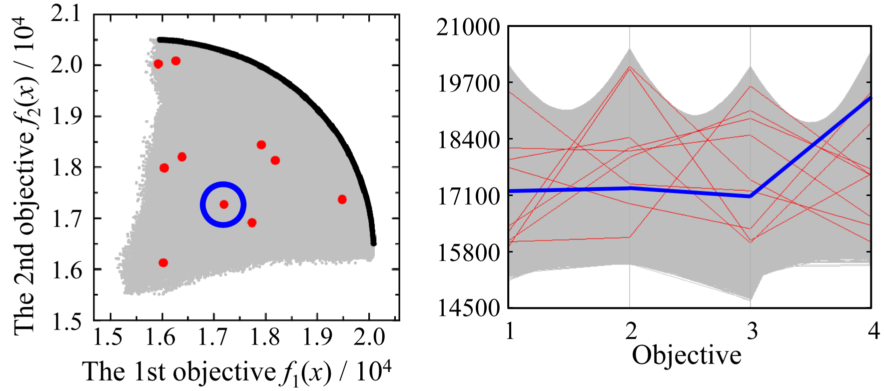


Obtained Solutions for a Four-Objective Problem

It may be very difficult for the decision maker to choose a single final solution from a large number of obtained non-dominate solutions.

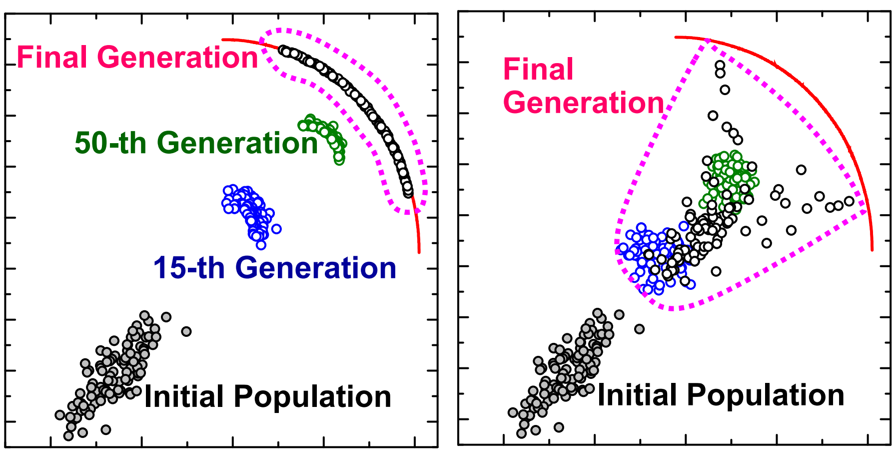
4. Selection of a Single Final Solution
Choice of a single final solution is difficult for DM

Presentation of only a small number of solutions may help the decision maker. How to select those solutions?



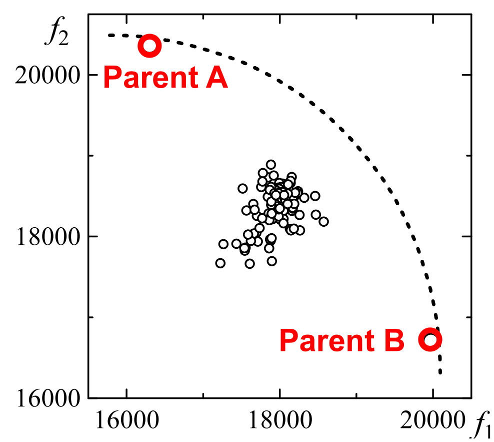
Ten solutions selected from 220,298 non-dominated solutions.

5. Examination of Search Behavior
Visual observation of many-objective search is difficult



2-Objective Problem **8-Objective Problem**

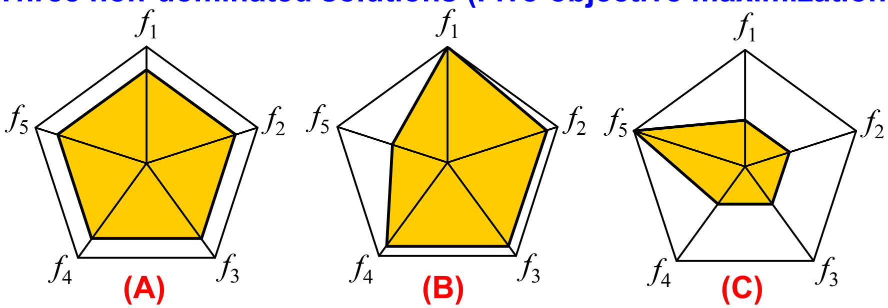
6. Large Diversity of Solutions in a Population
Usefulness of crossover may be degraded



Generated 100 offspring from two dissimilar parents (○) by uniform crossover for a 2-objective 500-item knapsack problem. Ishibuchi et al. IEEE Trans on EC (2015)

Difficulties of Many-Objective Problems

Three non-dominated solutions (Five-objective maximization)



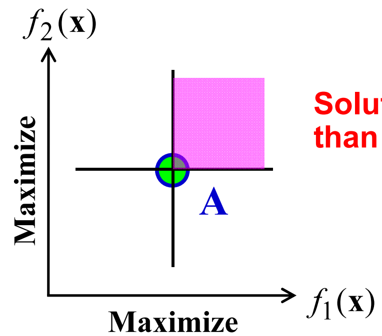
Good for all objectives. Very good except for f_5. Only f_5 is good.

These three solutions are non-dominated.

By increasing the number of objectives, almost all solutions become non-dominated.

Better Solution: Two-Objective

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

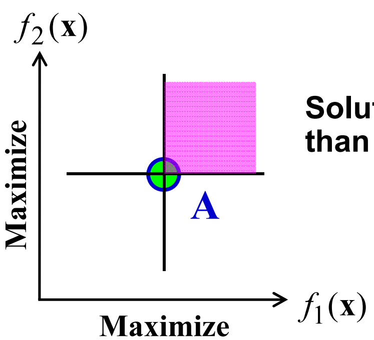


Solutions in this region are better than solution A. (1/4 of the space)

Pareto dominance-based comparison

Better Solution: Four-Objective

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

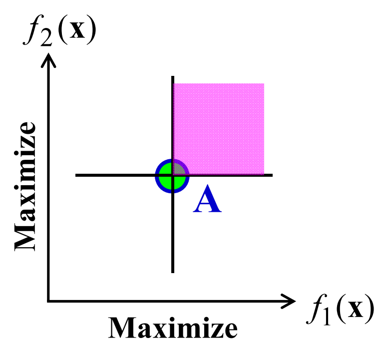


Solutions in this region are better than solution A. (1/16 of the space)

Pareto dominance-based comparison

Better Solution by Pareto Dominance

Pareto dominance-based comparison



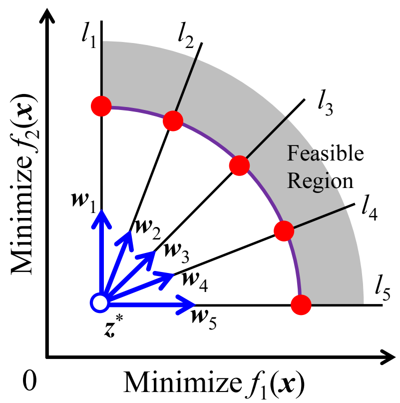
Percentage of the better region

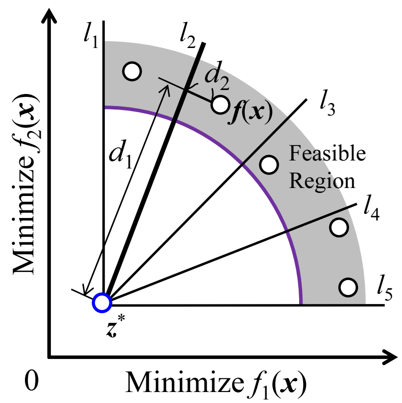
2 objectives	1/4	25%
3 objectives	1/8	13%
4 objectives	1/16	6%
5 objectives	1/32	3%
10 objectives	1/1024	0.1%
15 objectives	1/32768	0.003%
20 objectives	1/1048576	0.0001%

It is very difficult to find a better solution.

Use of Scalarizing Function (MOEA/D)

Recently MOEA/D has been very popular.
A scalarizing function is used in MOEA/D.

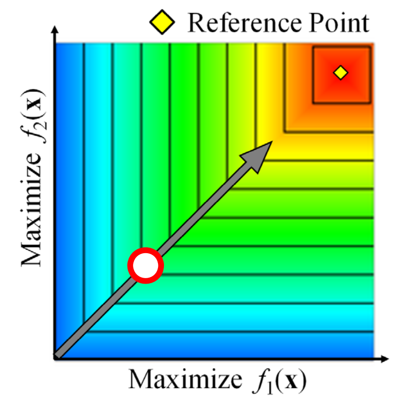




Use of Scalarizing Function

Weighted Tchebycheff

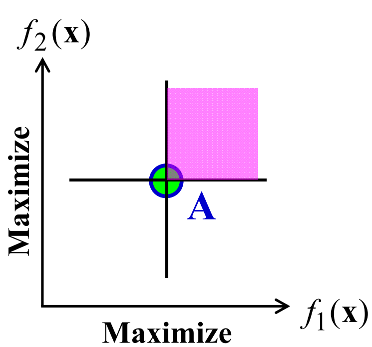
$$g^{TE}(\boldsymbol{x}\,|\,\boldsymbol{\lambda}, \boldsymbol{z}^*) = \max_{i=1,2,\ldots,m} \{\lambda_i \cdot |z_i^* - f_i(\boldsymbol{x})|\}$$



Contour lines of the Tchebycheff function

Use of Scalarizing Function

Weighted Tchebycheff



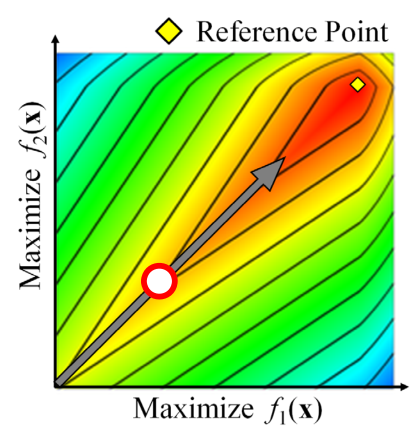
Percentage of the better region

2 objectives	1/4	25%
3 objectives	1/8	13%
4 objectives	1/16	6%
5 objectives	1/32	3%
10 objectives	1/1024	0.1%
15 objectives	1/32768	0.003%
20 objectives	1/1048576	0.0001%

Use of Scalarizing Function

PBI Function ($\theta = 5$)

$$g^{PBI}(\boldsymbol{x}\,|\,\boldsymbol{\lambda}, \boldsymbol{z}^*) = d_1 + \theta\, d_2$$

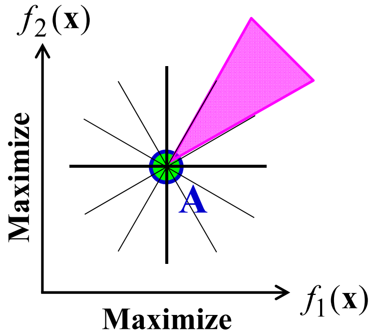


Contour lines of the PBI function

Use of Scalarizing Function

PBI Function (θ = 5)

Very Rough Calculation



Percentage of the better region

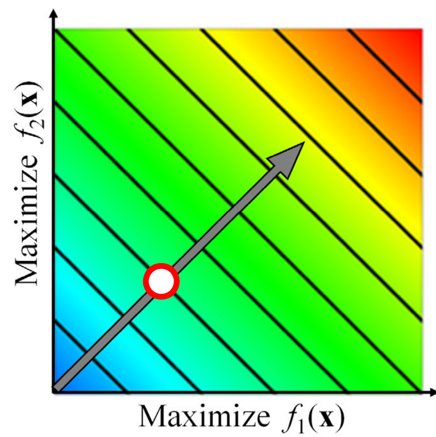
2 objectives	1/12	8%
3 objectives	1/36	3%
4 objectives	1/108	1%
5 objectives	1/324	0.3%
10 objectives	1/78732	0.001%
15 objectives		
20 objectives		

Much smaller than the case of the Pareto dominance.

Use of Scalarizing Function

Weighted Sum

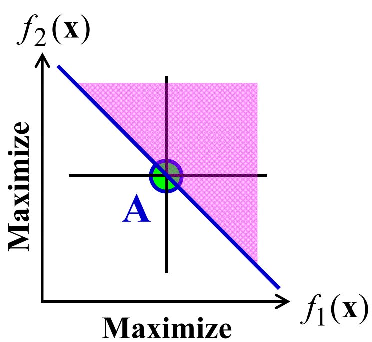
$$g^{WS}(\boldsymbol{x} \mid \boldsymbol{\lambda}) = \lambda_1 \cdot f_1(\boldsymbol{x}) + \lambda_2 \cdot f_2(\boldsymbol{x}) + \ \cdots \ + \lambda_m \cdot f_m(\boldsymbol{x})$$



Contour lines of the Weighted sum function

Use of Scalarizing Function

Weighted Sum



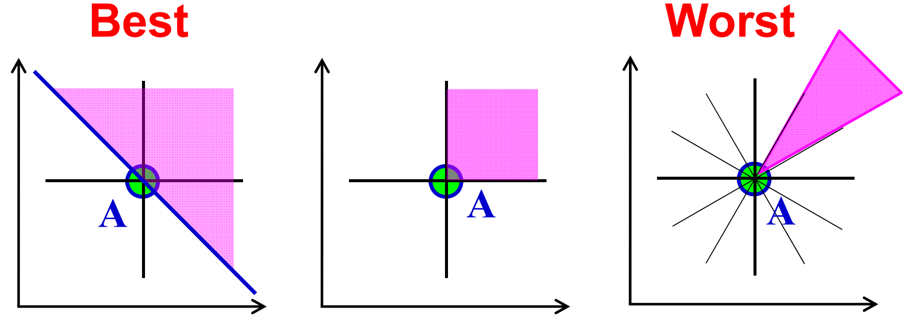
Percentage of the better region

2 objectives	1/2	50%
3 objectives	1/2	50%
4 objectives	1/2	50%
5 objectives	1/2	50%
10 objectives	1/2	50%
15 objectives	1/2	50%
20 objectives	1/2	50%

Always a half of the objective space is better.

Expected Performance of EMO Algorithms on Many-Objective Problems

Best **Worst**



Weighted Sum (MOEA/D-WS)

Pareto Dominance (NSGA-II) Tchebycheff (MOEA/D-Tch)

PBI Function (MOEA/D-PBI) (θ = 5)

Our Results on Knapsack Problems
Ishibuchi et al. IEEE TECV (2015)

Test Problems:
500-item knapsack problems with 2-10 objectives

Algorithms:
NSGA-II
MOEA/D with WS (Weighted Sum)
MOEA/D with Tchebycheff
MOEA/D with PBI ($\theta = 5$)

Performance Indicator:
Hypervolume

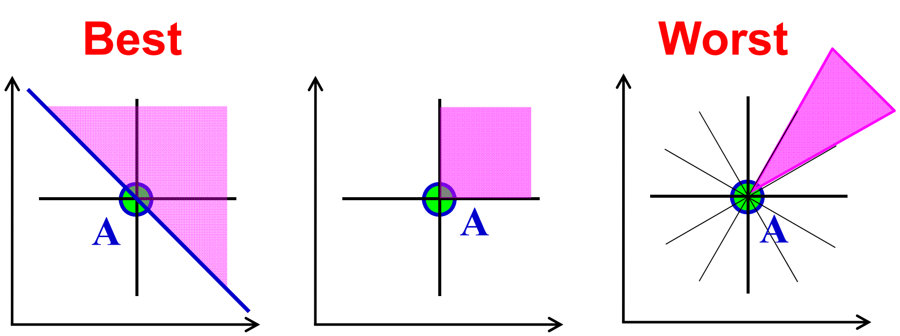
Expected difficulties are observed.

Our Results on Knapsack Problems
Ishibuchi et al. IEEE TECV (2015)

Average Hyper-Volume Value
(Normalized by the Result of the MOEA/D-WS)

EMO Algorithm	2-Obj	4-Obj	6-Obj	8-Obj	10-Obj
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
NSGA-II	96.5	86.2	77.8	72.0	65.5
MOEA/D: PBI (5)	100.9	89.3	73.8	67.4	61.9

Results = Expected Performance of EMO Algorithms on Many-Objective Problems



Weighted Sum (MOEA/D-WS) | Tchebycheff (MOEA/D-Tch) | PBI Function (MOEA/D-PBI) ($\theta = 5$)

Our Results on DTLZ Test Problems
Ishibuchi et al. IEEE TECV (2017)

Test Problems:
DTLZ1 - DTLZ4 Problems with 5-10 objectives

Algorithms:
NSGA-II
MOEA/D with WS (Weighted Sum)
MOEA/D with Tchebycheff
MOEA/D with PBI ($\theta = 5$)
NSGA-III
MOEA/DD

Performance Indicator:
Hypervolume

Totally different results are obtained.

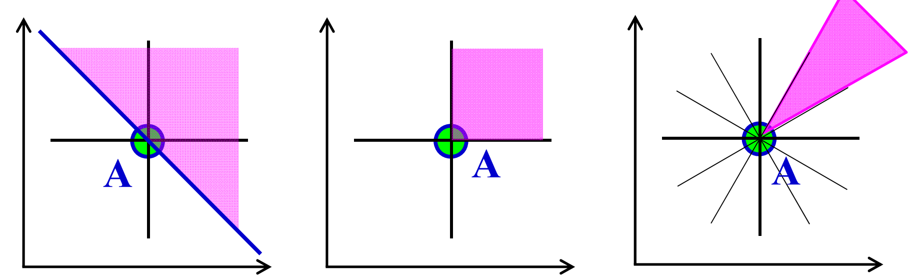
Our Results on DTLZ Test Problems
Ishibuchi et al. IEEE TECV (2017)

Average Hyper-Volume Value

Problem	M	NSGA-III	MOEA/DD	PBI	Tch	WS	NSGA-II
DTLZ 1	5	1.57677	1.57794	1.57768	1.51186	0.50052	0.00000
	8	2.13770	2.13730	2.13620	2.05463	0.96246	0.00000
	10	2.59280	2.59260	2.59220	2.51973	1.07913	0.00000
DTLZ 2	5	1.30317	1.30778	1.30728	1.14598	0.61944	0.67442
	8	1.96916	1.97862	1.97817	1.35469	0.68315	0.00004
	10	2.50878	2.51509	2.51500	1.69045	0.83883	0.00000
DTLZ 3	5	1.29894	1.30638	1.30398	1.14475	0.60143	0.00000
	8	1.95007	1.97162	1.74240	1.33166	0.66684	0.00000
	10	2.50727	2.51445	2.50933	1.69956	0.80348	0.00000
DTLZ 4	5	1.30839	1.30876	1.20680	1.00426	0.42941	1.00881
	8	1.98025	1.98083	1.86439	1.35100	0.71296	0.00000
	10	2.51524	2.51532	2.43536	1.56890	0.95488	0.00000

Results on DTLZ Test Problems
Totally different from the expected results

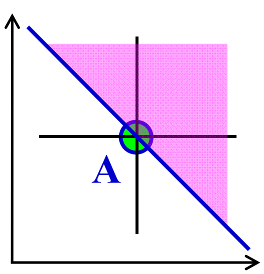
Worst | Best



Weighted Sum (MOEA/D-WS) | Tchebycheff (MOEA/D-Tch) | PBI Function (MOEA/D-PBI) (θ= 5)

Results on DTLZ Test Problems
Totally different from the expected results

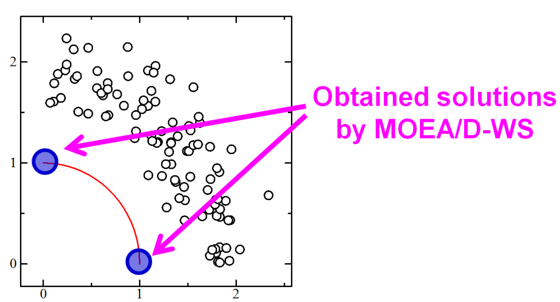
Worst



Weighted Sum (MOEA/D-WS)

Why ?

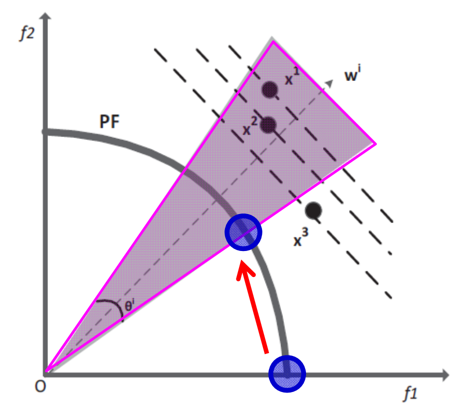
==> Because of the concave shape of the Pareto fronts !



A Promising Research Direction
Localized Weighted Sum

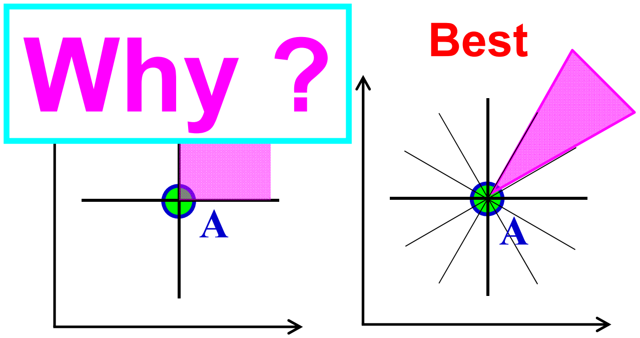
Localized weighted sum method for many-objective optimization IEEE TEVC 2018

Rui Wang, Zhongbao Zhou, Hisao Ishibuchi, Tianjun Liao, Tao Zhang



Results on DTLZ Test Problems

Totally different from the expected results



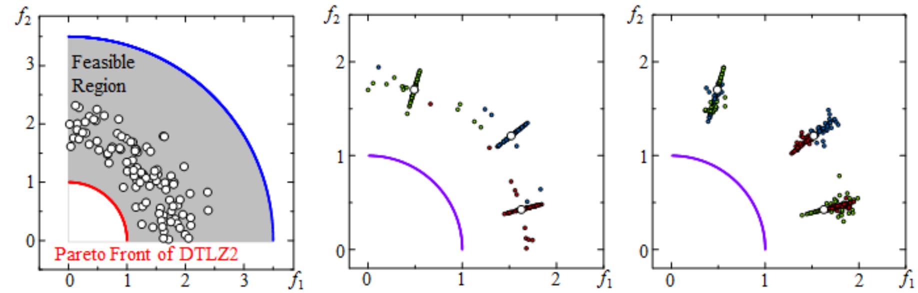
Tchebycheff
(MOEA/D-Tch)

PBI Function
(MOEA/D-PBI)
(θ = 5)

Reason

DTLZ test problems are very easy

DTLZ2



Feasible Region and Initial Solutions

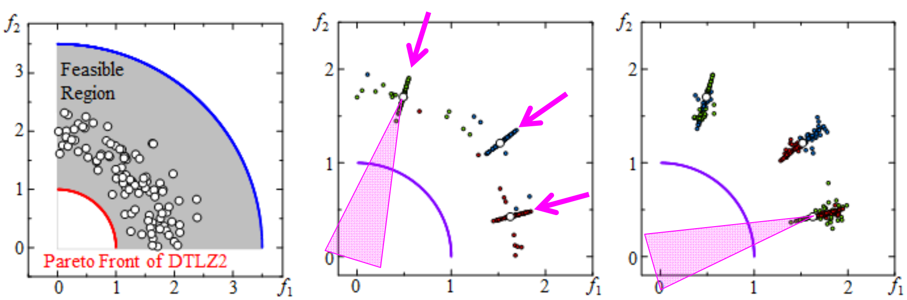
Generated Solutions by Mutation

Generated Solutions by Crossover

Reason

It is easy to find better solution.

DTLZ2



Feasible Region and Initial Solutions

Generated Solutions by Mutation

Generated Solutions by Crossover

Today's Plan

Difficulties in Evolutionary Many-Objective Optimization Studies

1. Difficulties related to many-objective search
2. Difficulties related to test problems
3. Difficulties related to performance evaluation

Typical Scenario of Many-Objective Optimization Papers

Motivation:
- Many-objective optimization problems are difficult.
- New algorithms are needed.

Proposal:
- We propose a new high-performance algorithm.

Computational Experiments:
- Better results are obtained by the proposed algorithm than the existing ones on DTLZ 1-4 and WFG 1-9 problems.

Test Problems in Recent Many-Objective Papers

Publication Year	Proposed Algorithm	Test Problems	Number of Objectives
2014	NSGA-III	DTLZ 1-4 WFG 6-7 S-DTLZ 1-2	3, 5, 8, 10, 15 3, 5, 8, 10, 15 3, 5, 8, 10, 15
2015	I-DBEA	DTLZ 1-4 DTLZ5(I, M) WFG 1-9	3, 5, 8, 10, 15 3, 5, 8, 10, 15 3, 5, 10, 15
2015	MOEA/DD	DTLZ 1-4 WFG 1-9	3, 5, 8, 10, 15 3, 5, 8, 10
2016	MOEA/D-DU EFR-RR	DTLZ 1-4, 7 WFG 1-9 S-DTLZ 1-2	2, 5, 8, 10, 13 2, 5, 8, 10, 13 2, 5, 8, 10, 13
2016	θ-DEA	DTLZ 1-4, 7 WFG 1-9 S-DTLZ 1-2	3, 5, 8, 10, 15 3, 5, 8, 10, 15 3, 5, 8, 10, 15

High-Performance Evolutionary Many-Objective Algorithms

2007 MOEA/D

2014 NSGA-III

2015 I-DBEA

2015 MOEA/DD

2016 θ - DEA

Better Results on DTLZ and WFG

(New algorithms are better than old ones).

Typical Scenario of Many-Objective Optimization Papers

Motivation:
- Many-objective optimization problems are difficult.
- New algorithms are needed.

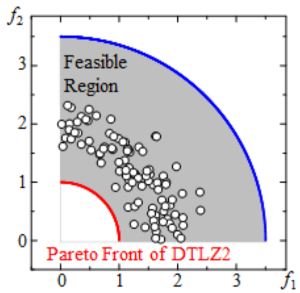
Proposal:
- We propose a new high-performance algorithm.

Computational Experiments:
- Better results are obtained by the proposed algorithm than the existing ones on DTLZ 1-4 and WFG 1-9 problems.

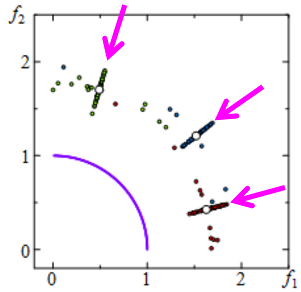
Test problems are easy and have special features.

Special Feature: Better new solutions can be easily created by

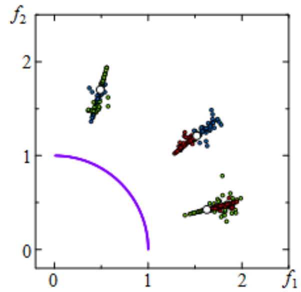
DTLZ2



Feasible Region and Initial Solutions

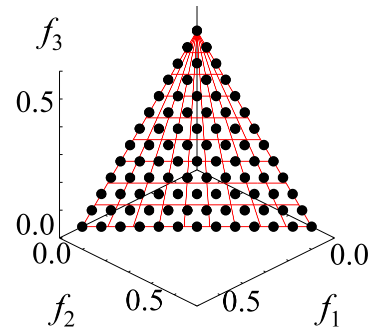


Generated Solutions by Mutation

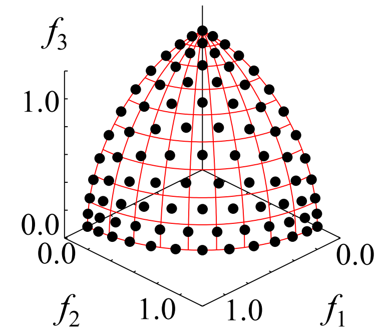


Generated Solutions by Crossover

Special Feature: DTLZ 1-4 and WFG 4-9 have triangular Pareto fronts



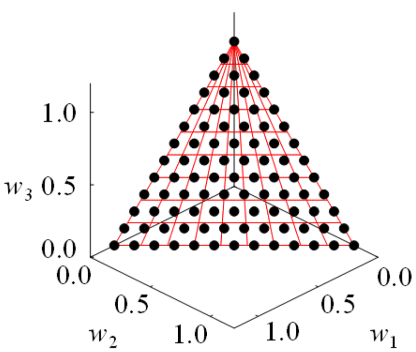
DTLZ 1



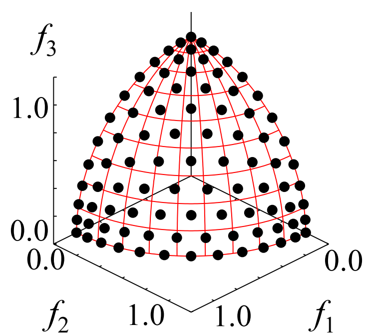
DTLZ 2

MOEA/D and Test Problems

MOEA/D looks perfect for DTLZ



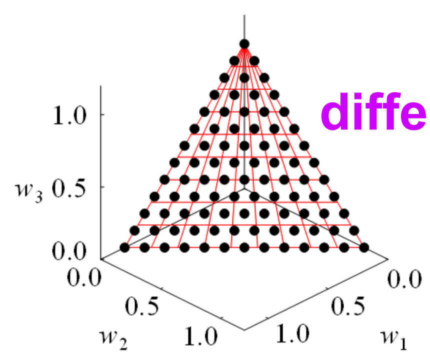
Weight Vectors



Pareto front (DTLZ 2)

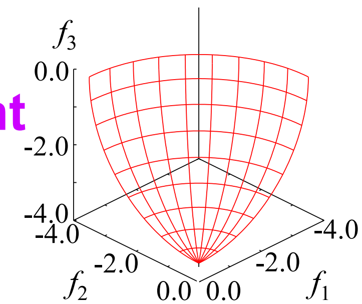
Shape of the Pareto front for MOEA/D:

The point is whether the shape of the Pareto front is similar to the shape of the weight vector distribution.



Weight Vectors

different

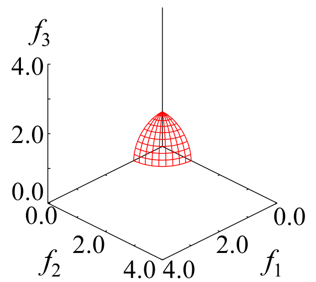


Pareto front (Minus-DTLZ 2)

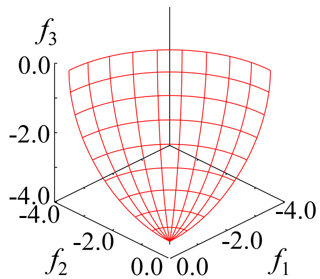
Our Idea: Min-DTLZ and Min-WFG
Ishibuchi et al. IEEE TEVC (2017)

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

Change from "minimization" to "maximization" is the same as the multiplication by (- 1).

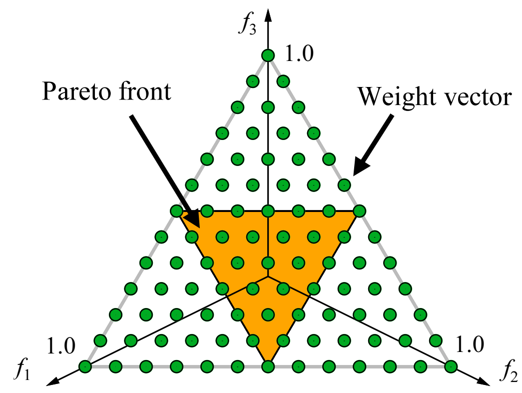


Pareto front of DTLZ2

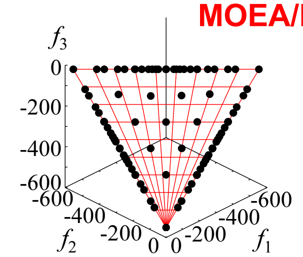


Pareto front of (-1) x DTLZ2

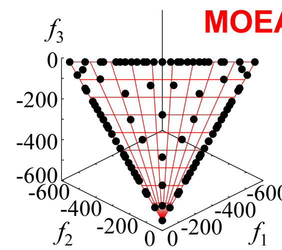
Experimental Results on (- 1) x DTLZ1



MOEA/D-PBI



MOEA/D-Tch



Our Results on Minus-DTLZ Test Problems
Ishibuchi et al. IEEE TECV (2017)

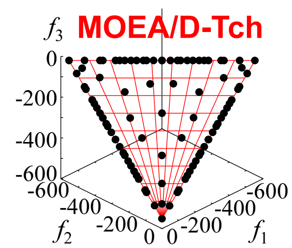
Average Hyper-Volume Value

Problem	*M*	NSGA-III	MOEA/DD	PBI	Tch	WS	NSGA-II
Minus	5	0.01265	0.00972	0.01739	0.01208	0.00083	0.01520
DTLZ 1	8	5.227E-05	0.881E-05	0.598E-05	3.215E-05	0.139E-05	3.568E-05
	10	1.185E-06	0.100E-06	0.079E-06	0.620E-06	0.025E-06	0.765E-06
Minus	5	0.13957	0.08794	0.15984	0.15556	0.14930	0.17147
DTLZ 2	8	4.454E-03	2.690E-03	5.978E-03	0.459E-03	1.560E-03	4.585E-03
	10	6.308E-04	1.836E-04	5.199E-04	0.052E-04	0.640E-04	3.797E-04
Minus	5	0.12951	0.08190	0.15902	0.15199	0.14891	0.16472
DTLZ 3	8	0.00414	0.00255	0.00596	0.00050	0.00156	0.00390
	10	0.00054	0.00018	0.00052	0.00001	0.00006	0.00033
Minus	5	0.12326	0.07242	0.12296	0.14878	0.14881	0.16970
DTLZ 3	8	4.582E-03	2.198E-03	2.020E-03	0.485E-03	1.563E-03	3.886E-03
	10	6.065E-04	2.569E-04	2.333E-04	0.043E-04	0.642E-04	3.006E-04

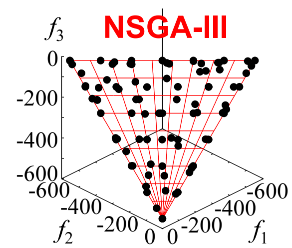
Experimental Results
(Hypervolume)

DTLZ and WFG			(-1) x DTLZ and (-1) x WFG
MOEA/D (1997)		Better	MOEA/D (1997)
NSGA-III (2014)			NSGA-III (2014)
MOEA/DD (2015)			MOEA/DD (2015)
θ-DEA (2016)	Better		θ-DEA (2016)

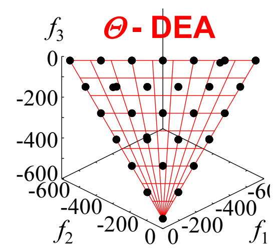
MOEA/D-Tch



NSGA-III



Θ - DEA



Today's Plan (Part 1)

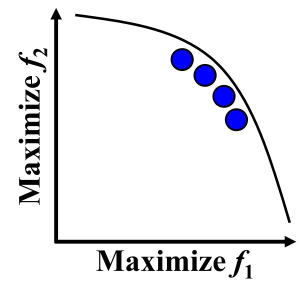
Difficulties in Evolutionary Many-Objective Optimization Studies

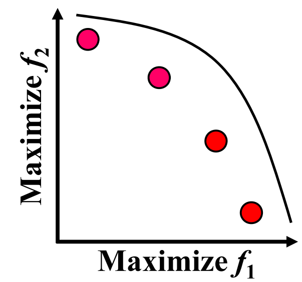
1. Difficulties related to many-objective search
2. Difficulties related to test problems
3. Difficulties related to performance evaluation

Two-objective Optimization

The final result of optimization is a solution set.
Comparison of solution sets is not easy.

Which is a better solution set?

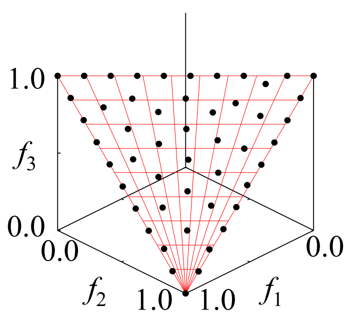




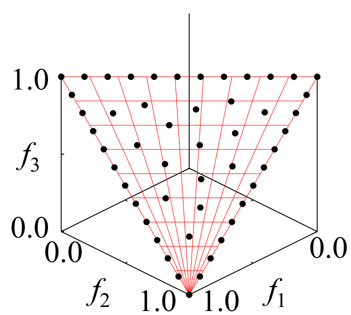
Three-objective Optimization

The final result of optimization is a solution set.
Comparison of solution sets is difficult:

Which is a better solution set?



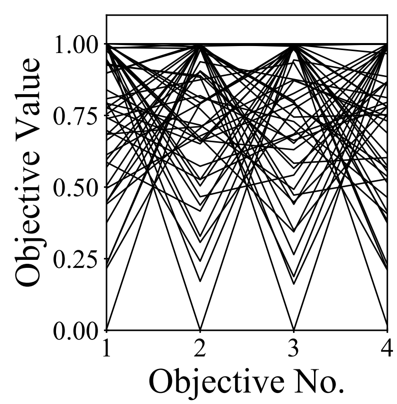
(a)



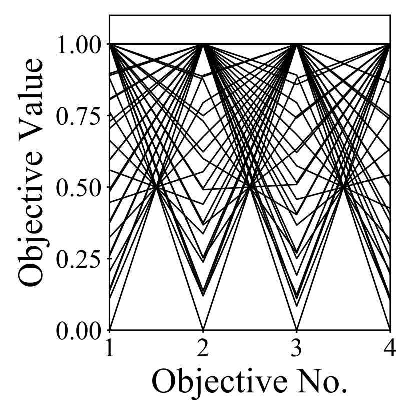
(b)

Four-objective Optimization

Which is the better solution set?



(a)

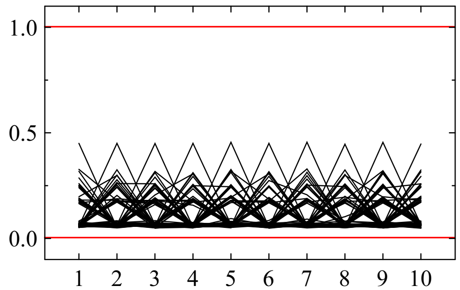


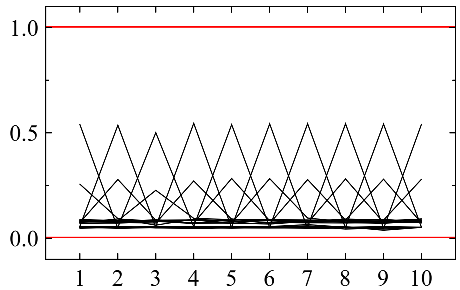
(b)

Ten-objective Optimization

Maximize $f_1(\boldsymbol{x}),\ f_2(\boldsymbol{x}),\ ...,\ f_{10}(\boldsymbol{x})$

The final result of optimization is a solution set.
Comparison of solution sets is very difficult.





Difficulties in Performance Evaluation

1. **How to Specify the Population Size**
2. **How to Specify the Reference Point for HV**
3. **How to Specify the Reference Points for IGD**

[1] H. Ishibuchi et al. , How to compare many-objective algorithms under different settings of population and archive sizes, *Proc. of CEC 2016*, pp. 1149-1156. (Proposal of the Basic Idea)

[2] R. Tanabe, H. Ishibuchi, A. Oyama, Benchmarking multi- and many-objective evolutionary algorithms under two optimization scenarios, *IEEE Access*, vol. 5, pp. 19597-19619, December 2017. (Performance Comparison Results)

Performance Indicators

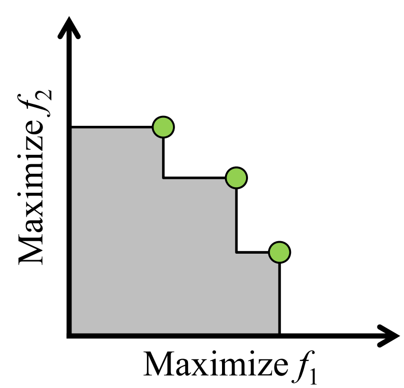
Frequently-Used Performance Indicators

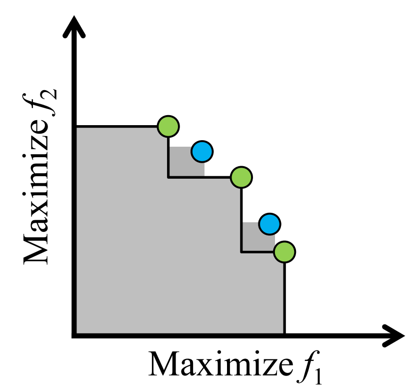
1. **Hypervolume Indicator**
2. **IGD (Inverted Generational Distance) Indicator**

Property of These Indicators:
By increasing the number of solutions, the evaluation of a solution set by these indicators can be improved.

Hypervolume

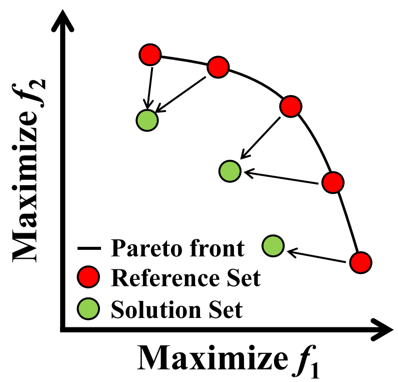
Hypervolume (HV) is the volume of the dominated region by the obtained solutions. The HV value can can be improved by adding new solutions.

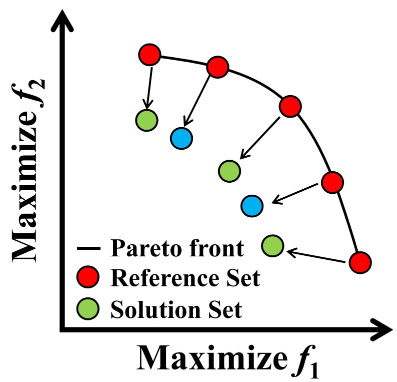




IGD: Inverted Generational Distance

Average distance from each reference point on the Pareto front to the nearest solution. The IGD value can be improved by adding new solutions.





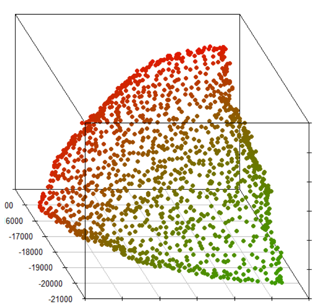
Specification of Population Size
How about the following settings?

Algorithm A:
- Crossover probability: 1.0
- Mutation probability: 1/*n* (*n*: string length)
- **Population size: 5,000**

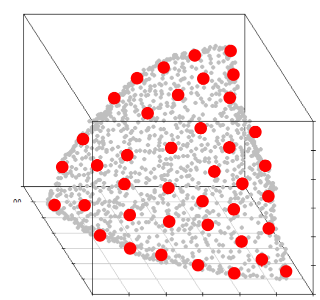
Algorithm B:
- Crossover probability: 0.2
- Mutation probability: 5/*n* (*n*: string length)
- **Population size: 50**

Comparison under these settings may be OK for single-objective optimization. However, for multi-objective optimization, ...

Obtained Solution Sets



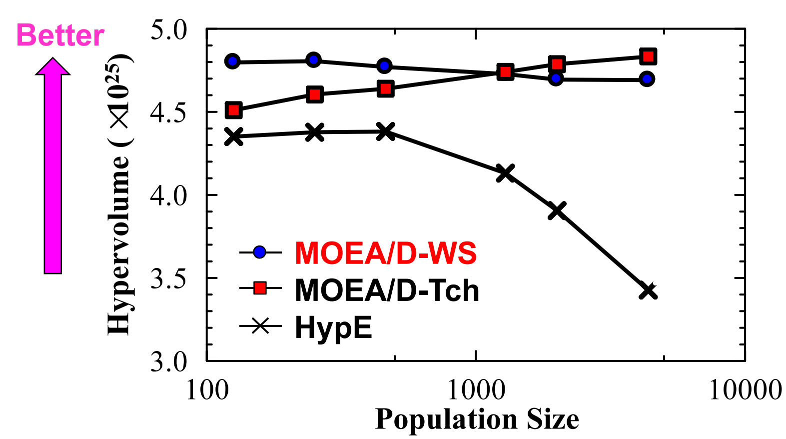
Algorithm A



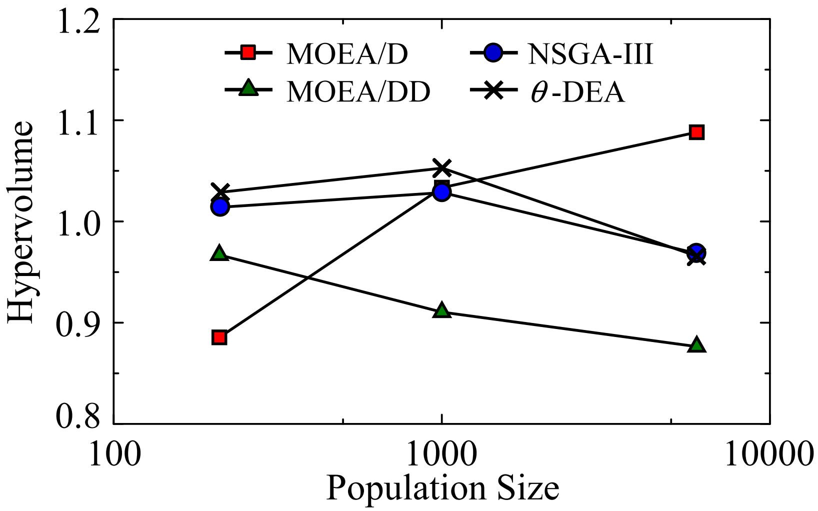
Algorithm B

Experimental Results
Under various settings of the population size

Results on a Six-objective 500-item Knapsack Problem



Other Results: Five-Objective WFG3
MOEA/D can be the worst and the best.

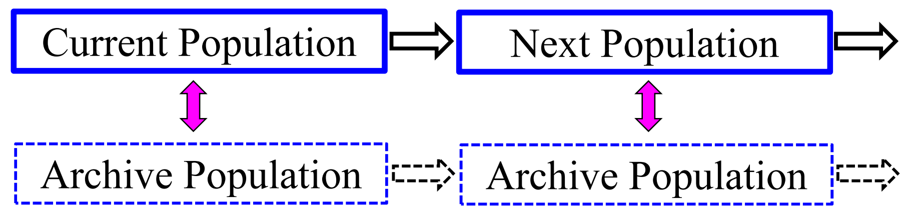


Performance of the final population of different size

Question

How to compare EMO algorithms with/without an archive population?

Some algorithms have an archive population whereas others do not have.



Our Idea (CEC 2016): Solution selection from all the examined solutions

Algorithm A:

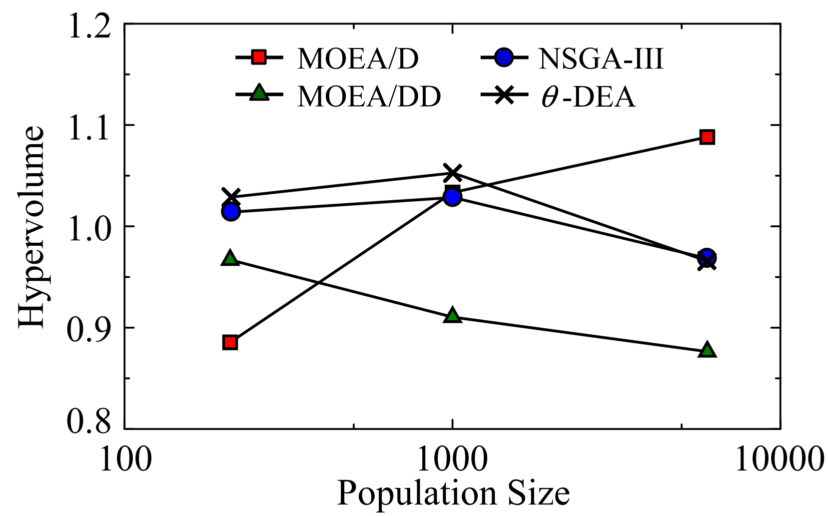
Crossover probability: 1.0
Mutation probability: 1/n (n: string length)
Population size: 100
Size of Archive Population: 1,000

Algorithm B:

Crossover probability: 0.2
Mutation probability: 5/n (n: string length)
Population size: 100
No Archive Population

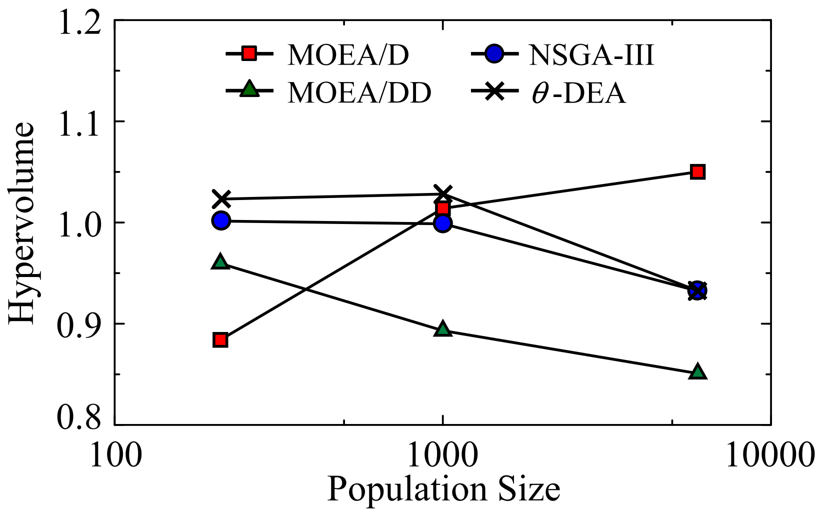
The comparison may be unfair ==> Solution selection from all the examined solutions.

Performance of the Final Population Five-Objective WFG3



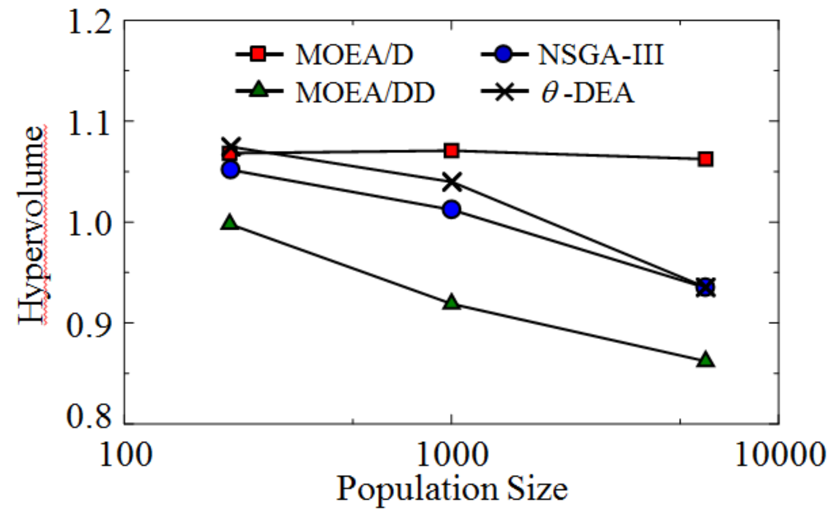
Performance of the final population of different size

Selection of 50 Solutions from the Final Population



Performance of the selected 50 solutions

Selection of 50 Solutions from all the Examined Solutions



Performance of the selected 50 solutions

Performance Comparison using Solution Selection Methods

R. Tanabe, H. Ishibuchi, and A. Oyama, "Benchmarking multi- and many-objective evolutionary algorithms under two optimization scenarios," IEEE Access, Dec 2017.

Two Optimization Scenarios:

(i) Use of the final population

(ii) Use of selected solutions from the examined solutions

Observation: Performance comparison results are different between the two optimization scenarios.

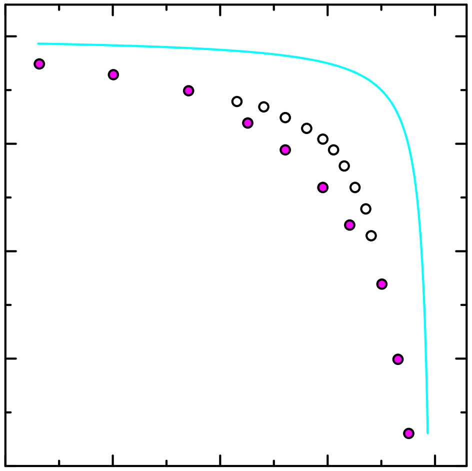
Difficulties in Perfoamance Evaluation

1. How to Specify the Population Size
2. How to Specify the Reference Point for HV
3. How to Specify the Reference Points for IGD

[1] H. Ishibuchi et al., Reference point specification in hypervolume calculation for fair comparison and efficient search, *Proc. of GECCO 2017*, pp. 585-592. (Proposal of the Basic Idea)

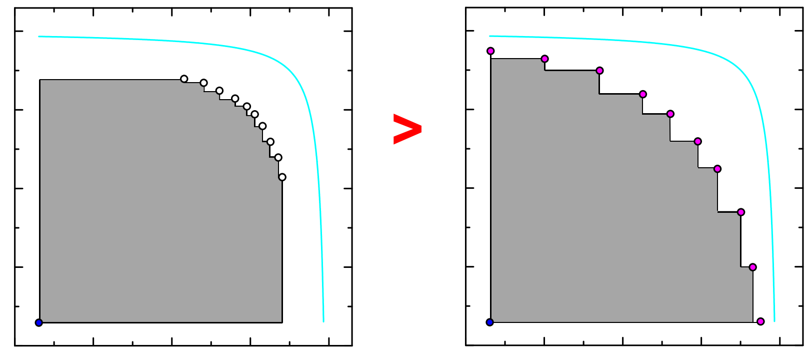
[2] H. Ishibuchi et al., How to specify a reference point in hypervolume calculation for fair performance comparison," *Evolutionary Computation* (2018). (Extended Journal Version)

Two Solution Sets (maximization)
Which has the larger hypervolume?



Hypervolume (HV)
Comparison results depends on the reference point

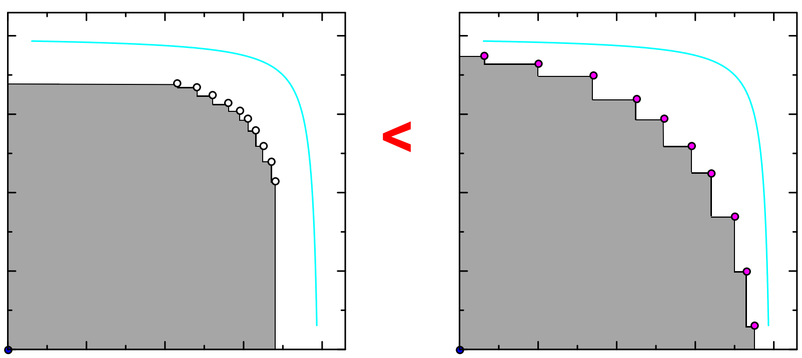
When the reference point is close to the Pareto front:



Better Solution Set

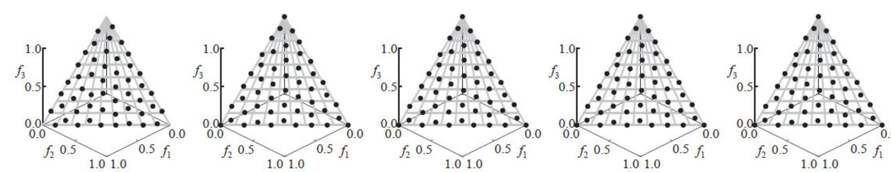
Hypervolume (HV)
Comparison results depends on the reference point

When the reference point is far from the Pareto front:



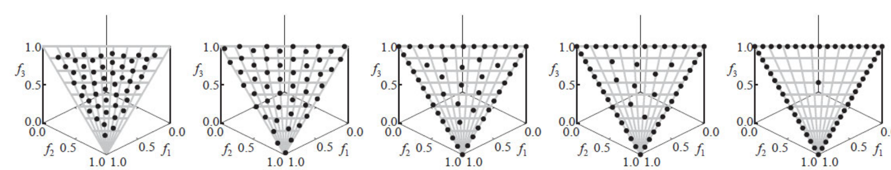
Better Solution Set

HV: Dependency of Optimal Distribution of Solutions on the Shape of the Pareto Front



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

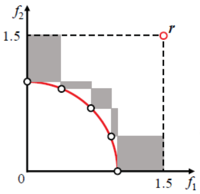
Figure 1: Obtained solution sets for the three-objective normalized DTLZ1.



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

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

HV: Dependency of Optimal Distribution of Solutions on the Shape of the Pareto Front



(a) Two-objective DTLZ2.

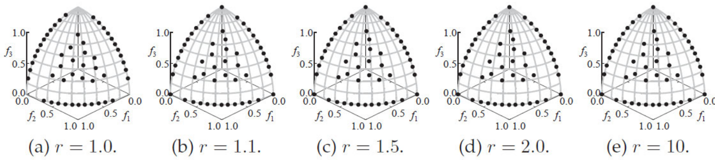
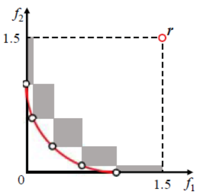


Figure 8: Obtained solution sets for the three-objective DTLZ2 problem.



(b) Two-objective Minus-DTLZ2.

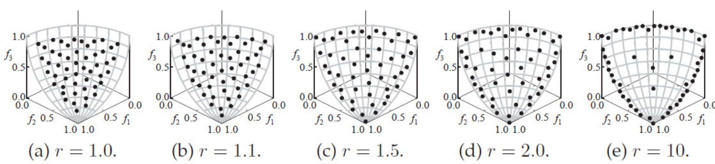


Figure 9: Obtained solution sets for the three-objective Minus-DTLZ2 problem.

HV: Dependency of Optimal Distribution of Solutions on the Shape of the Pareto Front

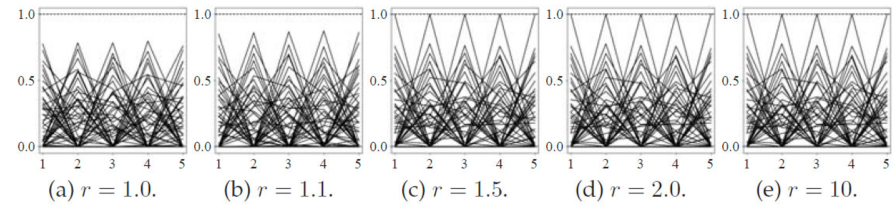


Figure 6: Obtained solution sets for the five-objective normalized DTLZ1 problem.

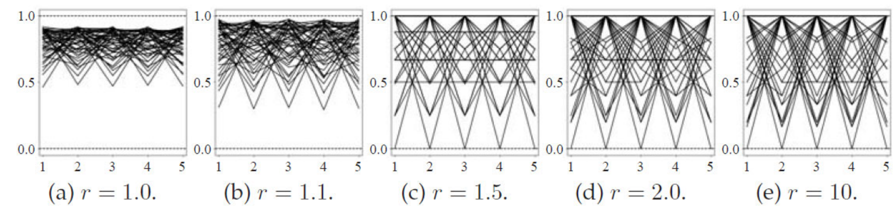


Figure 7: Obtained solution sets for the five-objective Minus-DTLZ1 problem.

Difficulties in Performance Evaluation

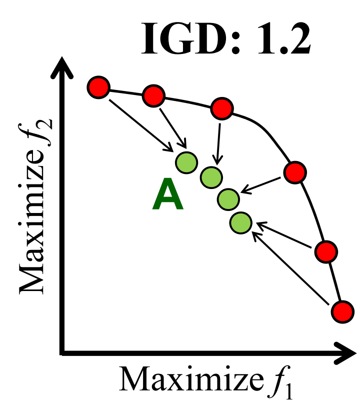
1. How to Specify the Population Size
2. How to Specify the Reference Point for HV
3. How to Specify the Reference Points for IGD

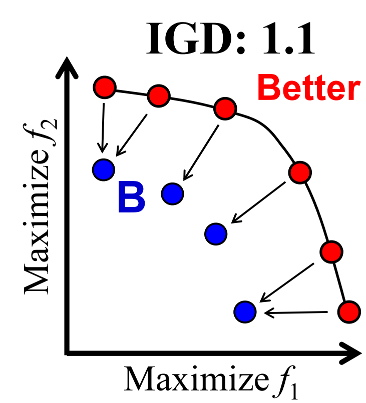
[1] H. Ishibuchi et al., Reference point specification in inverted generational distance for triangular linear Pareto front, *IEEE Trans. on Evolutionary Computation* (2018). (Reference Point Specification)

[2] H. Ishibuchi, H. Masuda, Y. Nojima, A study on performance evaluation ability of a modified inverted generational distance indicator," *Proc. of GECCO 2015*, pp. 695-702. (Modification of the IGD Indicator)

IGD-based performance comparison results depends on the reference point specifications

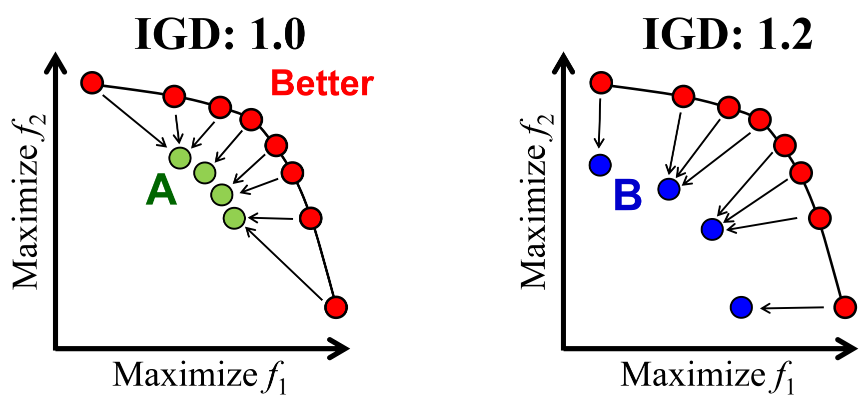
Specification of reference points is important.





IGD-based performance comparison results depends on the reference point specifications

Specification of reference points is important.

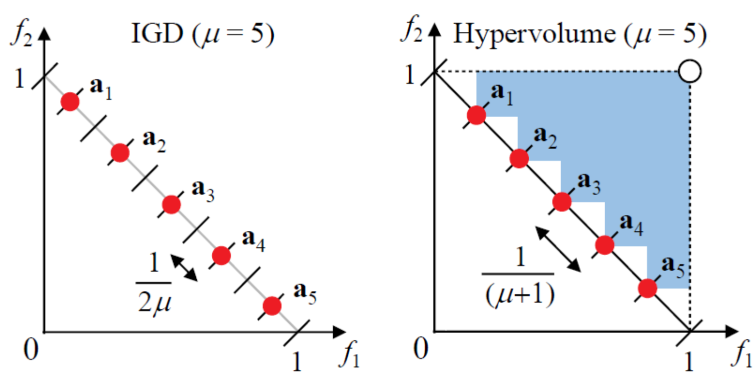


How to specify a set of reference points

Current Standard:
Use of a large number of uniformly distributed solutions.

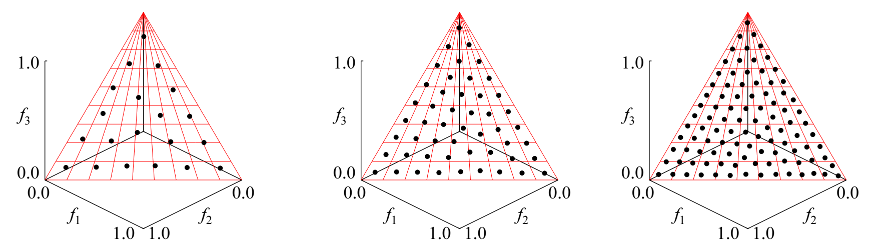
This is not always a good method as shown in the following slides.

Analysis of IGD from a Viewpoint of Optimal Distribution of Solutions

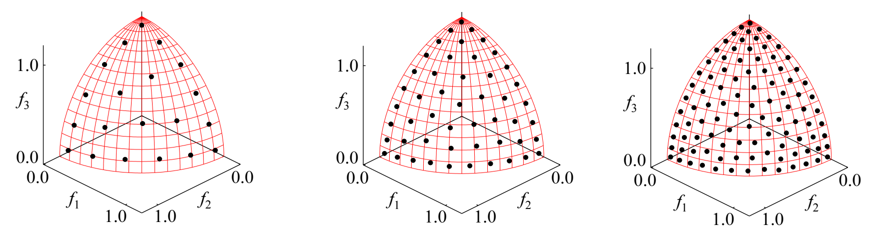


IEEE Trans. on Evolutionary Computation (2018)
Reference Point Specification in Inverted Generational Distance for Triangular Linear Pareto Front
Hisao Ishibuchi, Ryo Imada, Yu Setoguchi, and Yusuke Nojima

Optimal Distributions of Solutions for IGD are not always intuitive

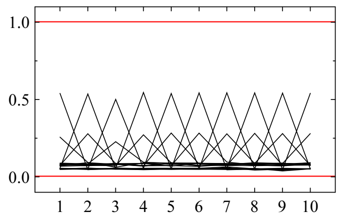


Population size 20 **Population size 50** **Population size 100**

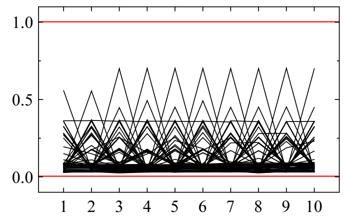


Population size 20 **Population size 50** **Population size 100**

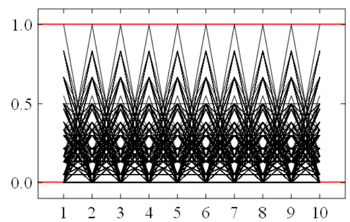
Optimal Distributions of Solutions for IGD are not always intuitive



Population size 20



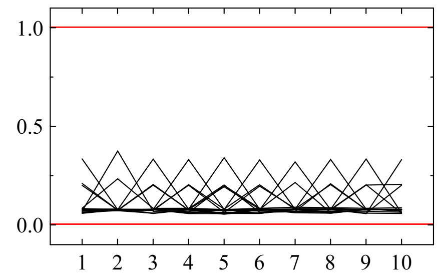
Population size 100



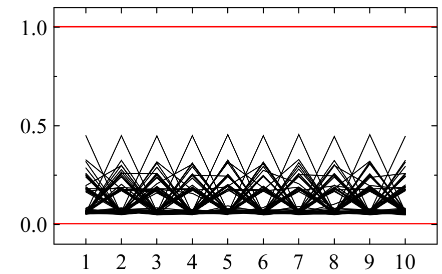
Reference points: 10,010 points

Optimal Distributions of Solutions for IGD are not always intuitive

When we randomly generate 100,000 reference points, the optimal distributions of solutions are as follows:



Population size 20



Population size 100

Part 2: Approaches

We explain why many-objective optimization is difficult for EMO (Evolutionary Multiobjective Optimization) algorithms, and how those difficulties can be handled. We also suggest some promising future research directions

Part 1 (Hisao Ishibuchi): Difficulties

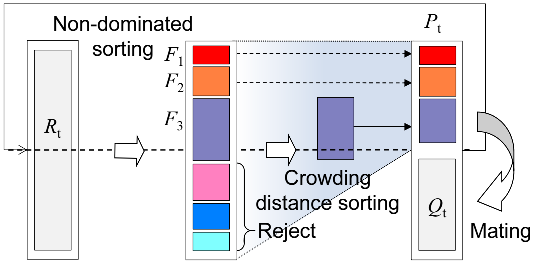
Part 2 (Hiroyuki Sato): Approaches

Part 3 (Hisao Ishibuchi): Future Directions

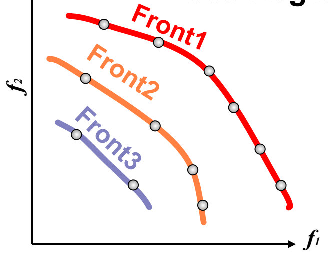
Approaches for Many-Objective Optimization

1. **Relaxed Dominance Based Approach**
2. **Indicator Based Approach**
3. **Decomposition Based Approach**
4. **Reference Based Approach**
5. **Dimensionality Reduction Approach**
6. **Efficient Solution Generation**

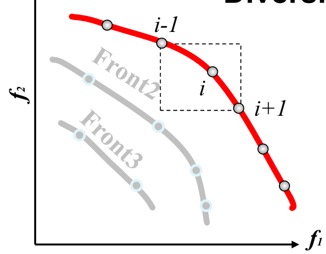
NSGA-II [Deb 02]



■ **Non-dominated sorting**
⇒ Convergence

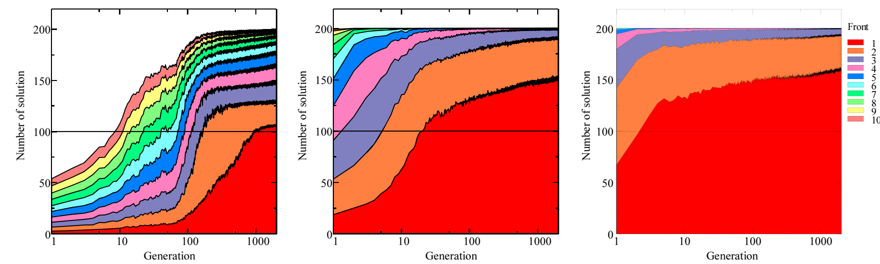


■ **Crowding Distance**
⇒ Diversity



Front Distribution over Generation

Results on knapsack problems with 500 items



2 objectives

5 objectives
Coarse

10 objectives
Coarse

In many-objective problems,

- The number of solutions belonging to Front 1 exceeds the size of parent solutions in early stage of the evolution.
- **Convergence of solutions toward Pareto front is deteriorated.**

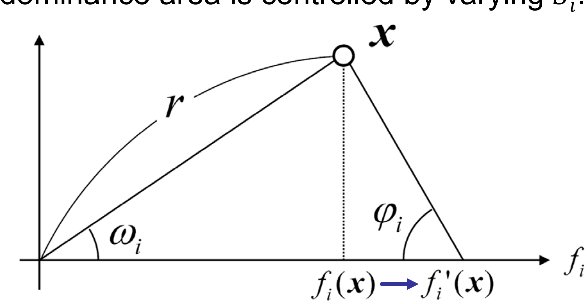
Approaches for Many-Objective Optimization

1. **Relaxed Dominance Based Approach**
2. **Indicator Based Approach**
3. **Decomposition Based Approach**
4. **Reference Based Approach**
5. **Dimensionality Reduction Approach**
6. **Efficient Solution Generation**

Relaxed Dominance Based Approach
Transform of Objective Values

Controlling Dominance Area of Solutions [Sato 07]

– The dominance area is controlled by varying S_i.

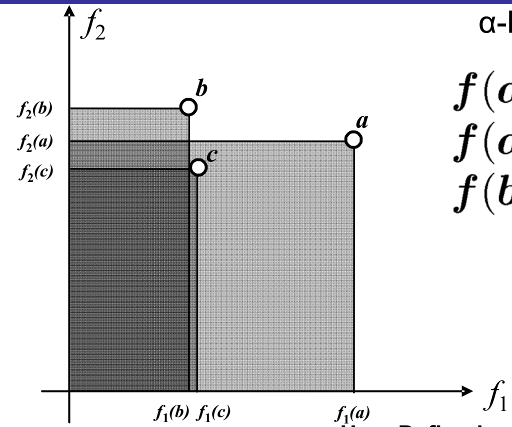


$$f_i'(x) = \frac{r \cdot \sin(\omega_i + S_i \cdot \pi)}{\sin(S_i \cdot \pi)} \qquad (i = 1, 2, \cdots, m)$$

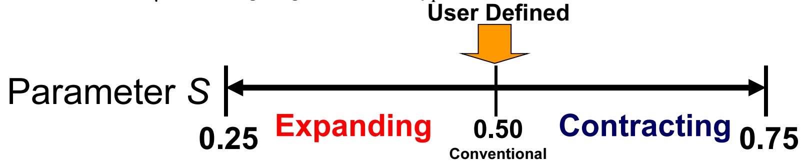
Where, $S_i = \varphi_i / \pi$

Controlling Dominance Area of Solutions (CDAS) [Sato 07]

α-Domination [Ikeda 01]

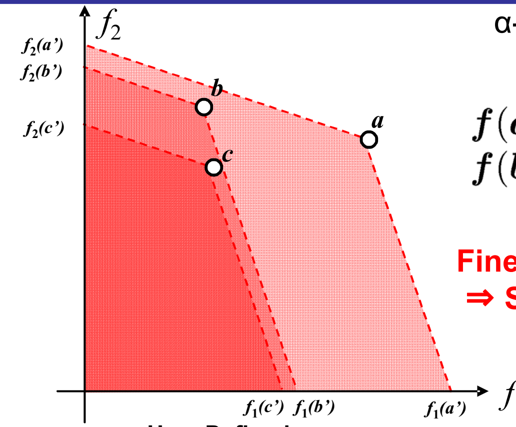


$$f(a) \succeq f(c)$$
$$f(a) \not\succeq f(b)$$
$$f(b) \not\succeq f(c)$$



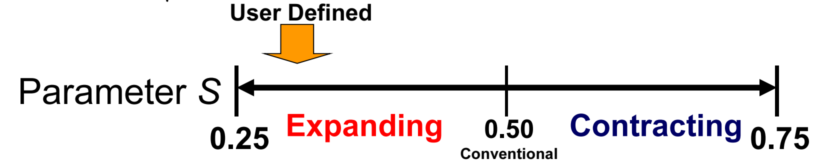
Controlling Dominance Area of Solutions (CDAS) [Sato 07]

α-Domination [Ikeda 01]



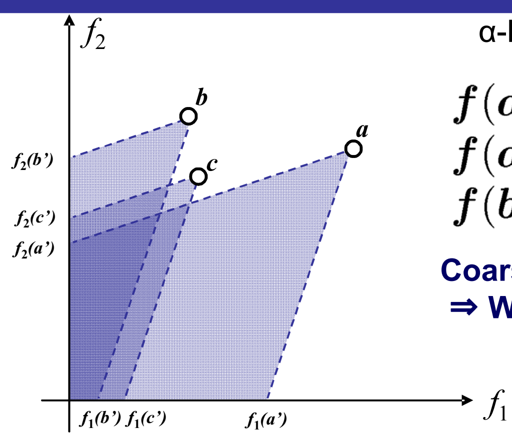
$$f(a') \succeq \{f(b'), f(c')\}$$
$$f(b') \succeq f(c')$$

Fine grained ranking
⇒ Strengthen selection



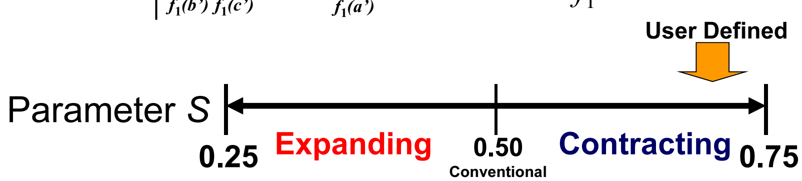
Controlling Dominance Area of Solutions (CDAS) [Sato 07]

α-Domination [Ikeda 01]



$$f(a') \not\succeq f(b')$$
$$f(a') \not\succeq f(c')$$
$$f(b') \not\succeq f(c')$$

Coarser ranking
⇒ Weaken selection

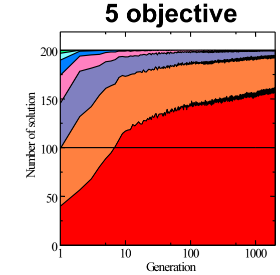


Front Distribution over Generation

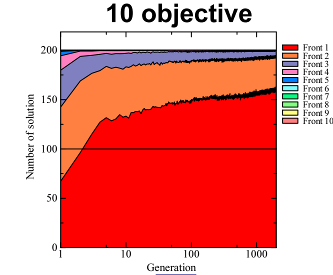
Conventional Dominance Area

Results on knapsack problems with 500 items

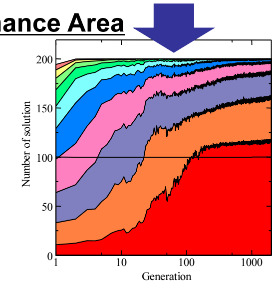
5 objective



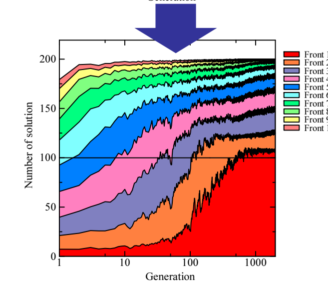
10 objective



Expanded Dominance Area



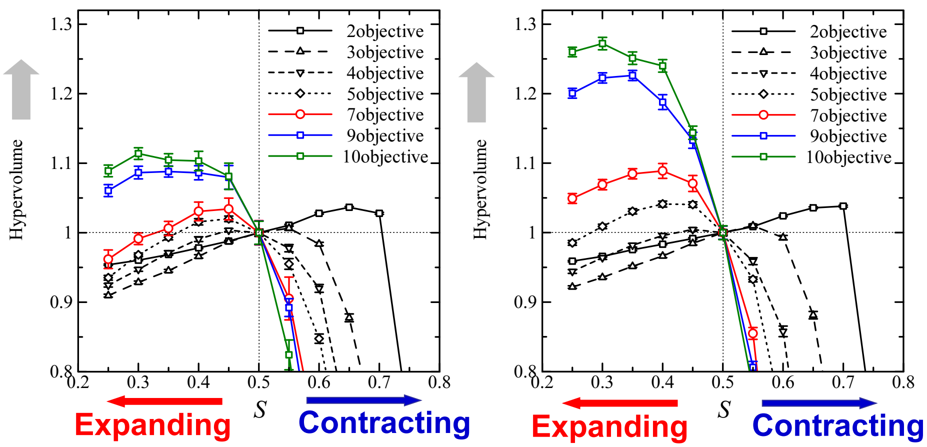
S^*=0.40



S^*=0.30

Hypervolume by Varying *S*

Results on knapsack problems with 500 items

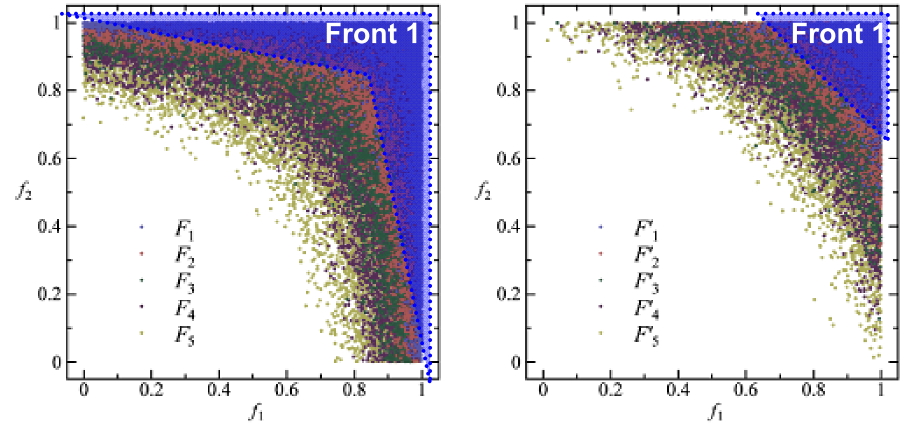


NSGA-II with CDAS **SPEA2 with CDAS**

Front Distribution in Objective Space

Simple experiment

- Randomly generated 100 points in $[0,1]^2$ are classified into non-dominated fronts.
- The same process is repeated 1,000 times.



Conventional Dominance (S=0.5) **Expanded Dominance (*S*=0.4)**

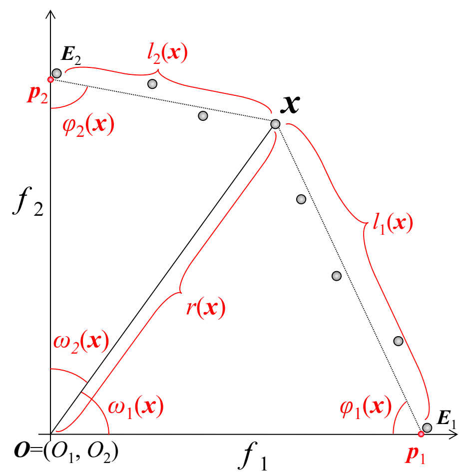
Problems in CDAS

1. An appropriate parameter *S* controlling dominance area of solutions must be found out experimentally.
2. The diversity of solutions deteriorates when we decrease *S* from 0.5.

- **To solve these problems**
  - A variant of CDAS called self-controlling dominance area of solutions (S-CDAS) was proposed.
  - The algorithm self-controls dominance area for each solution without the need of an external parameter.

Relaxed Dominance Based Approach
Modification of Objective Values

S-CDAS determines dominance area that extreme solutions $\boldsymbol{E}$ are never dominated.



■For a single solution $\boldsymbol{x}$, calculate $\boldsymbol{\varphi}(\boldsymbol{x})=\{\varphi_1, \varphi_2,\ldots, \varphi_m\}$. Here, $\varphi_i(\boldsymbol{x})$ is the angle determined by the solution $\boldsymbol{x}$ and the landmark vector p_i in the i-th objective function.

$$\varphi_i(\boldsymbol{x}) = \sin^{-1}\left\{\frac{r(\boldsymbol{x})\cdot\sin(\omega_i(\boldsymbol{x}))}{l_i(\boldsymbol{x})}\right\} \quad (i=1,2,\ldots,m)$$

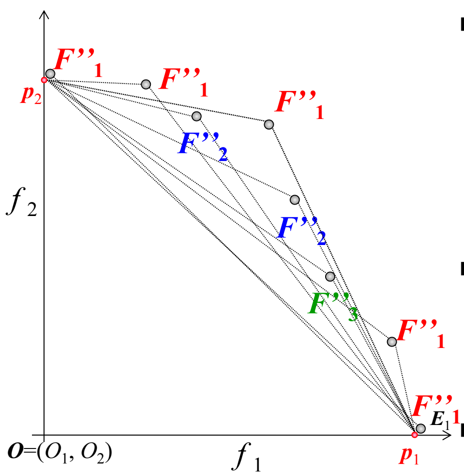
■ Modify objective values of all solutions $\boldsymbol{y}\in F_i$ by the following equation.

$$f_i'(\boldsymbol{y}) = \frac{r_i(\boldsymbol{y})\cdot\sin(\omega_i(\boldsymbol{y})+\varphi_i(\boldsymbol{x}))}{\sin(\varphi_i(\boldsymbol{x}))} \quad (i=1,2,\ldots,m)$$

■Check dominance relations between the solution $\boldsymbol{x}$ and all other solutions $\boldsymbol{y}\in F_j$. If a solution $\boldsymbol{y}\in F_j$ is dominated by $\boldsymbol{x}$, the counter (rank) of $\boldsymbol{y}$ is incremented.

Relaxed Dominance Based Approach
Modification of Objective Values

S-CDAS determines dominance area that extreme solutions $\boldsymbol{E}$ are never dominated.



■For a single solution $\boldsymbol{x}$, calculate $\boldsymbol{\varphi}(\boldsymbol{x})=\{\varphi_1, \varphi_2,\ldots, \varphi_m\}$. Here, $\varphi_i(\boldsymbol{x})$ is the angle determined by the solution $\boldsymbol{x}$ and the landmark vector p_i in the i-th objective function.

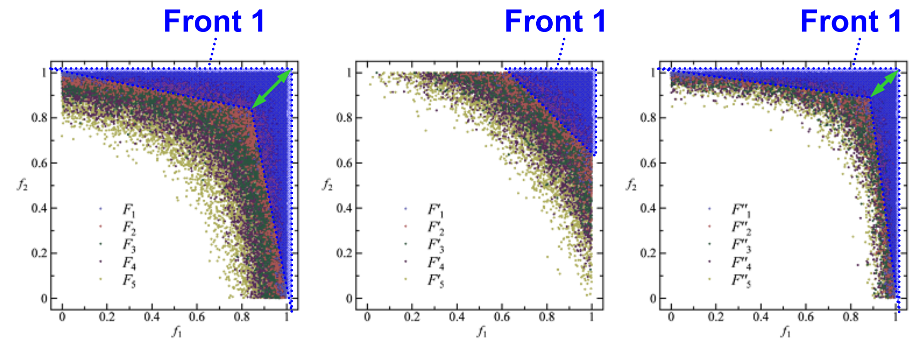
$$\varphi_i(\boldsymbol{x}) = \sin^{-1}\left\{\frac{r(\boldsymbol{x})\cdot\sin(\omega_i(\boldsymbol{x}))}{l_i(\boldsymbol{x})}\right\} \quad (i=1,2,\ldots,m)$$

■ Modify objective values of all solutions $\boldsymbol{y}\in F_i$ by the following equation.

$$f_i'(\boldsymbol{y}) = \frac{r_i(\boldsymbol{y})\cdot\sin(\omega_i(\boldsymbol{y})+\varphi_i(\boldsymbol{x}))}{\sin(\varphi_i(\boldsymbol{x}))} \quad (i=1,2,\ldots,m)$$

■Check dominance relations between the solution $\boldsymbol{x}$ and all other solutions $\boldsymbol{y}\in F_j$. If a solution $\boldsymbol{y}\in F_j$ is dominated by $\boldsymbol{x}$, the counter (rank) of $\boldsymbol{y}$ is incremented.

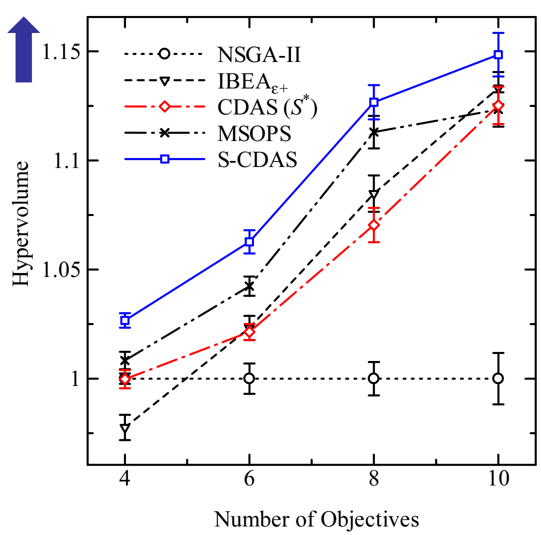
Front Distribution in Objective Space



Conventional Dominance **Expanded Dominance (CDAS with *S*=0.4)** **Expanded Dominance (S-CDAS)**

Performance Comparison with Conventional MOEAs

Results on knapsack problems with 500 items



Relaxed Dominance Based Approach
Counting Objective Approach

(1−k)-dominance [Farina 02]

Solution $\boldsymbol{x}$ is said to (1-k)-dominate solution $\boldsymbol{y}$ if and only if

$$\begin{cases} n_e < m, \\ n_b \geq \dfrac{m-n_e}{k+1}. \end{cases}$$

m: Number of objectives

n_b: Number of objectives where $\boldsymbol{x}$ is better than $\boldsymbol{y}$ ($|\{i\in\{1,2,\ldots,m\}\,|\, f_i(\boldsymbol{x})>f_i(\boldsymbol{y})\}|$)

n_e: Number of objectives where $\boldsymbol{x}$ is equal to $\boldsymbol{y}$ ($|\{i\in\{1,2,\ldots,m\}\,|\, f_i(\boldsymbol{x})=f_i(\boldsymbol{y})\}|$)

- k=0 is equivalent to the conventional dominance.
- The larger k, the more relaxed dominance producing a fine-grained dominance ranking.

M. Farina and P. Amato, "On the optimal solution definition for many-criteria optimization problems," In Proc. of the 2002 Annual Meeting of the North American Fuzzy Information Processing Society (NAFIPS' 02). IEEE, pp. 233-238, 2002.

Relaxed Dominance Based Approach
Counting Objective Approach

Ranking-dominance [Kukkonen 07]

Aggregates rank value on each objective.

$$R_{sum} = \sum_{i=1}^{m} Rank(fi(\boldsymbol{x})), \qquad R_{min} = \min_{i\in\{1,2,..,m\}} Rank(fi(\boldsymbol{x}))$$

	Objective values f_1	f_2	f_3	Ranks f_1	f_2	f_3	Aggregated Ranks R_{sum}	R_{min}
$\boldsymbol{x}_1$	5	200	15	3	2	5	10	2
$\boldsymbol{x}_2$	9	350	14	5	4	4	13	4
$\boldsymbol{x}_3$	4	100	13	2	1	3	6	1
$\boldsymbol{x}_4$	6	270	12	4	3	2	9	2
$\boldsymbol{x}_5$	3	400	11	1	5	1	7	1

Aggregation

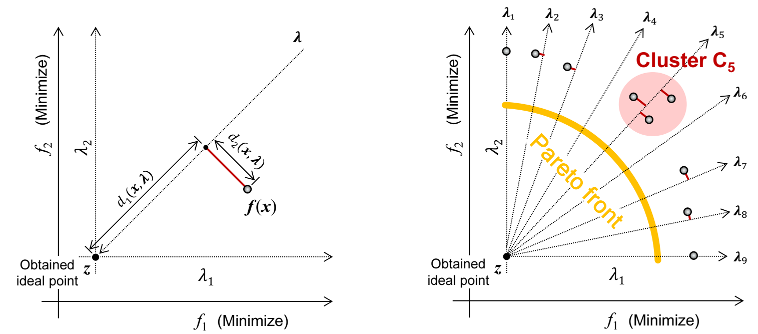
S. Kukkonen and J. Lampinen, "Ranking-dominance and many-objective optimization," In Proc. of the 2007 IEEE Congress on Evolutionary Computation (CEC'07). IEEE, Los Alamitos, CA, pp. 3983-3990, 2007.

Relaxed Dominance Based Approach
Combination with Reference Lines

θ-dominance [Yuan 16] in θ-DEA

– Given two solutions $\boldsymbol{x}$ and $\boldsymbol{y}$, $\boldsymbol{x}$ is said to θ-dominate $\boldsymbol{y}$, if and only if $\boldsymbol{x}$ and $\boldsymbol{y}$ are in the same cluster C_j and $F(\boldsymbol{x},\boldsymbol{\lambda_j}) < F(\boldsymbol{y},\boldsymbol{\lambda_j})$.

$$F(\boldsymbol{x},\boldsymbol{\lambda_j}) = d_1(\boldsymbol{x},\boldsymbol{\lambda_j}) + \theta d_2(\boldsymbol{x},\boldsymbol{\lambda_j})$$



Y. Yuan, H. Xu, B. Wang, and X. Yao., "A New Dominance Relation-Based Evolutionary Algorithm for Many-Objective Optimization," IEEE Transactions on Evolutionary Computation, 20, 1, pp. 16–37, 2016.

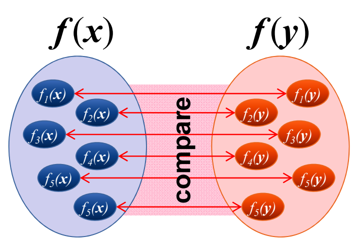
Relaxed Dominance Based Approach
Pareto Partial Dominance

Pareto Dominance

Solution $\boldsymbol{x}$ dominates $\boldsymbol{y}$, if

$$\forall i \in M : f_i(\boldsymbol{x}) \geq f_i(\boldsymbol{y}) \quad \wedge$$
$$\exists i \in M : f_i(\boldsymbol{x}) > f_i(\boldsymbol{y})$$

is satisfied. $\quad M = \{1,2,\ldots,m\}$



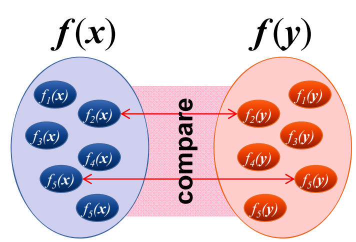
Pareto Partial Dominance

[Sato 12]

Solution $\boldsymbol{x}$ partially dominate $\boldsymbol{y}$, if

$$\forall i \in R \subset M : f_i(\boldsymbol{x}) \geq f_i(\boldsymbol{y}) \quad \wedge$$
$$\exists i \in R \subset M : f_i(\boldsymbol{x}) > f_i(\boldsymbol{y})$$

is satisfied. $\quad M = \{1,2,\ldots,m\}$

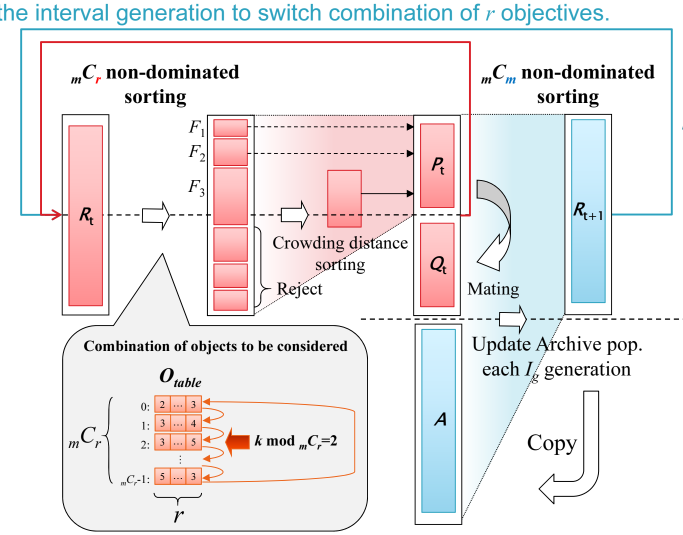


There are ${}_mC_r$ combinations when we select r objectives from m objective functions.

Pareto Partial Dominance MOEA

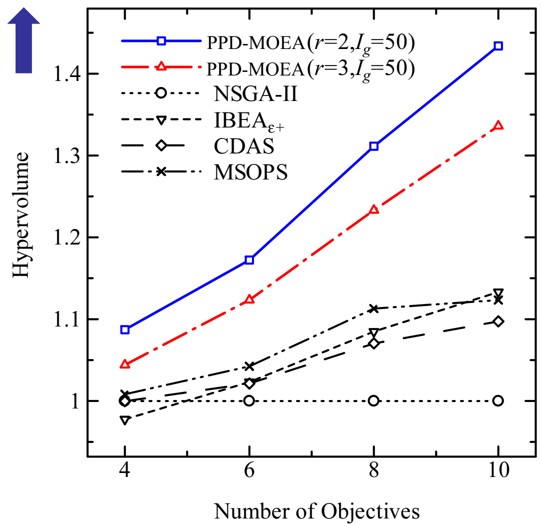
Parameters: [Sato 12]

- r : the number of objectives to be considered in partial dominance.
- I_g: the interval generation to switch combination of r objectives.



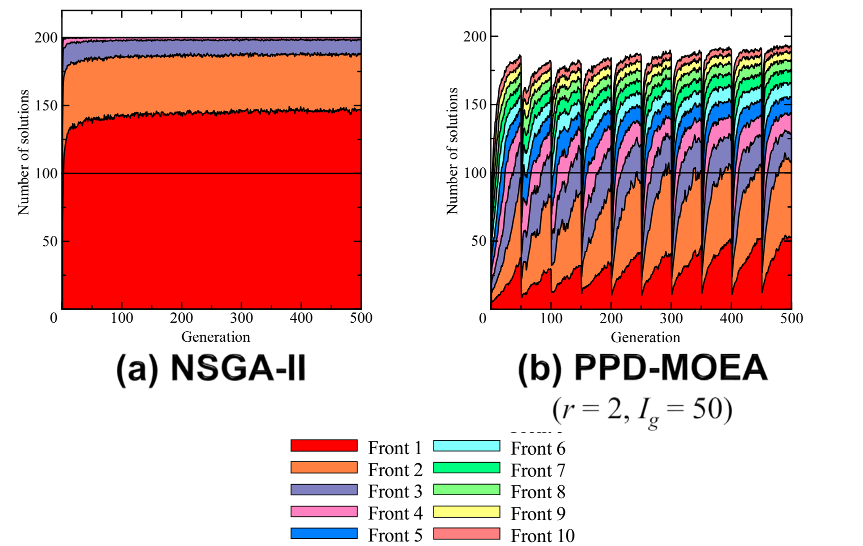
Performance Varying the Number of Objectives

Results on knapsack problems with 500 items



Front Distribution Over Generations (8 objectives)

Results on knapsack problems with 500 items



Approaches for Many-Objective Optimization

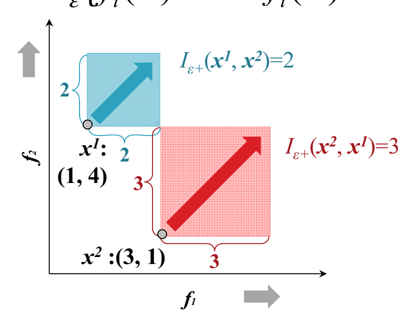
1. **Relaxed Dominance Based Approach**
2. **Indicator Based Approach**
3. **Decomposition Based Approach**
4. **Reference Based Approach**
5. **Dimensionality Reduction Approach**
6. **Efficient Solution Generation**

Indicator Based Approach
Distance Based Indicator

IBEA$_{\varepsilon+}$ [Zitzler 04]

introduces fine grained ranking of solutions by calculating fitness value based on indicators which measure the degree of superiority for each solution in the population.

$$I_{\varepsilon+}(A,B) = \min_{\varepsilon}\{f_i(A) + \varepsilon \geq f_i(B) : i \in 1,\ldots,m\}$$



E. Zitzler and S. Künzli, "Indicator-Based Selection in Multiobjective Search," in Proc. of Parallel Problem Solving from Nature - PPSN VIII. PPSN 2004. LNCS, Vol. 3242, pp. 832–842, 2004.

Indicator Based Approach
Hypervolume

SMS-EMOA [Beume 07]

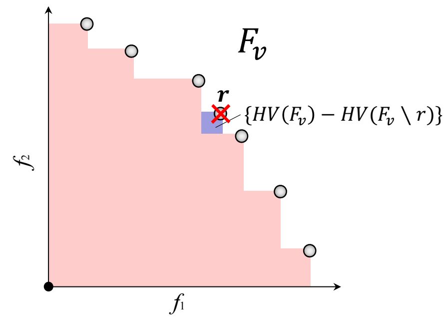
– (μ+1) algorithm using Hypervolume contribution

SMS-EMOA

$t \leftarrow 0$
$P_t \leftarrow \text{init}()$
Repeat
  $q_{t+1} \leftarrow \text{Generate}(P_t)$
  $P_{t+1} \leftarrow \text{Reduce}(P_t \cup q_{t+1})$
  $t \leftarrow t+1$
Until termination condition fulfilled

Reduce (Q)

$\{F_1, F_2, \ldots, F_v\} \leftarrow$ Non-dominated sort (Q)
$r \leftarrow \text{argmin}_{s \in F_v}\{HV(F_v) - HV(F_v \setminus s)\}$
return $(Q \setminus r)$



N. Beume, B. Naujoks, M. Emmerich, "SMS-EMOA: Multiobjective selection based on dominated hypervolume," European Journal of Operational Research, Vol. 181, Issue 3, pp. 1653-1669, 2007.

Indicator Based Approach
Hypervolume

SMS-EMOA [Beume 07]

– The computational cost to calculate Hypervolume is exponentially increased by increasing the number of objectives.

HypE (Hypervlume Estimation) [Bader 11]

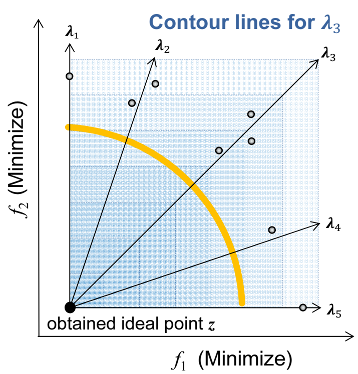
– Monte Carlo sampling to approximate hypervolume values.

J. Bader and E. Zitzler, "HypE: An algorithm for fast hypervolume-based many-objective optimization," Evolutionary Computation, Vol. 19, Issue 1, pp. 45–76, 2011.

Indicator Based Approach
R2 Indicator

MOMBI [Gómez 13]

– R2 indicator to evaluate a set of solutions is utilized to rank solutions in the population.



R2 value for a solution set $\mathcal{P}$

$$R2(\mathcal{P}, \mathcal{L}) = \frac{1}{|\mathcal{L}|} \sum_{\boldsymbol{\lambda} \in \mathcal{L}} \min_{\boldsymbol{x} \in \mathcal{P}} \left\{ \max_{i \in \{1,2,\cdots,m\}} \lambda_i \cdot f_i'(\boldsymbol{x}) \right\}$$

$\mathcal{L}$: Weight vector set $\{\boldsymbol{\lambda}_1, \boldsymbol{\lambda}_2, \ldots\}$

Rank i solution set $\mathcal{F}_i$

$$\mathcal{F}_i = \bigcup_{\boldsymbol{\lambda} \in \mathcal{L}} \underset{\boldsymbol{x} \in \mathcal{P} \setminus \mathcal{Q}}{\text{argmin}} \left\{ \max_{i \in \{1,2,\cdots,m\}} \lambda_i \cdot f_i'(\boldsymbol{x}) \right\}$$

$\mathcal{Q}$: Lower ranked solutions than $\mathcal{F}_i$

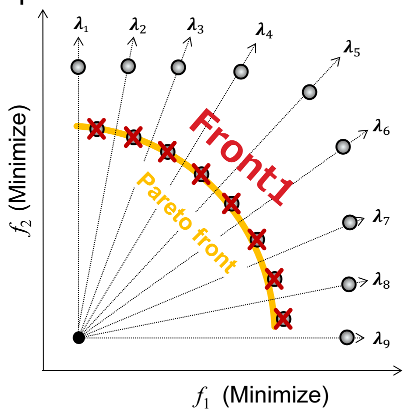
R. H. Gómez and C. A. C. Coello, "MOMBI: A new metaheuristic for many-objective optimization based on the R2 indicator," in Proc. of IEEE Congress on Evolutionary Computation, pp. 2488-2495, 2013.

Approaches for Many-Objective Optimization

1. **Relaxed Dominance Based Approach**
2. **Indicator Based Approach**
3. **Decomposition Based Approach**
4. **Reference Based Approach**
5. **Dimensionality Reduction Approach**
6. **Efficient Solution Generation**

Decomposition Based Approach

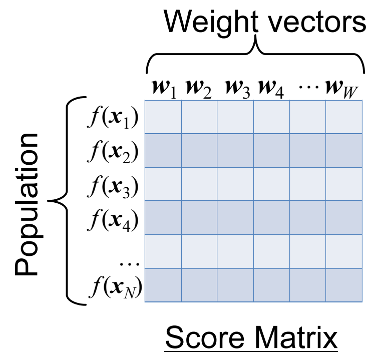
- The Pareto front (objective space) is decomposed by a set of weight vectors or reference lines.
- Each weight vector or reference line specifies an approximation part of the Pareto front.



Decomposition Based Approach
Weight Based Ranking

MSOPS [Hughes 03]

aggregates fitness vector with multiple weight vectors, and reflects the ranking of solutions calculated for each weight vector in parent selection.



(1) MSOPS fills up all the elements of the score matrix by calculating

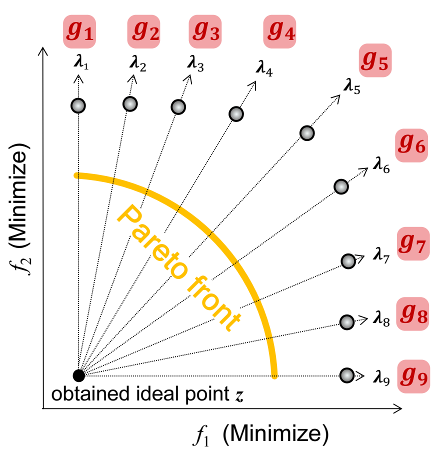
$$s = \max_{i \in \{1,2,\ldots,m\}} w_i \cdot f_i(\boldsymbol{x}).$$

(2) For each column, all solutions are ranked in descending order based on score values.
(3) For each row, the indices of column are sorted in ascending order based on rank value.
(4) MSOPS selects solutions which have higher ranks for each weight vectors as parent solutions.

E. J. Hughes, "Multiple single objective Pareto sampling," In Proc. of the 2003 IEEE Congress on Evolutionary Computation (CEC'03), Vol. 4, pp. 2678–2684, 2003.

MOEA/D [Zhang 07]

- MOEA/D decomposes a multi-objective optimization problem into a number of single-objective optimization problems.
- Each single objective optimization problem is defined by a **scalarizing function g** using a weight vector λ_i $(i = 1,2,\ldots,N)$.



After the Proposal of MOEA/D…

Variants of MOEA/D focusing on

1. Solution Generation
- Continuous: DE and PSO [Li 09, Liuet 10, Martinez 11, Moubayed 10]
- Discrete: ACO and SA [Ke 13, Li 11]

2. Weight Vectors
- Adaptive control of the weight distribution [Jiang 11, Hamada 13,14]
- Self-adaptation of neighborhood size [Zhao 12]

3. Combination with Dominance
- NSGA-III [Deb 14], MOEA/DD [Li 15], θ-DEA [Yuan 16], etc

4. Parent Selection
- Control the number of solution generations for each scalarizing function [Zhang 09, Chiang 11]

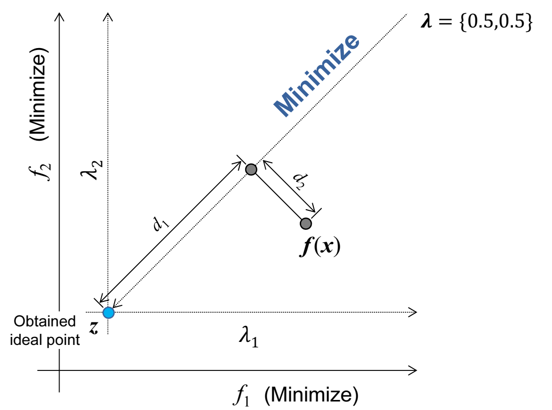
5. Scalarizing function
- Original: Tchebycheff [Bowman Jr. 76], Weighted Sum [Gass 55], PBI
- NBI-style Tchebycheff [Zhang 10]
- Adaptive Selection of Scalarizing function [Ishibuchi 09]

PBI Scalarizing Function

Minimize g^{pbi} toward **the obtained ideal point** z

Minimize $g^{pbi}(\boldsymbol{x}|\boldsymbol{\lambda}) = d_1 + \theta d_2$ θ: Parameter

$$d_1 = \frac{||(\boldsymbol{f}(\boldsymbol{x})-\boldsymbol{z})^T\boldsymbol{\lambda}||}{||\boldsymbol{\lambda}||}, \quad d_2 = ||\boldsymbol{f}(\boldsymbol{x}) - \left(\boldsymbol{z} - d_1\frac{\boldsymbol{\lambda}}{||\boldsymbol{\lambda}||}\right)||$$

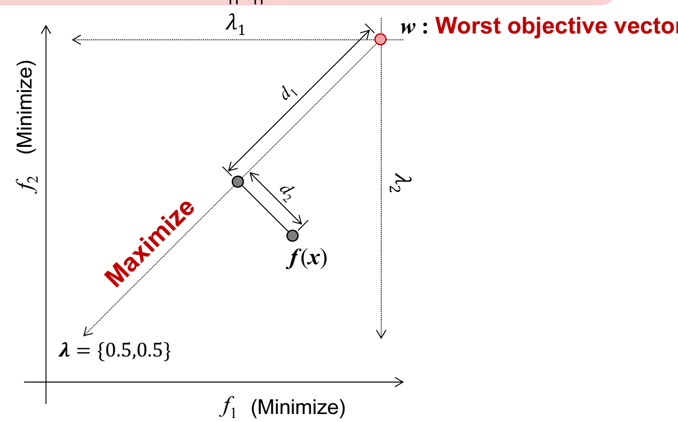


Inverted PBI Scalarizing Function

Maximize g^{ipbi} from **the worst objective vector** w

Maximize $g^{ipbi}(\boldsymbol{x}|\boldsymbol{\lambda}) = d_1 - \theta d_2$ θ: Parameter

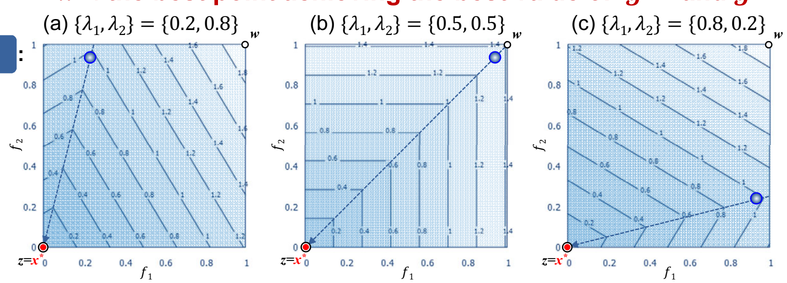
$$d_1 = \frac{||(\boldsymbol{w}-\boldsymbol{f}(\boldsymbol{x}))^T\boldsymbol{\lambda}||}{||\boldsymbol{\lambda}||}, \quad d_2 = ||(\boldsymbol{w} - \boldsymbol{f}(\boldsymbol{x})) - d_1\frac{\boldsymbol{\lambda}}{||\boldsymbol{\lambda}||}||$$



Contour Lines of PBI and Inverted PBI

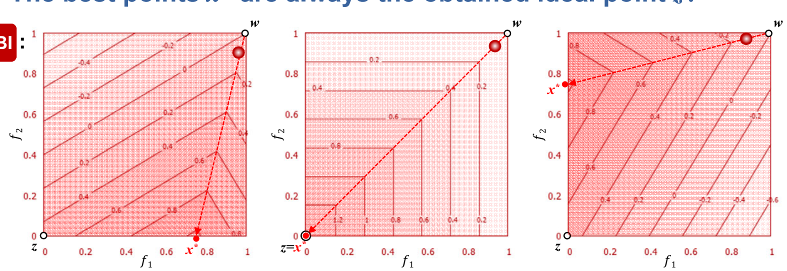
x^* : the best point achieving the best value of g^{pbi} and g^{ipbi}

PBI: using $\theta = 1.0$



The best points x^* are always the obtained ideal point z.

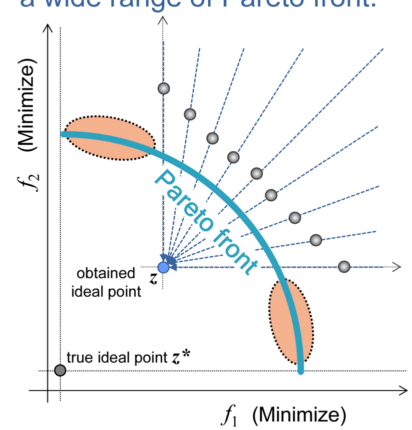
Inverted PBI: using $\theta = 1.0$



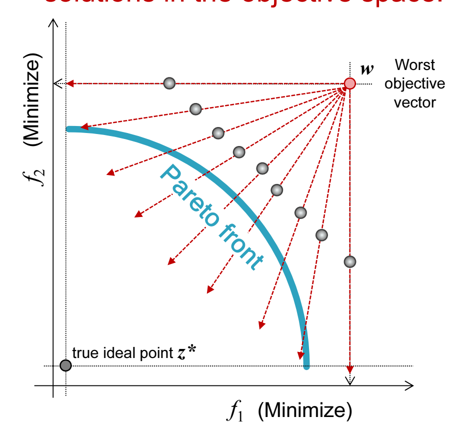
The best points x^* are widely distributed in the objective space.

Expected Improvement of Spread

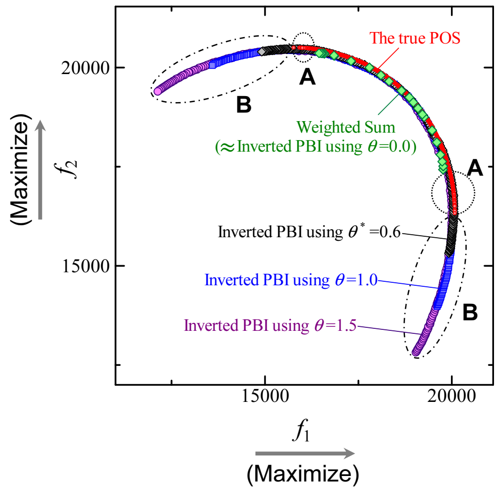
PBI faces the difficulty to approximate a wide range of Pareto front.



Inverted PBI can improve the spread of solutions in the objective space.



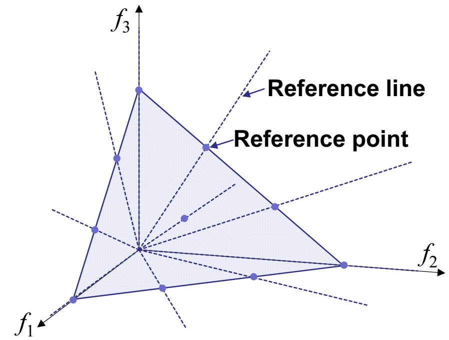
Obtained Population in MOKP with 2 objectives



NSGA using Reference Lines

NSGA-III [Deb 14]

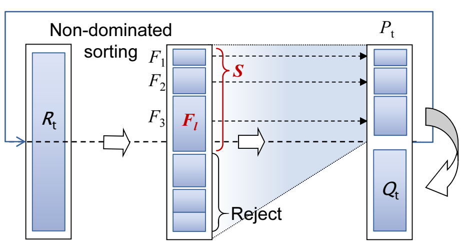
– Is designed for solving many-objective problems.
– Introduces reference lines to maintain the distribution of solutions in the objective space.

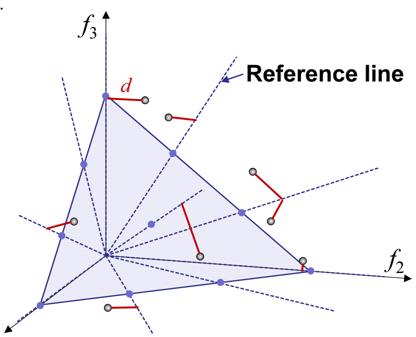


K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints," IEEE Transactions on Evolutionary Computation, 18, 4, pp. 577-601, 2014.

NSGA using Reference Lines

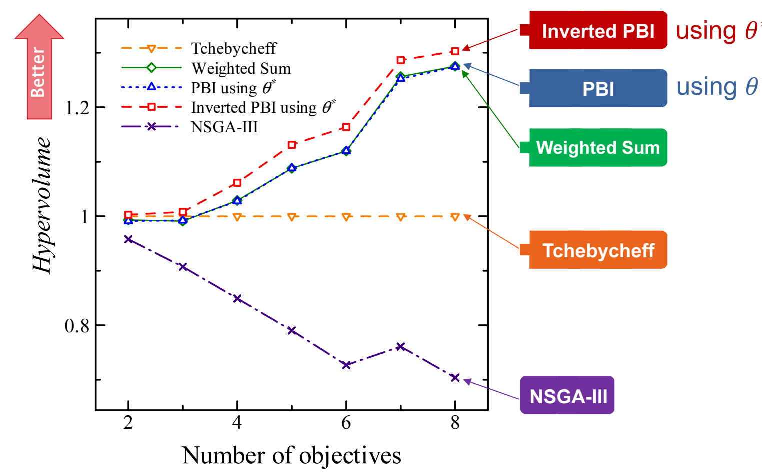
– Each solution in $\boldsymbol{S}$ is associated with a reference line having the minimum perpendicular distance d.
– The reference line with the minimum number of associated solution in $S \backslash F_l$ is focused.
  - If the reference line has no associated solution in $S \backslash F_l$ and F_l has solution(s) associated with the focused reference line, select the solution with the minimum perpendicular distance d from them.
  - If the reference line has associated solution(s) in $S \backslash F_l$ and F_l has solution(s) associated with the focused reference line, randomly select a solution from them.





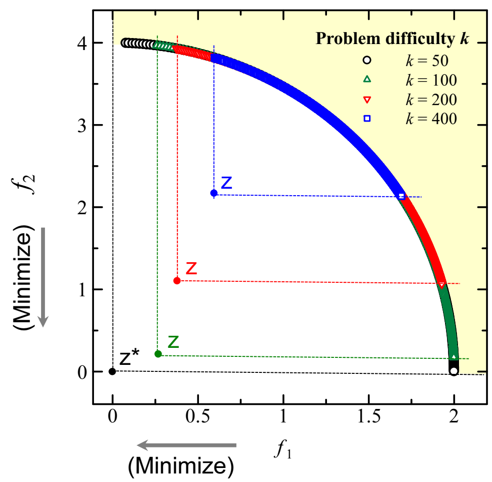
K. Deb and H. Jain, "An evolutionary many-objective optimization algorithm using reference-point-based nondominated sorting approach, part I: Solving problems with box constraints," IEEE Transactions on Evolutionary Computation, 18, 4, pp. 577-601, 2014.

Hypervolume on MOKPs

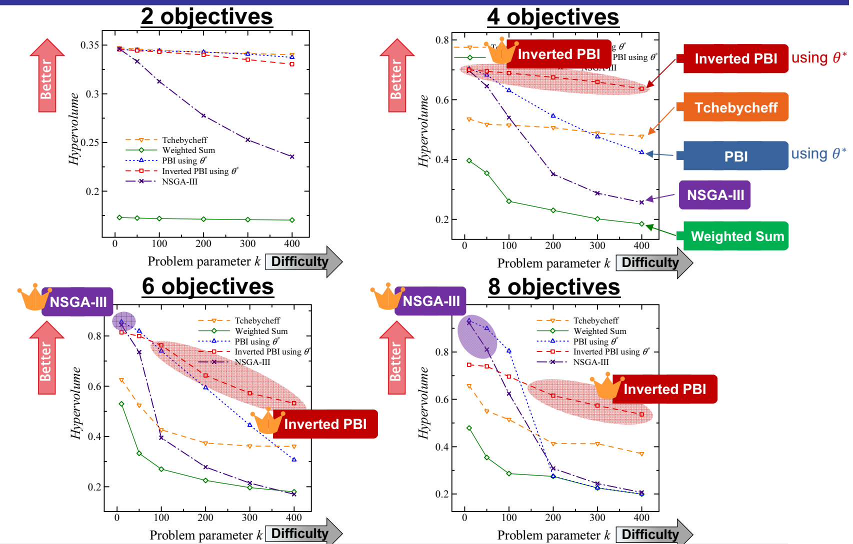


Problem Difficulty k in WFG4

Solutions on four WFG4 problems with different problem difficulties k.
The difficulty to obtain a widely spread solutions is increased by increasing k.



Hypervolume on WFG4 Problems

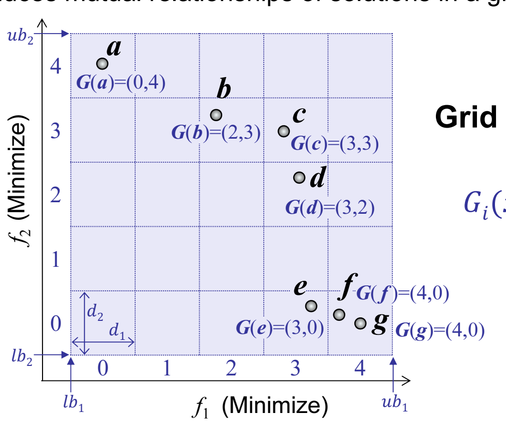


The inverted PBI achieves high search performance in problems with
- **many-objectives** and
- **the difficulty to obtain a widely spread solutions** in the objective space.

Decomposition Based Approach
Grid-based Decomposition

GrEA [Yang 13]

Introduces mutual relationships of solutions in a grid environment.



Grid coordinate

$$G_i(\boldsymbol{x}) = \left\lfloor \frac{f_i(x) - lbi}{d_i} \right\rfloor$$

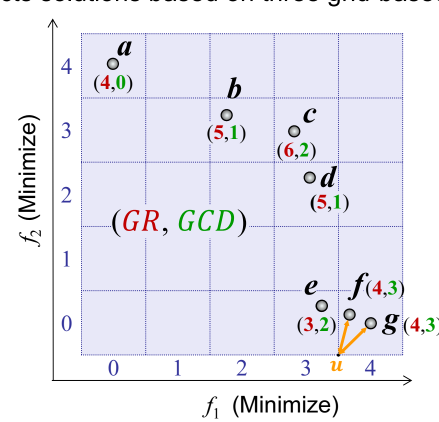
$$(i = 1,2, \dots, m)$$

S. Yang, M. Li, X. Liu and J. Zheng, "A Grid-Based Evolutionary Algorithm for Many-Objective Optimization," IEEE Transactions on Evolutionary Computation, Vol. 17, No. 5, pp. 721-736, 2013.

Decomposition Based Approach
Grid-based Decomposition

GrEA [Yang 13]

Selects solutions based on three grid-based criteria.



1st: Convergence

$$GR(\boldsymbol{x}) = \sum_{i=1}^{m} G_i(\boldsymbol{x})$$

2nd: Diversity

$$GCD(\boldsymbol{x}) = \sum_{y \in N(x)} \left(m - \sum_{i=1}^{m} |G_i(x) - Gi(y)| \right)$$

3rd:

$$GCPD(\boldsymbol{x}) = \sqrt{\sum_{i=1}^{m} (f_i(x) - ui)^2}$$

$\boldsymbol{u}$: the utopia point in its hyper-box

S. Yang, M. Li, X. Liu and J. Zheng, "A Grid-Based Evolutionary Algorithm for Many-Objective Optimization," IEEE Transactions on Evolutionary Computation, Vol. 17, No. 5, pp. 721-736, 2013.

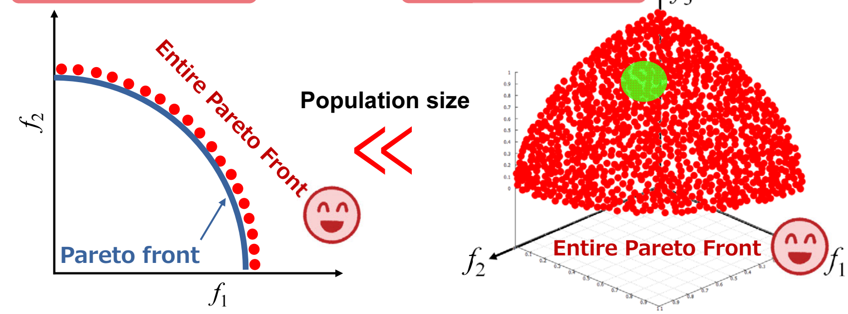
Approaches for Many-Objective Optimization

1. Relaxed Dominance Based Approach
2. Indicator Based Approach
3. Decomposition Based Approach
4. **Reference Based Approach**
5. Dimensionality Reduction Approach
6. Efficient Solution Generation

Difficulties in Approximation of Entire Pareto Front

- Any dimensional Pareto front must be approximated by a set of points.

2 objectives 3 objectives



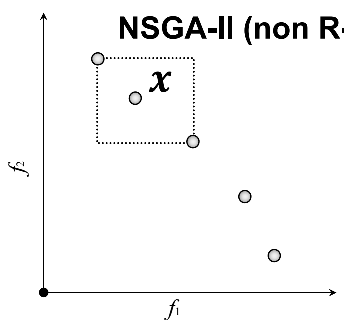
To finely approximate high-dimensional Pareto front, a huge size of population is needed, and computational cost is increased.

- **Handling of Huge Population**
  [Kowatari 12, Ishibuchi 15, Jaimes 16, Tatsukawa 16]
- **User-preference based search**
  [Deb 06, Auger 09, Gong 11, Narukawa 15, etc.]

Reference Based Approach
Goal Point Approach

R-NSGA-II [Deb 06]

– The diversity maintenance mechanism using crowding distance is replaced with a search focusing mechanism using distance to the **reference point**.

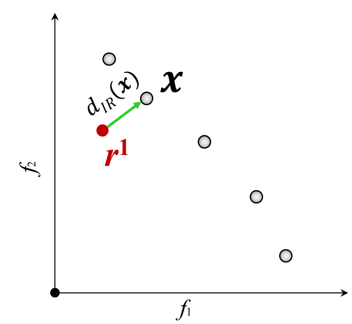


K. Deb, J. Sundar, N. Udaya Bhaskara Rao, and S. Chaudhuri, "Reference point based multi-objective optimization using evolutionary algorithms," International Journal of Computational Intelligence Research, Vol. 2, Issue 3, pp. 273–286, 2006.

Reference Based Approach
Goal Point Approach

R-NSGA-II [Deb 06]

– The diversity maintenance mechanism using crowding distance is replaced with a search focusing mechanism using distance to the **reference point**.



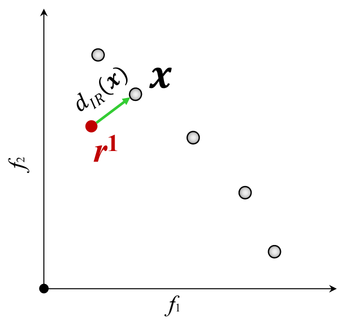
$$d_{IR}(\boldsymbol{x}) = \sqrt{\sum_{i=1}^{m}\left(\frac{f_i(\boldsymbol{x}) - ri}{f_i^{max}(\boldsymbol{x}) - f_i^{max}(\boldsymbol{x})}\right)^2}$$

K. Deb, J. Sundar, N. Udaya Bhaskara Rao, and S. Chaudhuri, "Reference point based multi-objective optimization using evolutionary algorithms," International Journal of Computational Intelligence Research, Vol. 2, Issue 3, pp. 273–286, 2006.

Reference Based Approach
Goal Point Approach

R-NSGA-II [Deb 06]

- Solutions with a smaller d_{IR} are preferred.
- For randomly choose two solutions, if they are in the same front, one has smaller d_{IR} wins.



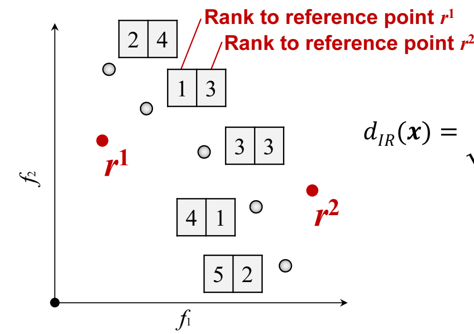
$$d_{IR}(\boldsymbol{x}) = \sqrt{\sum_{i=1}^{m}\left(\frac{f_i(\boldsymbol{x}) - ri}{f_i^{max}(\boldsymbol{x}) - f_i^{max}(\boldsymbol{x})}\right)^2}$$

K. Deb, J. Sundar, N. Udaya Bhaskara Rao, and S. Chaudhuri, "Reference point based multi-objective optimization using evolutionary algorithms," International Journal of Computational Intelligence Research, Vol. 2, Issue 3, pp. 273–286, 2006.

Reference Based Approach
Goal Points Approach

R-NSGA-II [Deb 06]

- R-NSGA-II can treat multiple reference points.
- For each reference, solutions are ranked based on d_{IR}.



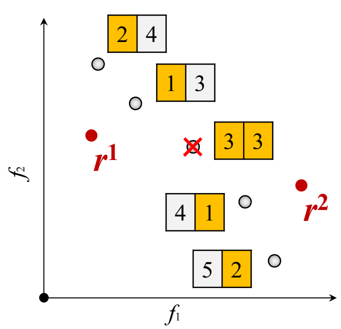
$$d_{IR}(\boldsymbol{x}) = \sqrt{\sum_{i=1}^{m}\left(\frac{f_i(\boldsymbol{x}) - ri}{f_i^{max}(\boldsymbol{x}) - f_i^{max}(\boldsymbol{x})}\right)^2}$$

K. Deb, J. Sundar, N. Udaya Bhaskara Rao, and S. Chaudhuri, "Reference point based multi-objective optimization using evolutionary algorithms," International Journal of Computational Intelligence Research, Vol. 2, Issue 3, pp. 273–286, 2006.

Reference Based Approach
Goal Points Approach

R-NSGA-II [Deb 06]

- R-NSGA-II can treat multiple reference points.
- For each reference, solutions are ranked based on d_{IR}.
- **Minimum rank** of each solution is used for comparison.



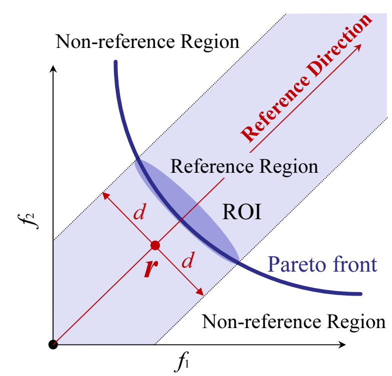
$$d_{IR}(\boldsymbol{x}) = \sqrt{\sum_{i=1}^{m}\left(\frac{f_i(\boldsymbol{x}) - ri}{f_i^{max}(\boldsymbol{x}) - f_i^{max}(\boldsymbol{x})}\right)^2}$$

K. Deb, J. Sundar, N. Udaya Bhaskara Rao, and S. Chaudhuri, "Reference point based multi-objective optimization using evolutionary algorithms," International Journal of Computational Intelligence Research, Vol. 2, Issue 3, pp. 273–286, 2006.

Reference Based Approach
Preference Region Approach

Use of Preference Radius [Hu 17]

- Preference Region is specified by the reference point, direction, and **radius** d.



Algorithm

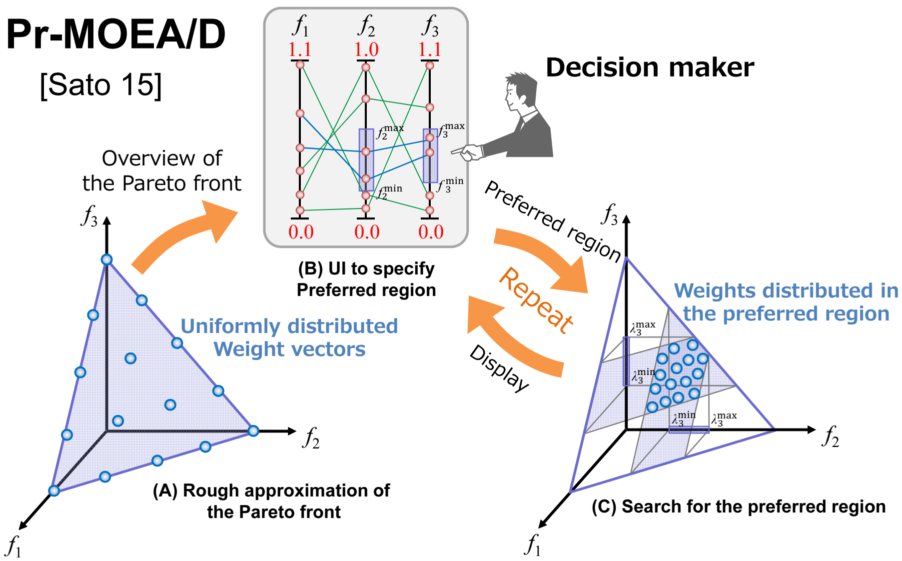
- Solutions are divided into
  1. Non-preferred solution set
  2. Preferred solution set
- Among preferred solutions, solutions are selected by using dominance.
- If the number of selected solutions is lower than the upper limit, non-preferred solutions which have smaller distances to the reference directions are selected.

J. Hu, G. Yu, J. Zheng, and J. Zou, "A preference-based multi-objective evolutionary algorithm using preference selection radius," Soft Computing, Vol. 21, Issue 17, pp. 5025-5051, 2017.

Reference Based Approach
Preference Region Approach

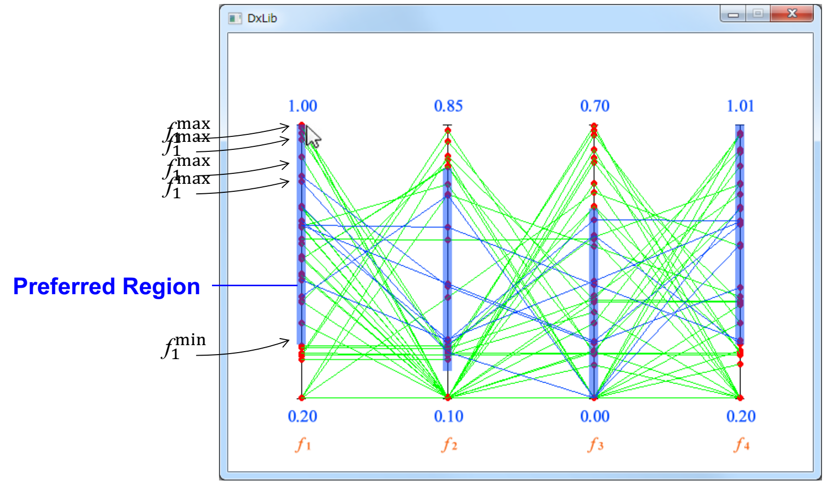
Pr-MOEA/D

[Sato 15]



(B) UI to Specify Preferred Region

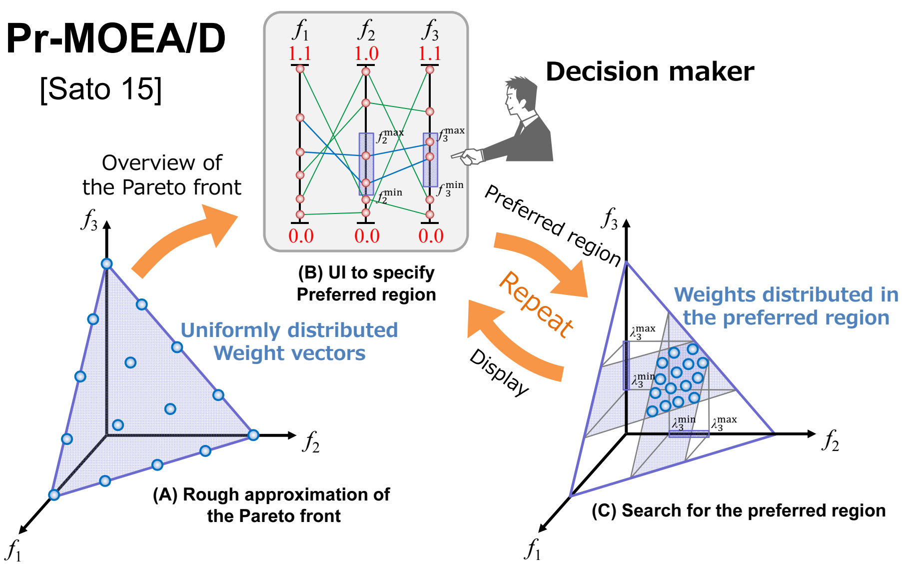
Parallel Coordinates UI



Reference Based Approach
Preference Region Approach

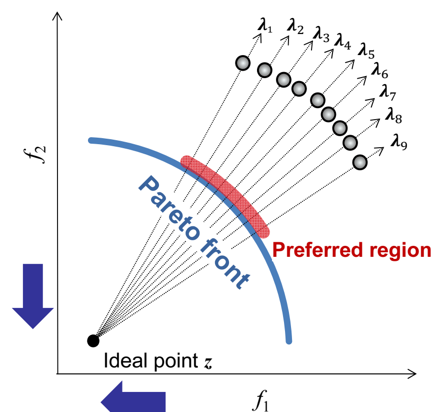
Pr-MOEA/D

[Sato 15]



(C) Fine-grained Approximation of Preferred region

Pr-MOEA/D [Sato 15]



- Objective vector $\boldsymbol{f}$ is transformed into a scalar value g with a weight $\boldsymbol{\lambda}$

Scalarizing function :

Minimize $g(\boldsymbol{x}|\boldsymbol{\lambda}^j) = d_1 + \theta d_2$

$$d_1 = \frac{\left\|((\boldsymbol{f}(\boldsymbol{x}) - z)^T \boldsymbol{\lambda}^j)\right\|}{\|\boldsymbol{\lambda}^j\|},$$

$$d_2 = \left\|\boldsymbol{f}(\boldsymbol{x}) - (z - d_1 \boldsymbol{\lambda}^j)\right\|$$

$(j = 1,2,\ldots,N)$

N : Population size

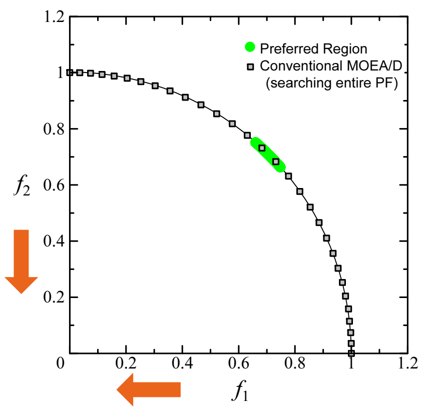
m : Number of objectives

Preference Region Based Search:

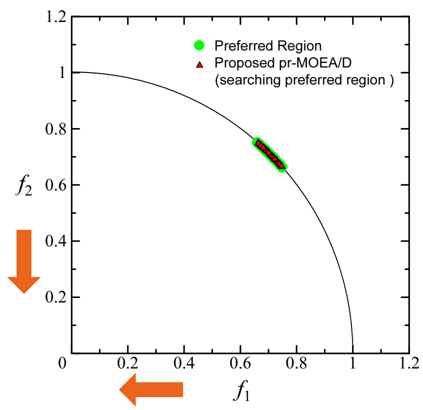
Weight distribution is changed and a part of Pareto front is approximated.

Obtained Solutions (2 objectives)

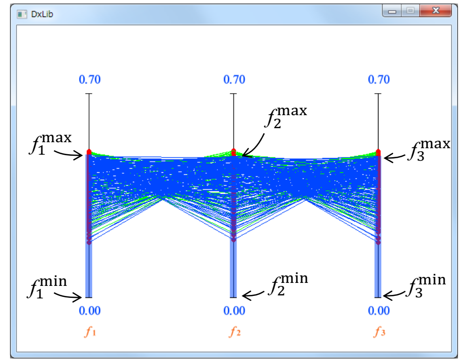
Conventional MOEA/D

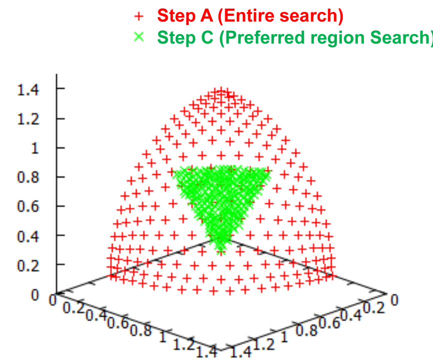


Pr-MOEA/D



Obtained Solutions (3 objectives)





Approaches for Many-Objective Optimization

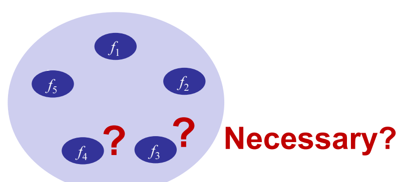
1. **Relaxed Dominance Based Approach**
2. **Indicator Based Approach**
3. **Decomposition Based Approach**
4. **Reference Based Approach**
5. **Dimensionality Reduction Approach**
6. **Efficient Solution Generation**

Dimensionality Reduction

- **[Ishibuchi 11] showed that search performances of NSGA-II and SPEA2 were not severely degraded by the increase in the number of objectives when they were highly correlated or dependent.**

H. Ishibuchi, N. Akedo, H. Ohyanagi and Y. Nojima, "Behavior of EMO algorithms on many-objective optimization problems with correlated objectives," in Proc. of 2011 IEEE Congress of Evolutionary Computation (CEC), pp. 1465-1472, 2011.

- **Dimensionality reduction of objectives is a way to reduce the difficulty of many-objective optimization.**
  - **[Saxena 05-10], [Brockhoff 06-11], [Jaimes 08-09], [Guo 12, 13]…**



Dimensionality Reduction

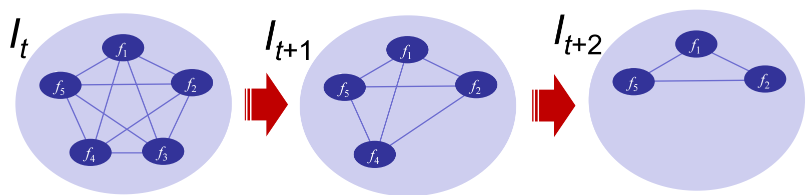
PCA-NSGA-II [Deb 05]

Step 1: Set an iteration counter $t = 0$ and initial set of objectives $I_0 = \{f_1, f_2, \ldots, f_m\}$.

Step 2: Initialize a random population for all objectives in the set I_t, run EMO, and obtain a population P_t.

Step 3: Perform a PCA analysis on P_t using I_t to choose a reduced set of objectives I_{t+1}.

Step 4: **If $I_{t+1} = I_t$, stop and declare the obtained front.** Else set $t = t + 1$ and go to **Step 2**.



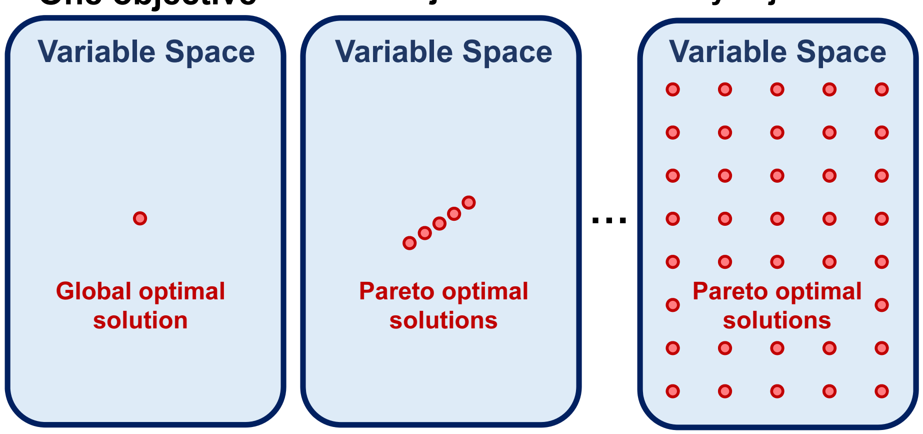
Approaches for Many-Objective Optimization

1. **Relaxed Dominance Based Approach**
2. **Indicator Based Approach**
3. **Decomposition Based Approach**
4. **Reference Based Approach**
5. **Dimensionality Reduction Approach**
6. **Efficient Solution Generation**

Efficient Solution Generation

Search range should be wider in many-objective problems.

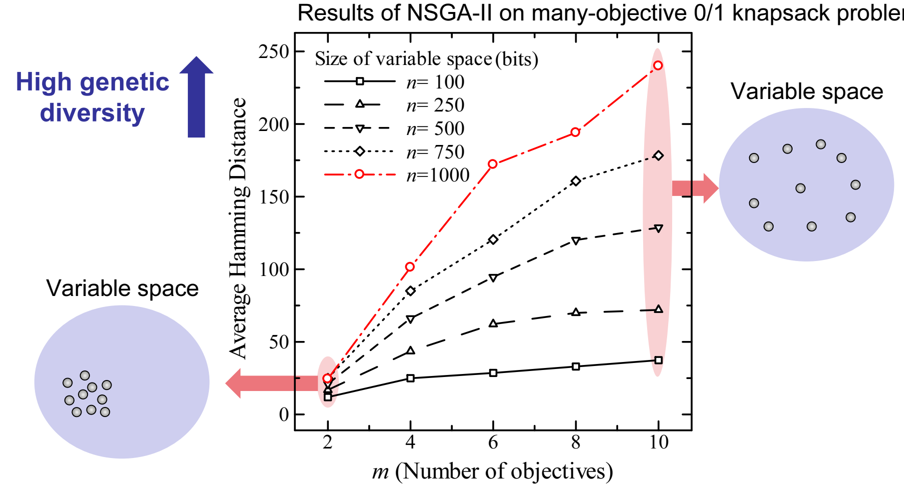
One objective | **Two objectives** | **Many objectives**



Note that it depends on problem characteristics

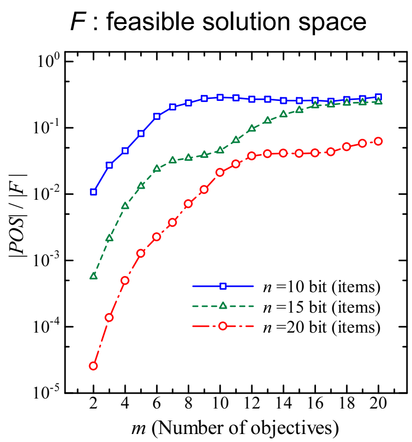
Increasing Genetic Diversity in MaOPs

Results of NSGA-II on many-objective 0/1 knapsack problem

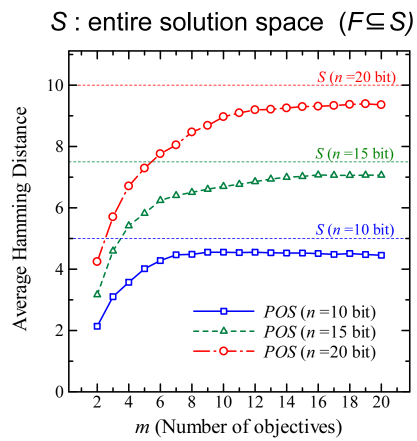


Genetic diversity in the population becomes **noticeably diverse by increasing the number of objectives.**

Pareto Optimal Solutions on Many-objective Knapsack Problems



(a) Ratio of true Pareto optimal solutions POS in feasible solution space *F*



(b) Average hamming distance of true POS

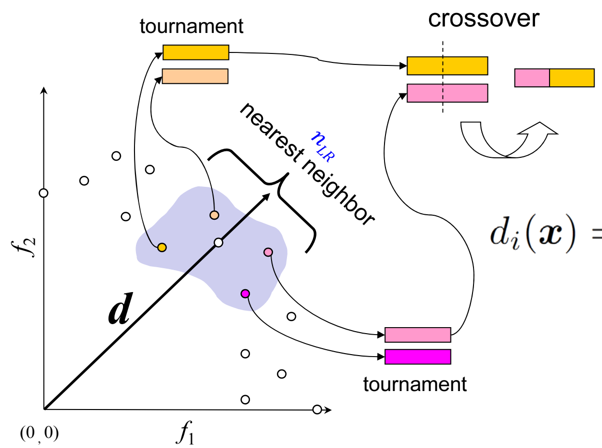
Local Mating in Objective Space

- To improve effects of genetic operation, there are several studies that apply crossover for two parents located near each other in the objective space.
  - NCGA (Neighborhood Cultivation GA) [Watanabe 02]
  - Local Recombination [Sato 07]
  - Mating scheme to control the diversity and convergence [Ishibuchi 04]
  - MOEA/D [Zhang 07]

Local Mating in Objective Space

Local Recombination [Sato 07]

selects pairs of parents by considering nearness of the search direction $\boldsymbol{d}$ of solutions, using a locality parameter n_{LR}.



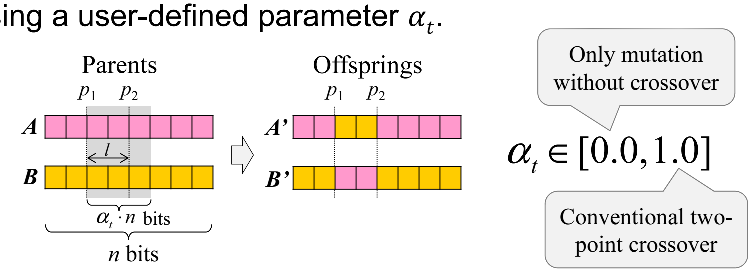
$$d_i(\boldsymbol{x}) = f_i(\boldsymbol{x}) / \sum_{j=1}^{m} f_j(\boldsymbol{x})$$

$$(i = 1, 2, \cdots, m)$$

Control of Crossed Genes in Variable Space

CCG_{TX} : CCG for Two-point Crossover

CCG_{TX} controls the maximum length of crossed genes by using a user-defined parameter α_t.

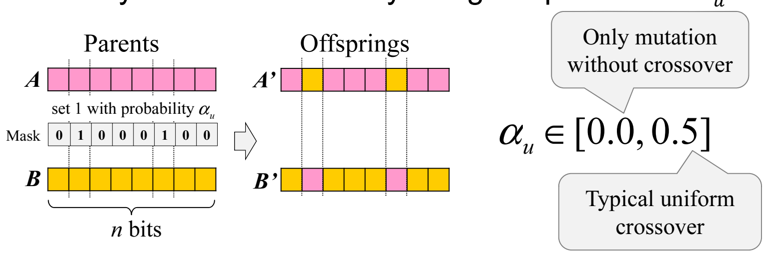


1. Randomly choose the 1st crossover point p_1.
2. Randomly determine the length of the crossed genes l in the range $[0, \alpha_t \cdot n]$.

Control of Crossed Genes in Variable Space

CCG$_{UX}$: CCG for Uniform Crossover

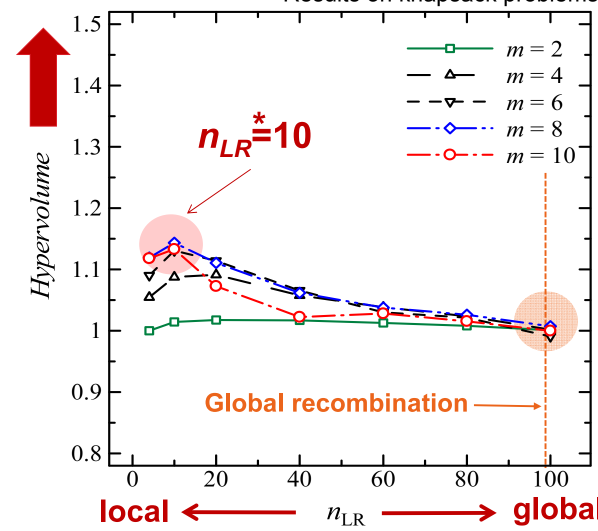
To control the number of crossed genes, CCG$_{UX}$ controls the probability of 1 in the mask by using the parameter α_u.



1. For all mask bit, set 1 with probability α_u.
2. If mask bit is 1, the gene is copied from other parent.

Effects of Local Mating Focusing on Objective Space

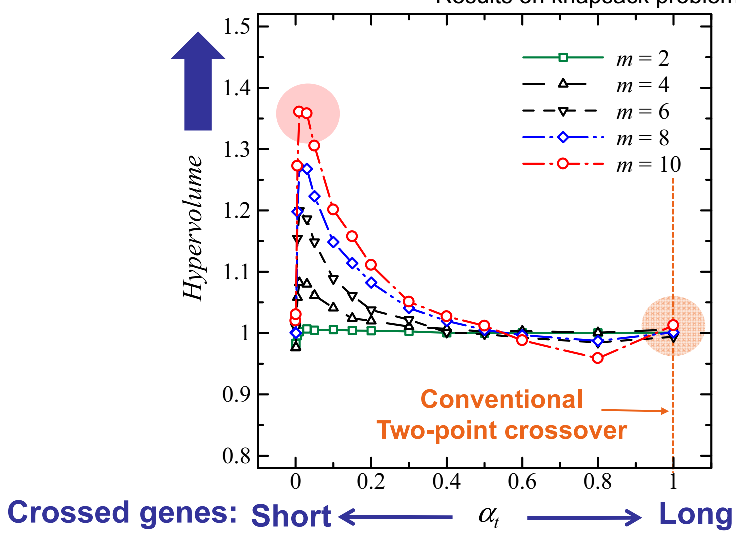
Results on knapsack problems with 500 items



NSGA-II

Effects of CCG$_{TX}$ Focusing on Variable Space

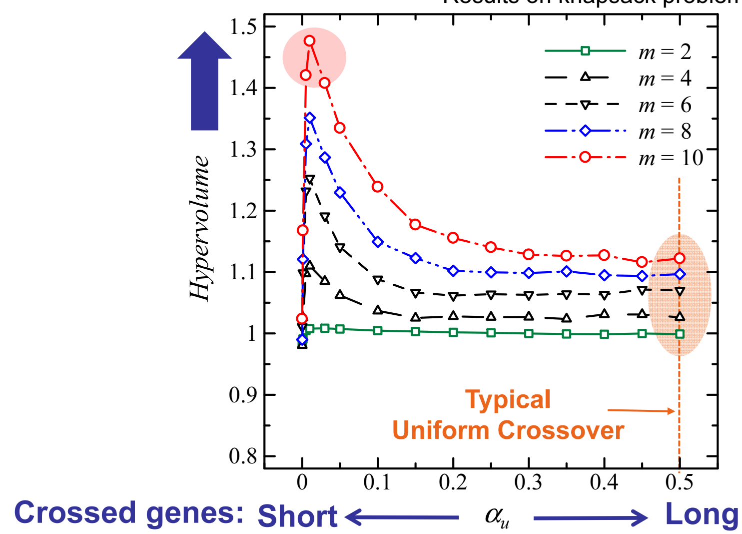
Results on knapsack problems with 500 items



NSGA-II

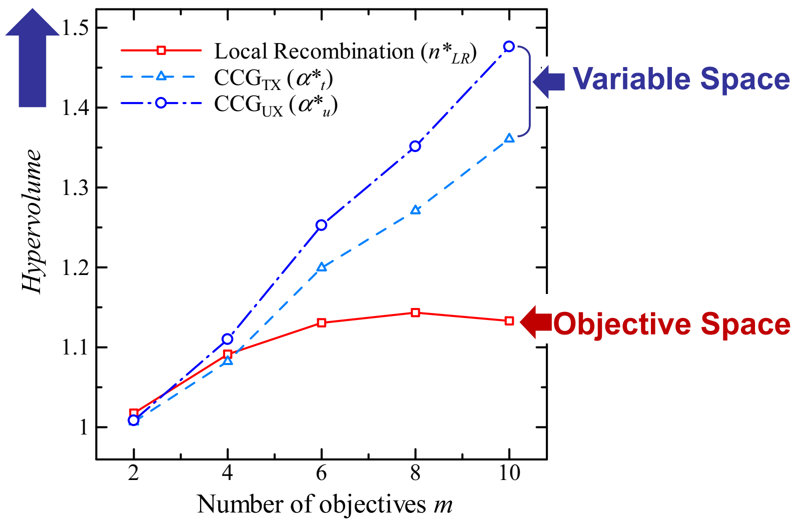
Effects of CCG$_{UX}$ Focusing on Variable Space

Results on knapsack problems with 500 items



NSGA-II

Comparison of Best *HVs* (NSGA-II)



Approaches for Many-Objective Optimization

1. **Relaxed Dominance Based Approach**
2. **Indicator Based Approach**
3. **Decomposition Based Approach**
4. **Reference Based Approach**
5. **Dimensionality Reduction Approach**
6. **Efficient Solution Generation**

Part 3: Future Directions

We explain why many-objective optimization is difficult for EMO (Evolutionary Multiobjective Optimization) algorithms, and how those difficulties can be handled. We also suggest some promising future research directions

Part 1 (Hisao Ishibuchi): Difficulties

Part 2 (Hiroyuki Sato): Approaches

Part 3 (Hisao Ishibuchi): Future Directions

Future Directions

1. Adaptation of Weight/Reference Vectors
A number of approaches have already been proposed. However, we still have a number of issues to be further addressed.

2. New Algorithm Development
From a practical viewpoint, it is a good idea to choose a final solution set from all the examined solutions. Several approaches have already been proposed for solution selection. However, design of many-objective algorithms in this framework has not been discussed in many studies.

Future Directions

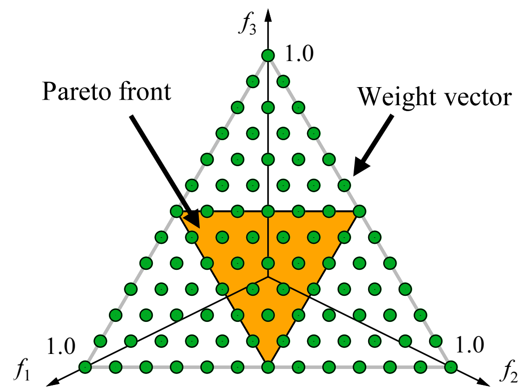
3. Test problems

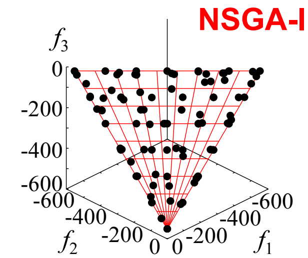
A wide variety of many-objective test problems are needed. Especially, realistic test problems are needed. It is also needed to analyze the relation between test problems and real-world problems.

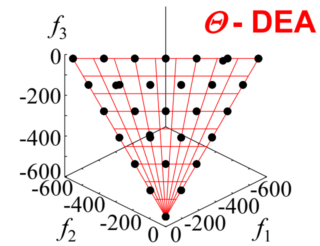
4. Performance Evaluation Methods

How to evaluation a solution set of a many-objective problem is an important research issue. Well-known and frequently-used performance indicators are not always appropriate for many-objective problems.

Adaptation of Weight/Reference Vectors







Adaptation of Weight/Reference Vectors

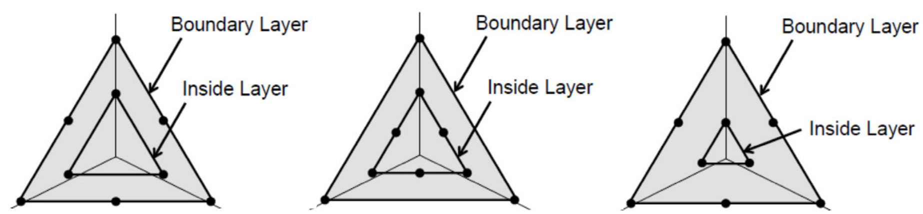
Question:

What is a good distribution of 200 reference vectors in a 10-dimensional objective space? We need 10 million solutions to cover the entire Pareto front.

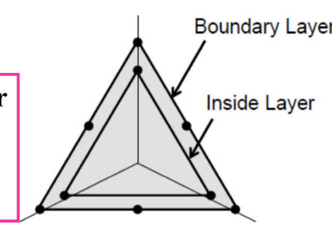
***k*-Objective Problem**	$5^{(k-1)}$
2-Objective Problem	5
3-Objective Problem	25
10-Objective Problem	10 million

Adaptation of Weight/Reference Vectors

For example, it is not unclear how to specify the boundary vectors and the inside vectors in the two-layer method.

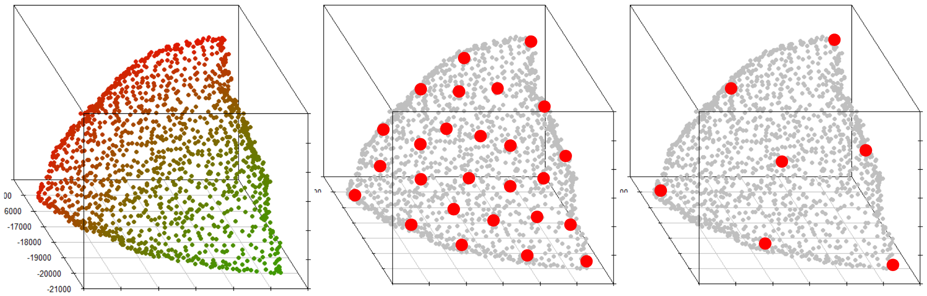


H. Ishibuchi et al., Two-Layered Weight Vector Specification in Decomposition-Based Multi-Objective Algorithms for Many-Objective Optimization Problems, IEEE CEC 2019.



Solution Selection

How many solutions do you need?

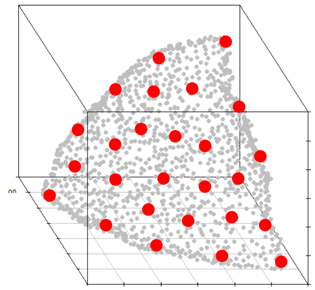


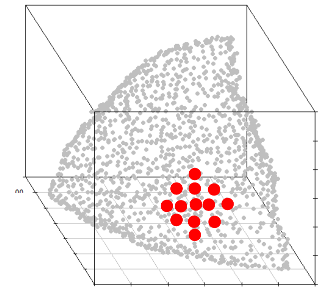
One Extreme (Left): A large number of solutions are needed when the Pareto front is to be carefully examined.

Another Extreme (Right): Only a small number of solutions are needed when a single solution is to be quickly chosen.

Solution Selection

Which is your goal?





Left: To find a well-distributed solution set over the entire Pareto front (very difficult for many-objective problems).

Right: To find a well-distributed solution set over a small region of the Pareto front (Q: how to specify the region of interest?) ==> Interactive approach.

Algorithm Design

(1) How to search for non-dominated solutions from which candidate solutions are selected.
The search result is all non-dominated solutions among the examined solutions. The question is how to search for a wide variety of good non-dominated solutions under this framework.

(2) How to chose candidate solutions which are presented to the decision maker.
Research topics may include (i) choice of a selection criterion (e.g., IGD), (ii) design of efficient algorithms, (ii) interaction with the decision maker.

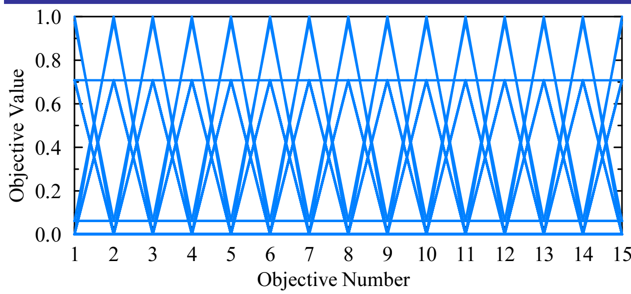
Many-Objective Test Problems

Current Trend:
DTLZ and WFG test problems have been used in evolutionary many-objective optimization studies.

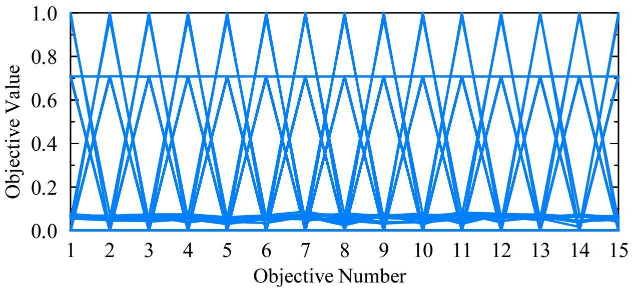
Reported Results:
Very good results have been reported for DTLZ and WFG test problems with many objectives (e.g., 10 objectives, 15 objectives).

==> They may be very easy test problems while many-objective problems should be difficult.

Results on a 15-Objective DTLZ 2 Problem
Perfect Solution Set by MOEA/D-PBI

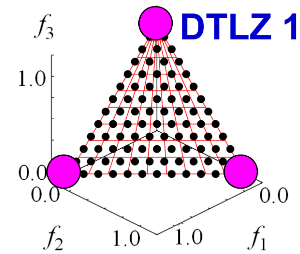


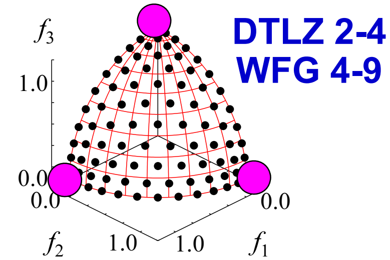
Reference points for IGD calculation on the Pareto front (Final Goal)



Obtained solutions (Actual results)

Normalized Pareto Fronts in $[0, 1]^m$ of DTLZ 1-4 and WFG 4-9





Pareto Fronts (p = 1 or p = 2):

$$(f_1)^p + (f_2)^p + ... + (f_m)^p = 1$$

Important Feature: Any extreme solution can minimize (m - 1) objectives. For example, (1, 0, ..., 0) is the best solution for all objectives except for the first objective.

Normalized Pareto Fronts in $[0, 1]^m$ of DTLZ 1-4 and WFG 4-9

Pareto Fronts (p = 1 or p = 2):

$$(f_1)^p + (f_2)^p + ... + (f_m)^p = 1$$

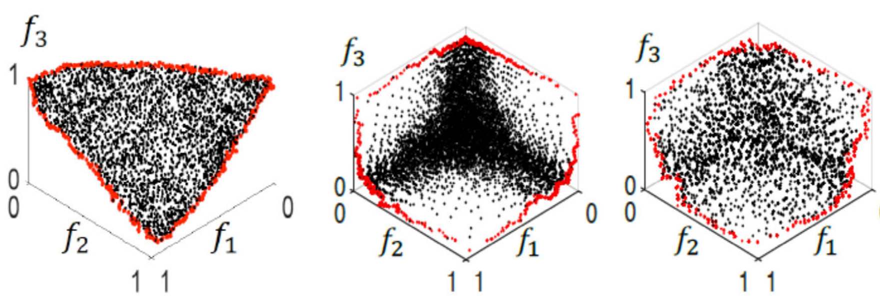
Important Feature: Any extreme solution can minimize (m - 1) objectives. For example, (1, 0, ..., 0) is the best solution for all objectives except for the first objective.

⇩

A multiobjective problem with (m - 1) objectives has the single best solution for all objectives.
==> There is no conflict among (m - 1) objectives.

Strange, Unrealistic, ...

Normalized Pareto Fronts in $[0, 1]^m$ of Other Test Problems



(a) Knapsack. (b) Assignment. (c) TSP.

H. Ishibuchi et al., Regular Pareto Front Shape is not Realistic, IEEE CEC 2019.

Wide Variety of Test Problems including realistic test problems

DTLZ and WFG have been still frequently used.

==> New test problems are needed.

In the last few years, some new test problem sets have been proposed. See

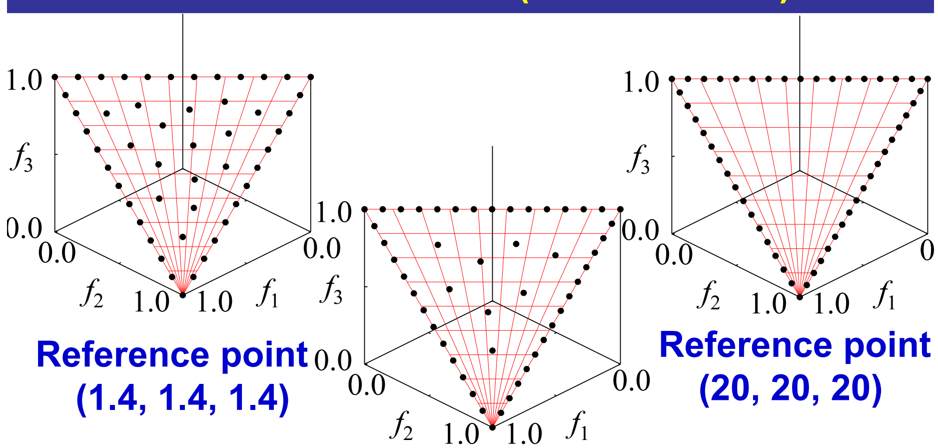
S. Zapotecas-Martınez, C. A. C. Coello, H. Aguirre, K. Tanaka, "A Review of Features and Limitations of Existing Scalable Multi-Objective Test Suites", IEEE TEVC, Vol. 23, No. 1, 130 - 142, Feb. 2019.

Hypervolume (HV) and IGD

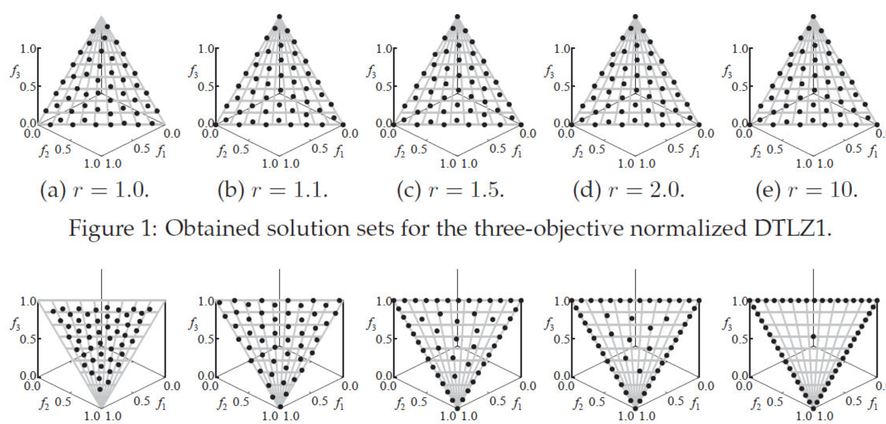
4. Performance Evaluation Methods

How to evaluation a solution set of a many-objective problem is an important research issue. Well-known and frequently-used performance indicators are not always appropriate for many-objective problems.

Optimal Distribution of Solutions for HV Inverted DTLZ1 Problem (Minimization)



HV: Dependency of Optimal Distribution of Solutions on the Shape of the Pareto Front



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

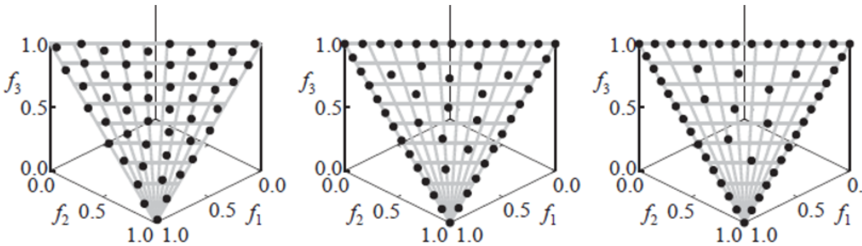
Figure 1: Obtained solution sets for the three-objective normalized DTLZ1.

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

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

Optimal Distribution of Solutions depends on the reference point specification (HV)

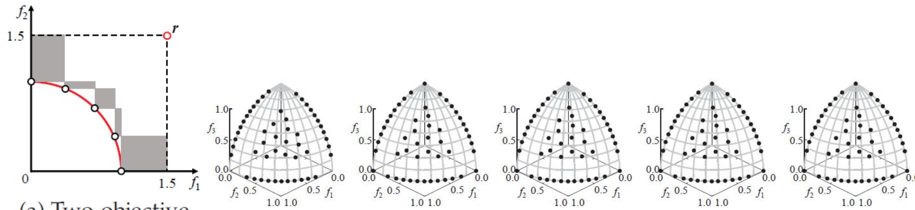
==> This means that the best weight (reference) vector specification in MOEA/D, NSGA-III, MOEA/DD etc. depends on the reference point specification.



(b) $r = 1.1$. (c) $r = 1.5$. (d) $r = 2.0$.

More boundary vectors are needed.

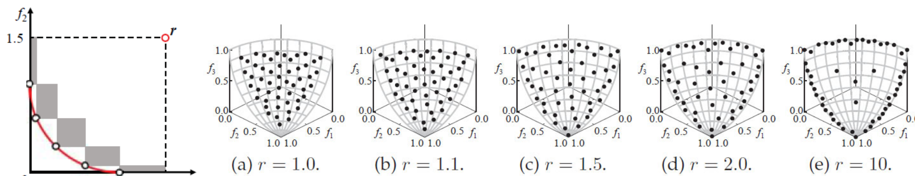
HV: Dependency of Optimal Distribution of Solutions on the Shape of the Pareto Front



(a) Two-objective DTLZ2.

(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.

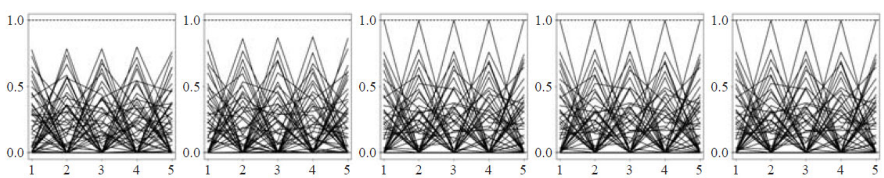


(b) Two-objective Minus-DTLZ2.

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

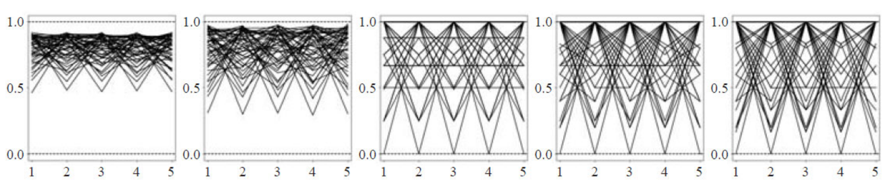
Figure 9: Obtained solution sets for the three-objective Minus-DTLZ2 problem.

HV: Dependency of Optimal Distribution of Solutions on the Shape of the Pareto Front



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

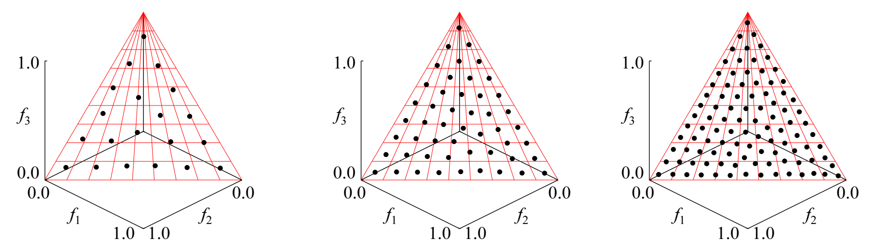
Figure 6: Obtained solution sets for the five-objective normalized DTLZ1 problem.



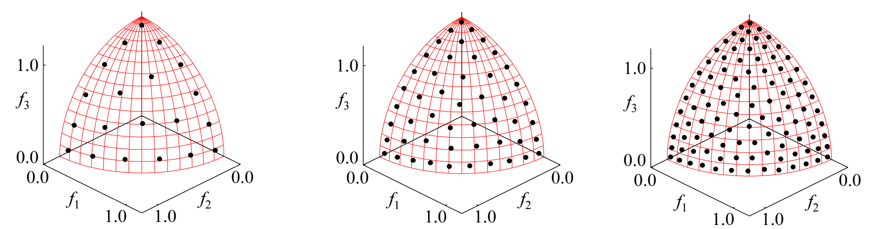
(a) $r = 1.0$. (b) $r = 1.1$. (c) $r = 1.5$. (d) $r = 2.0$. (e) $r = 10$.

Figure 7: Obtained solution sets for the five-objective Minus-DTLZ1 problem.

Optimal Distributions of Solutions for IGD are not always intuitive (Uniform Reference Points)

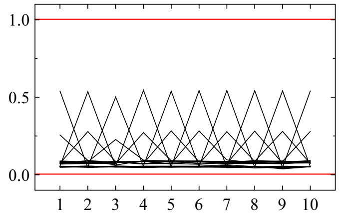


Population size 20 **Population size 50** **Population size 100**

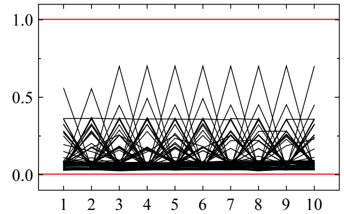


Population size 20 **Population size 50** **Population size 100**

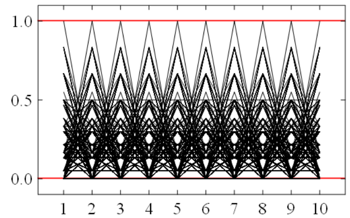
Optimal Distributions of Solutions for IGD are not always intuitive (Uniform Reference Points)



Population size 20



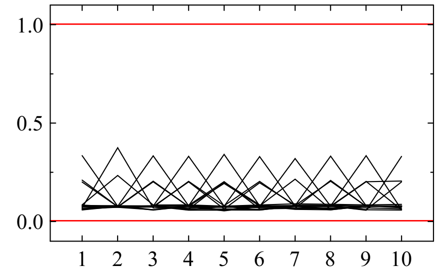
Population size 100



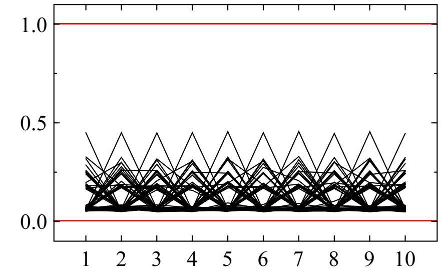
Reference points: 10,010 points

Optimal Distributions of Solutions for IGD are not always intuitive (Random Reference Points)

When we randomly generate 100,000 reference points, the optimal distributions of solutions are as follows:



Population size 20



Population size 100

Other Topics

1. **Objective Selection:** All objectives are not always equally important. Some objectives can be removed. Objective selection is (i) to improve the efficiency of many-objective search, and (ii) to help the solution selection by decreasing the number of non-dominated solutions.
2. **Normalization:** Objective space normalization is included in many EMO algorithms. Its necessity is clear. But, it also has some potential negative effects.

EMO-Related Future Events

September 16-20, 2019, Netherlands

Lorentz Workshop

Organizers: M. Emmerich, B. Naujoks, D. Brockhoff, R. Purshouse

Many Criteria Optimization and Decision Analysis

January 12-17, 2020, Germany

Dagstuhl Seminar

Organizers: C. M. Fonseca, K. Klamroth, G. Rudolph, M. M. Wiecek

Scalability in Multiobjective Optimization

Scalability
A number of research topics

- a large number of objectives (many-objective)
- a large number of variables (large-scale)
- a large number of constraints (many-constraints)
- high percentage of infeasible solutions
- a number of overlapping Pareto solutions in the objective space (multi-modal).
- a number of local Pareto fronts (multi-modal).
- expensive fitness evaluation (==> surrogate)
- search for a huge number of non-dominated solution for knowledge extraction

EMO 2021, Shenzhen, China
March 28-31, 2021, SUSTech