Tackling Many Objectives

Xin Yao
Department of Computer Science and Engineering
Southern University of Science and Technology (SUSTech)
Shenzhen, China
and
School of Computer Science, University of Birmingham, UK

What is SUSTech?

- Southern University of Science and Technology (SUSTech)
- The youngest public university in China
- Started officially in 2012 as a research intensive international university
- Only 6 year old

Where is SUSTech?

- Shenzhen.
- But where is Shenzhen?

Shenzhen: Silicon Valley in China

"Shenzhen is the city that most resembles the Silicon Valley."

--- *Business Week (2014)*

The most competitive city economically in China

- ✓ High technology
- ✓ Finance services
- ✓ Logistics industry
- ✓ Culture industry

Department of Computer Science and Engineering (CSE)

- Only established in 2016
- Currently 20 tenure track faculty members, recruited from all over the world (many more researchers)
- Growth plan: 55 tenure track faculty members

OPAL (OPtimisation And Learning) Group

- **5 tenure track professors**
  - **Ran Cheng, Hisao Ishibuchi, Yuhui Shi, Ke Tang, Xin Yao**
- **~10 research professors and postdoc researchers.**
- **International recognition**
  - **3 IEEE fellows (H-indices: 90, 66, 47): ~120K Google Scholar citations;**
    - **Hisao Ishibuchi, Yuhui Shi, Xin Yao;**
  - **IEEE Computational Intelligence Society (CIS) *Evolutionary Computation Pioneer Award* (Xin) and *Fuzzy Systems Pioneer Award* (Hisao)**
  - **IEEE CIS *Early Career Award* (Ke Tang)**
  - **IEEE CIS *Outstanding PhD Dissertation Award* (Ran Cheng)**
  - **JSPS Award (Hisao)**
  - **Royal Society Wolfson Research Merit Award (Xin)**
  - **Royal Society Newton Advanced Fellowship (Ke)**

Some Research Topics

计算机科学与工程系
Department of Computer Science and Engineering

- **Evolutionary learning and optimisation**
  - Multi-objectivity
  - Dynamics
  - Uncertainty
- **Ensemble machine learning**
  - Online learning
  - Class imbalance learning
- **Neural network learning**
- **Real-world applications**
  - Smart logistics
  - Search-based software engineering
  - Fault diagnosis
  - Reconfigurable computing architectures

We Have Many Openings

- There is nothing more exiting than being part of a growing story.
- We recruit
  - Faculty members
  - Postdoctoral research fellows
  - PhD students
  - Visitors
- We teach in English.
- Email: xiny@sustc.edu.cn

Outline

- Introduction
- Objective Reduction
- Alternative Dominance Relationship
- Improved Two-Archive Algorithm (Two_Arch2)
- Conclusions and Future Work

What is multi-objective optimisation?

- **More than one objective to be optimised, with or without constraints.**

$$
\begin{aligned}
&\text{min/max } f_m(\boldsymbol{x}), && m = 1, 2, \cdots, M \\
&\text{subject to } g_j(\boldsymbol{x}) \geq 0, && j = 1, 2, \cdots, J \\
&\qquad h_k(\boldsymbol{x}) = 0, && k = 1, 2, \cdots, K \\
&\qquad \underbrace{x_i^{(L)}}_{\text{lower bound}} \leq x_i \leq \underbrace{x_i^{(U)}}_{\text{upper bound}}, && i = 1, 2, \cdots, n
\end{aligned}
$$

Multi-objective Evolutionary Algorithms

- **MOEAs have been widely used in the last 20 years for multi-objective optimisation.**
- **They can provide a *set* of non-dominated solutions in a *single* run without requiring the set of weights.**
- **They do not require the objective functions to be convex, smooth, or even continuous.**
- **They can handle nonlinear constraints.**
- **They can deal with uncertainty and dynamics.**

An Indicator of MOEA’s Impact

- **Deb, K., Pratap. A, Agarwal, S., and Meyarivan, T. (2002). A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transaction on Evolutionary Computation*, 6(2), 181-197. (29,491 Google Scholar citations)**

Wow!

- **That’s impressive.**

Unfortunately

- **NSGA-II and other early MOEAs work well only with 2 or 3 objectives.**
- **They do not work well when the number of objectives goes beyond that.**
- **There is a scalability issue in terms of the number of objectives.**
- **In this talk, we consider Many Objective Optimisation, indicating the number of objectives is greater than three.**

Two Possible Approaches to Problem-Solving

1. Develop more sophisticated solutions to complex problems.

2. Simplify a complex problem so that an existing solution can be applied.

Can we simplify MaOPs into MOPs?

Can we reduce the number of objectives?

Outline

- Introduction
- Objective Reduction
- Alternative Dominance Relationship
- Improved Two-Archive Algorithm (Two_Arch2)
- Conclusions and Future Work

Objective Reduction

- If two objectives are positively correlated, we need to optimise only one of them.
- There are many methods that could be used to reduce the number of objectives.
- We give one example here.

Nonlinear Correlation Information Entropy (NCIE)

- **NCIE is an entropy measure.**
- **NCIE firstly divides variables *X* and *Y* into *b*b* uniform rank grids. Then, the probabilities p_{ij} can be approximated by counting the samples in those grids. In other words, p_{ij} in the *ij*-th grid can be calculated by the number of solutions in *ij*-the grid (n_{ij}/N).**
- **Parameter *b* can be set as *N^0.5*.**

$$H^r(X) = -\sum_{i=1}^{b} \frac{n_i}{N} log_b(\frac{n_i}{N})$$

$$H^r(X,Y) = -\sum_{i=1}^{b}\sum_{j=1}^{b} \frac{n_{ij}}{N} log_b(\frac{n_{ij}}{N})$$

$$NCIE(X,Y) = H^r(X) + H^r(Y) - H^r(X,Y)$$

Objective Reduction Based on NCIE

- **Correlation analysis is based on the matrix of modified NCIE R^N of *the non-dominated population*.**

$$R^N = \{Sgn(cov_{ij})NCIE_{ij}\}, (1 \le i, j \le m)$$

- **Objective selection aims to choose the most conflicting objectives.**
  - **Our approach is applied in every generation of MOEAs to update the correlation information among objectives.**

H. Wang and X. Yao, "Objective Reduction Based on Nonlinear Correlation Information Entropy," *Soft Computing*, June 2016, Volume 20, Issue 6, pp 2393–2407.

Objective Selection: An Example

- **Select the most conflicting objective**
- **Remove the objectives that are positively correlated to the selected objective**

	f_1	f_2	f_3	f_4	f_5
f_1	1.0000	0.4959	0.4244	0.5348	-0.3552
f_2	0.4959	1.0000	0.3972	0.4686	-0.3381
f_3	0.4244	0.3972	1.0000	0.4765	-0.4352
f_4	0.5348	0.4686	0.4765	1.0000	-0.4488
f_5	-0.3552	-0.3381	-0.4352	-0.4488	1.0000
$\sum NCIE < 0$	-0.3552	-0.3381	-0.4352	-0.4488	-1.5773

- ✓ f_5 is selected, because it has the most conflicting degree with other objectives.
- ✓ There is no objective positively correlated to f_5, thus, there is not a redundant objective with f_5 in the remaining objectives.
- ✓ f_4 is selected, because it has the largest absolute sum of NCIEs to other objectives. f_1 , f_2, and f_3 are omitted, they are all positively correlated to f_4 .
- ✓ Output $\{f_5, f_4\}$

Objective reduction can remove redundant objectives, but what if there is no redundancy?

Outline

- Introduction
- Objective Reduction
- Alternative Dominance Relationship
- Improved Two-Archive Algorithm (Two_Arch2)
- Conclusions and Future Work

Why Are Many Objectives Hard?

- The number of non-dominated solutions increases exponentially as the number of objectives grows.
- As a result, there is no selection pressure in MaOEAs to drive the evolutionary search.
- Can we use alternative dominance relationship other than Pareto dominance?

Θ-dominance --- Intuition

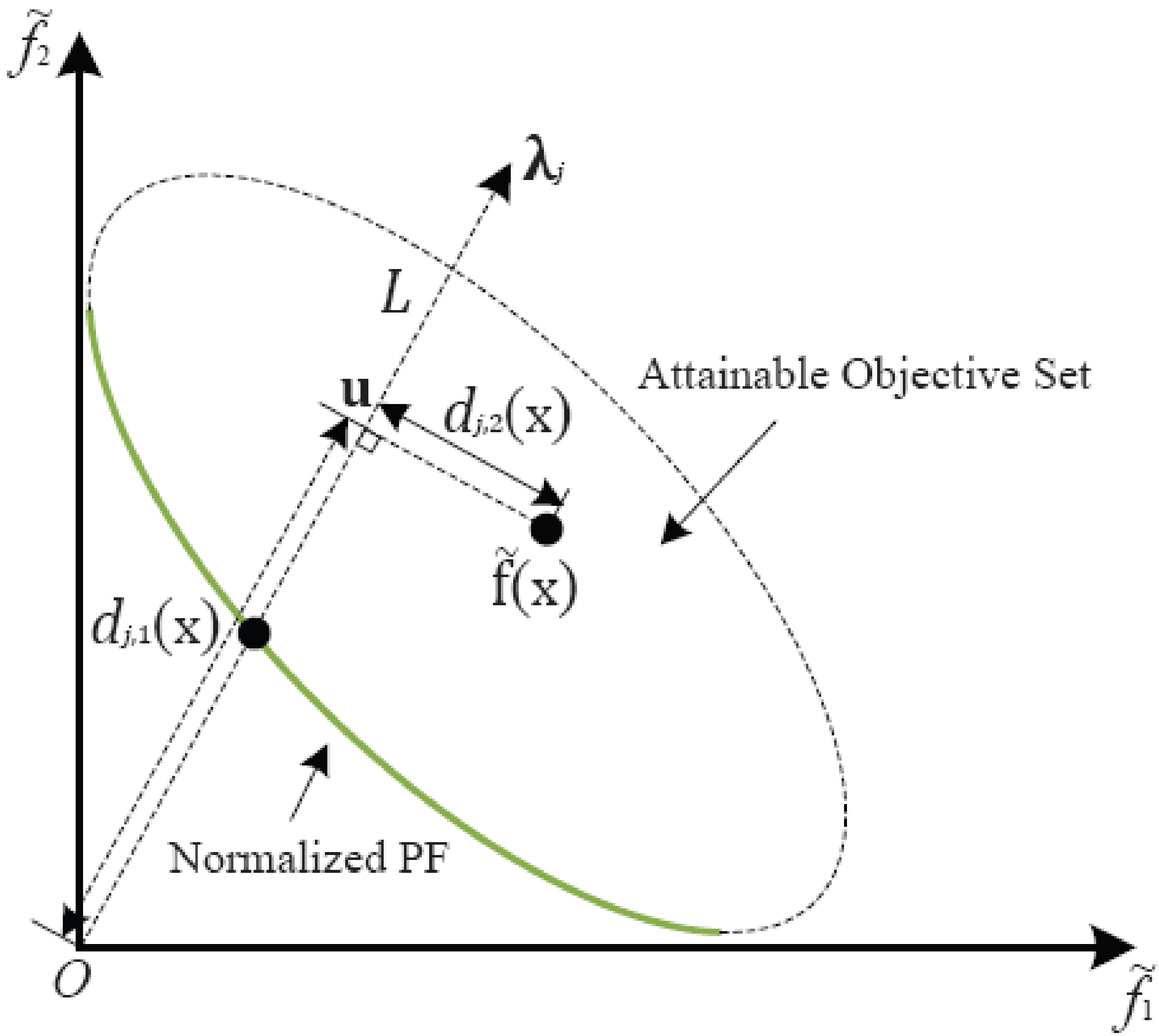


Fig. 3. Illustration of distances $d_{j,1}(\mathbf{x})$ and $d_{j,2}(\mathbf{x})$.

- ***f*'s are normalised objective functions.**
- **λ is the reference direction (point).**

•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):16-37, February 2016.

Θ-dominance --- Definition

Definition 7: Given two solutions $\mathbf{x}, \mathbf{y} \in S_t$, $\mathbf{x}$ is said to θ-dominate $\mathbf{y}$, denoted by $\mathbf{x} \prec_\theta \mathbf{y}$, iff $\mathbf{x} \in C_j$, $\mathbf{y} \in C_j$, and $\mathcal{F}_j(\mathbf{x}) < \mathcal{F}_j(\mathbf{y})$, where $j \in \{1, 2, \ldots, N\}$.

$$\mathcal{F}_j(\mathbf{x}) = \bar{d}_{j,1}(\mathbf{x}) + \theta d_{j,2}(\mathbf{x})$$

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):16-37, February 2016.

Balancing Convergence and Diversity

- The form of $F_j(x)$ indicates that the balance between convergence and diversity is very important in MaOEAs.
- **Why not manipulating the balance explicitly?**
  - **Y. Yuan, H. Xu, B. Wang, B. Zhang and X. Yao, "Balancing Convergence and Diversity in Decomposition-Based Many-Objective Optimizers," *IEEE Transactions on Evolutionary Computation*, 20(2):180-198, April 2016.**
- **Don't know how to strike the balance?**
  - **B. Li, K. Tang, J. Li and X. Yao, ``Stochastic Ranking Algorithm for Many-Objective Optimization Based on Multiple Indicators,'' *IEEE Transactions on Evolutionary Computation*, 20(6):924-938, December 2016.**

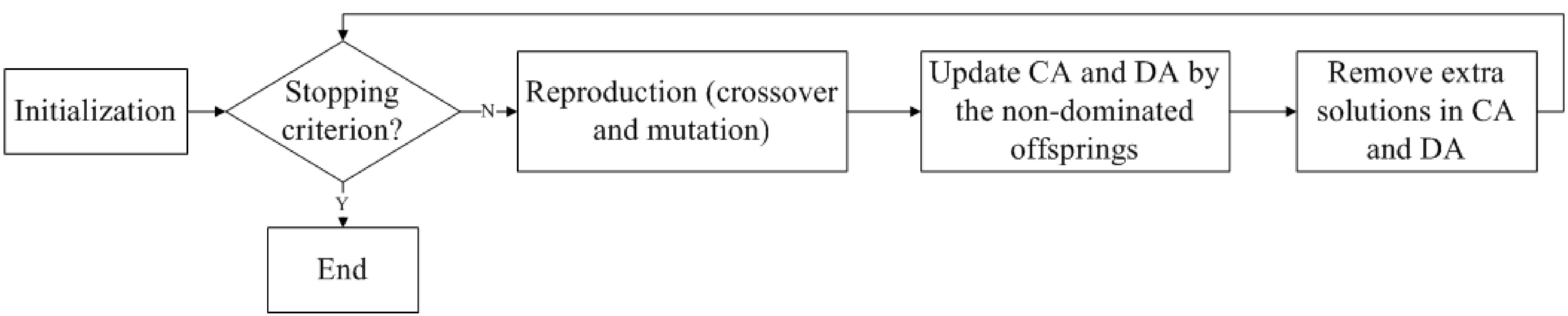
What if alternative dominance relationships still do not provide a satisfactory solution to a MaOP?

Outline

- Introduction
- Objective Reduction
- Alternative Dominance Relationship
- Improved Two-Archive Algorithm (Two_Arch2)
- Conclusions and Future Work

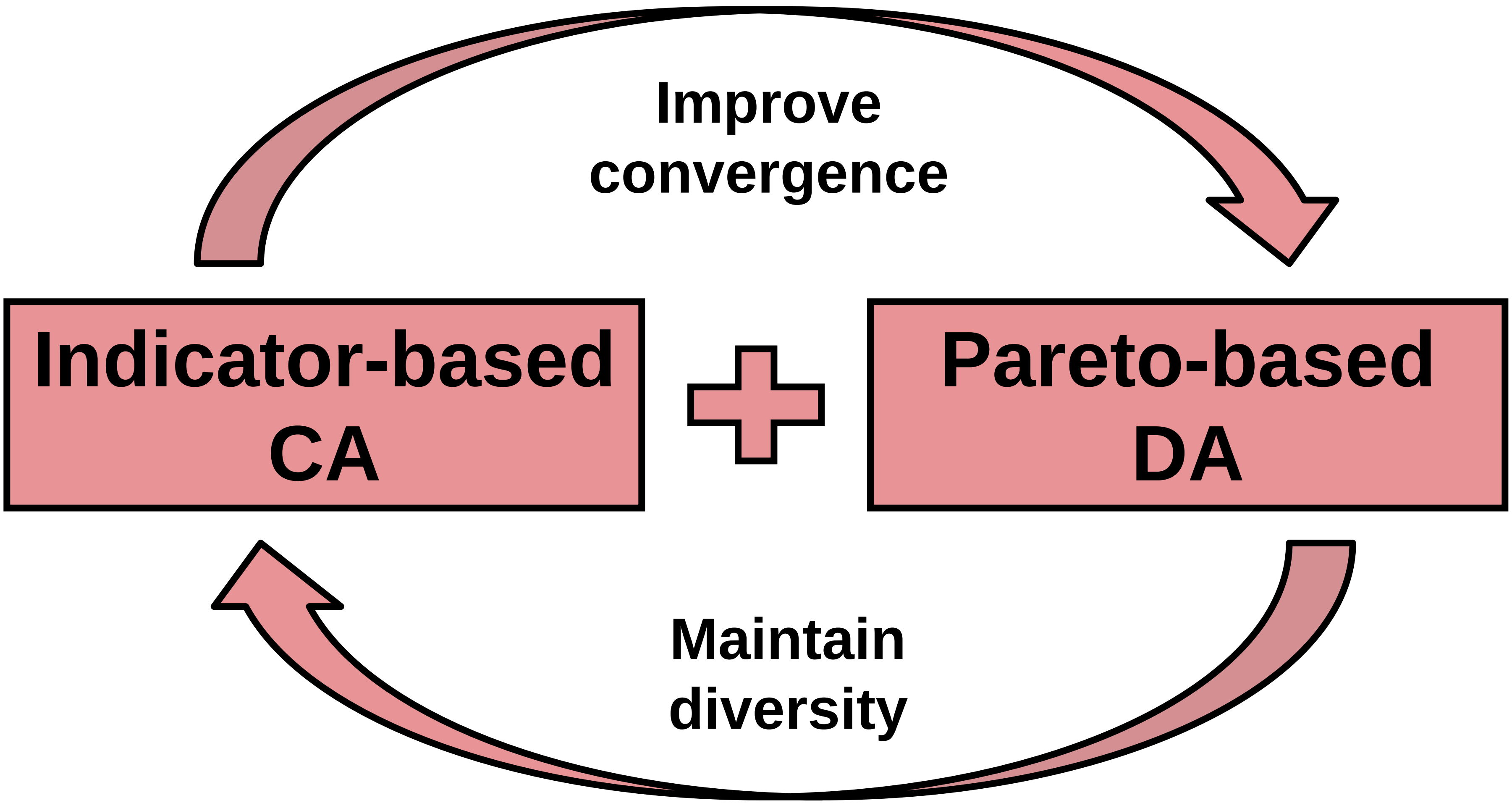
Two-Archive Algorithm

- Two-Archive algorithm (Two_Arch) maintains two archives (CA and DA) to promote convergence and diversity separately.
  - K. Praditwong and X. Yao, "A New Multi-objective Evolutionary Optimisation Algorithm: The Two-Archive Algorithm," *Proc. of the 2006 International Conference on Computational Intelligence and Security (CIS'2006)*, 3-6/11/2006, Ramada Pearl Hotel, Guangzhou, China. IEEE Press, Volume 1, pp.286-291.



<u>Improved</u> Two-Archive Algorithm:

Main Idea



Two_Arch2: Main Steps

Step 1: Initialization.
Step 2: Output <u>DA</u> if the stopping criterion is met, otherwise continue.
Step 3: Generate new solutions from CA and DA by crossover and mutation.
Step 4: Update CA and DA separately, go Step 2.

- **H. Wang, L. Jiao and X. Yao, "Two_Arch2: An Improved Two-Archive Algorithm for Many-Objective Optimization," *IEEE Transactions on Evolutionary Computation*, 19(4):524-541, August 2015.**

Convergence Archive (CA)

- **The quality indicator $I_{\varepsilon+}$ in IBEA is used in selection of CA. $I_{\varepsilon+}$ is an indicator that describes the minimum distance that one solution needs to dominate another solution in the objective space.**

$$I_{\varepsilon+}(x_1, x_2) = min_{\varepsilon}(f_i(x_1) - \varepsilon \leq f_i(x_2), 1 \leq i \leq m)$$

- **The fitness is assigned as below, the solution with the smallest fitness is removed from CA first.**

$$F(x_1) = \sum_{x_2 \in P \setminus \{x_1\}} -e^{-I_{\varepsilon+}(x_2, x_1)/0.05}$$

Diversity Archive (DA)

- **Update DA**
  - **When DA overflows, boundary solutions (solutions with maximal or minimal objective values) are firstly selected.**
  - **In the iterative process, the most different solution from the current DA is added until reaching the size.**
- **L_p-norm distance is adopted as the similarity measure in DA.**
- **DA is used as the final output of Two_Arch2.**

Degraded Euclidean Distance (Distance Concentration) in High-Dimensional Space

- **The Euclidean distance (L_2-norm) degrades its similarity indexing performance in a high-dimensional space.**
- **Most of existing diversity maintenance methods use the Euclidean distance to measure similarity among solutions for MaOPs.**

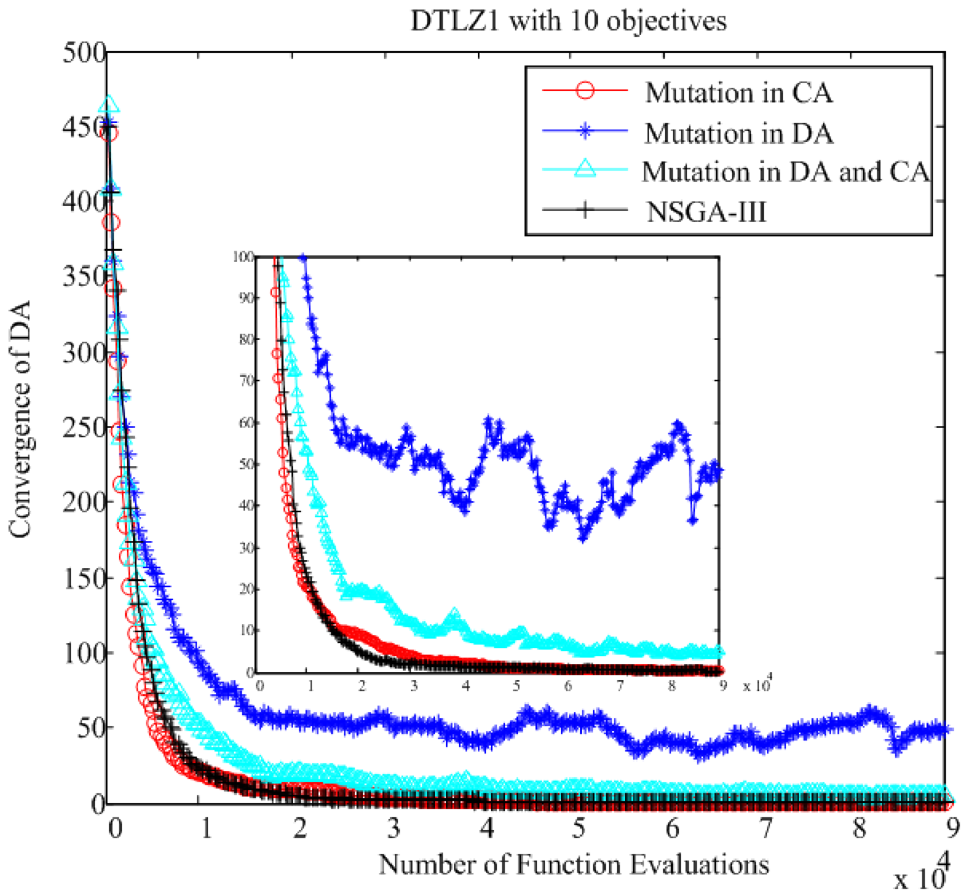
C. C. Aggarwal, A. Hinneburg and D. A. Keim, "On the surprising behavior of distance metrics in high dimensional space." Springer, 2001.

Similarity in High-Dimensional Space

- **The fractional distances (L_p-norm, p<1) perform better in a high-dimensional space.**
- **$L_{1/m}$-norm is employed in Two_Arch2, where *m* is the number of objectives.**

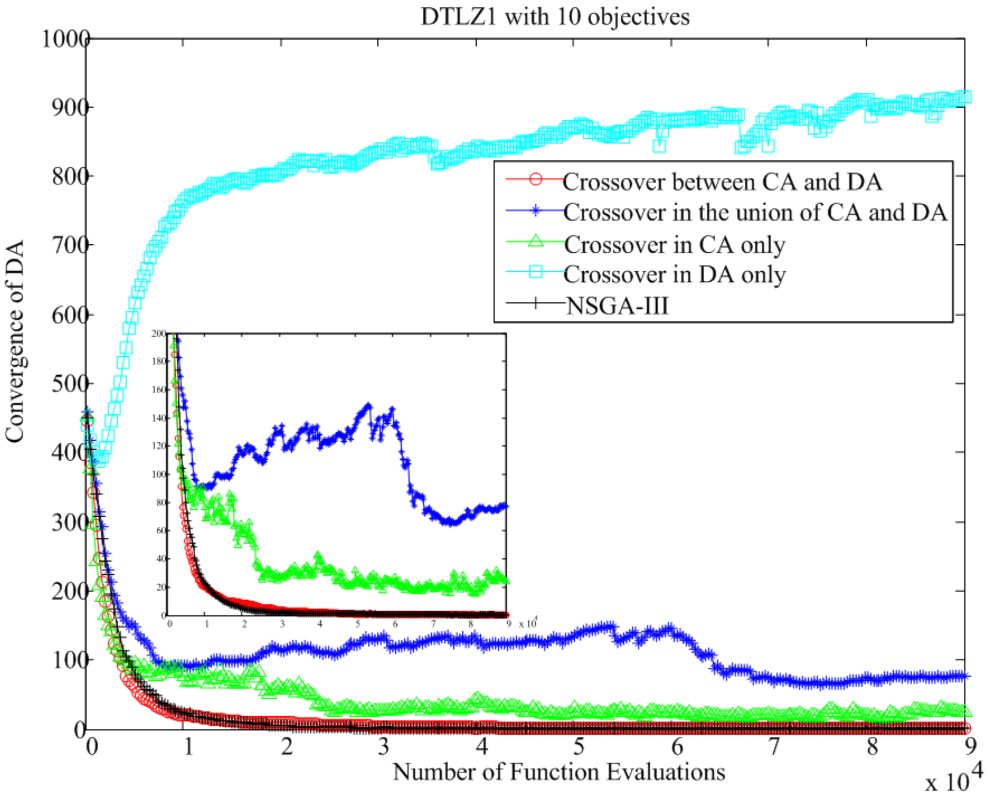
Interaction between CA and DA: Mutation

- Mutation to DA does not speed up convergence, and disturbs the guidance of CA to DA.
- Mutation is applied to CA only in Two_Arch2.



CA leads convergence

Interaction between CA and DA: Crossover

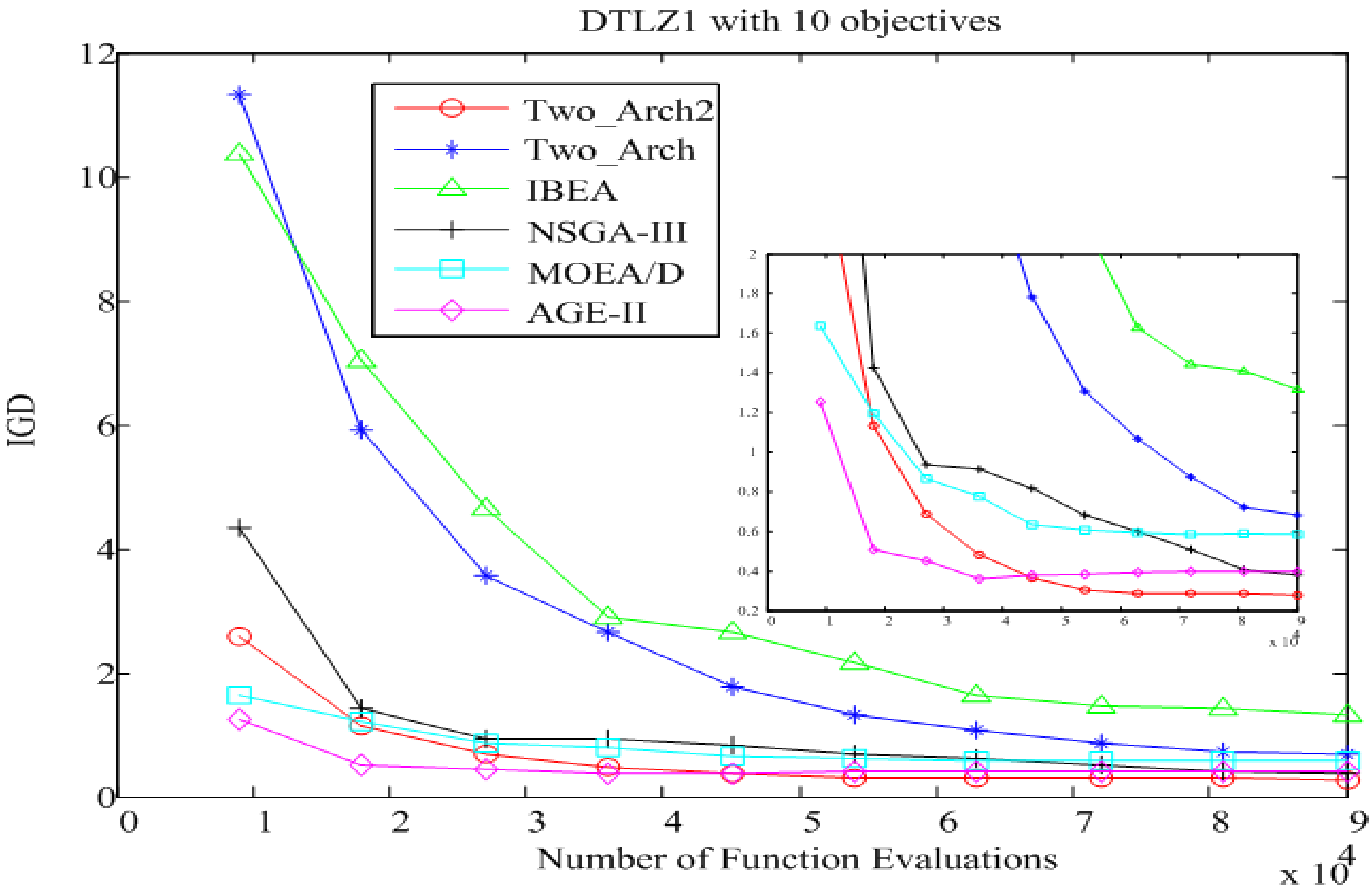


- The crossover between CA and DA has the fastest convergence speed.
- The crossover between CA and DA is employed in Two_Arch2.

Experimental Comparisons

- **Two_Arch2: Developed here**
- **Two_Arch: a reference to show the improvement of Two_Arch2 on MaOPs**
- **IBEA: indicator-based ($I_{\varepsilon+}$) MOEA with good convergence but poor diversity**
- **NSGA-III: newly-proposed MOEA with reference points for MaOPs**
- **MOEA/D: aggregation function-based MOEA**
- **AEG-II: Pareto-based MOEA with the ε-grid approximation in the objective space**

DTLZ1 with 10 Objectives



More Problems, More Objectives

- **More experimental results are in**
  - **H. Wang, L. Jiao and X. Yao, "Two_Arch2: An Improved Two-Archive Algorithm for Many-Objective Optimization,"** ***IEEE Transactions on Evolutionary Computation*****, 19(4):524-541, August 2015.**

Including Matlab code.

Outline

- **Introduction**
- **Objective Reduction**
- **Alternative Dominance Relationship**
- **Improved Two-Archive Algorithm (Two_Arch2)**
- **Conclusions and Future Work**

Conclusions

- **There are three major approaches to dealing with a large number of objectives:**
  1. **Objective reduction**
  2. **Alternative dominance relationship**
  3. **New algorithms**

- **This talk touches on only a tiny proportion of all the work. For more comprehensive review:**
  - **B. Li, J. Li, K. Tang and X. Yao, "Many-Objective Evolutionary Algorithms: A Survey,"** ***ACM Computing Surveys*****, 48(1), Article 13, 35 pages, September 2015.**

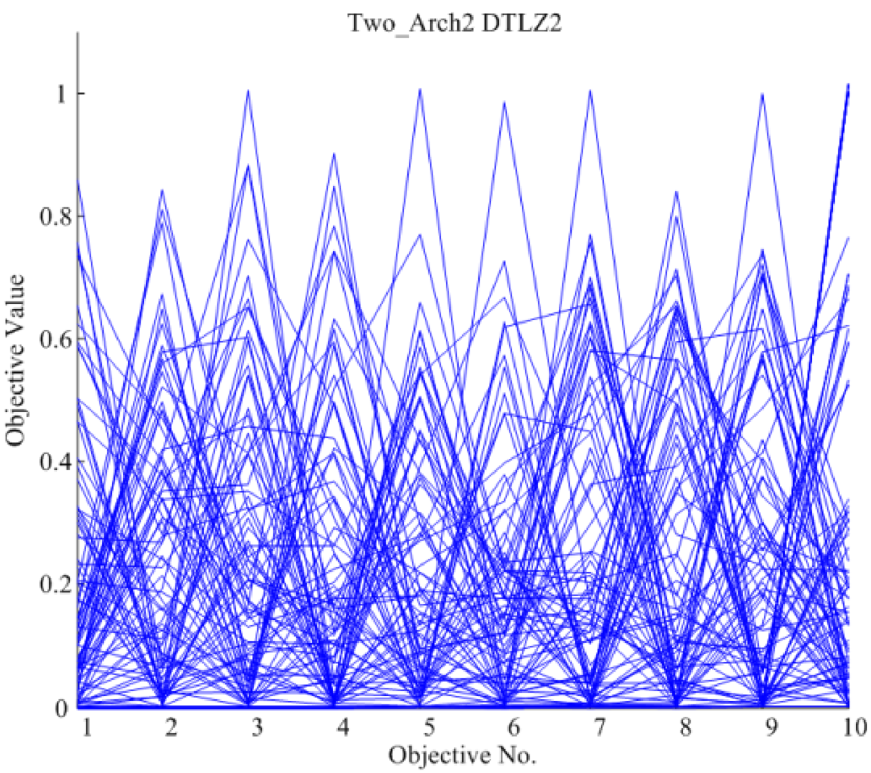
Future Work

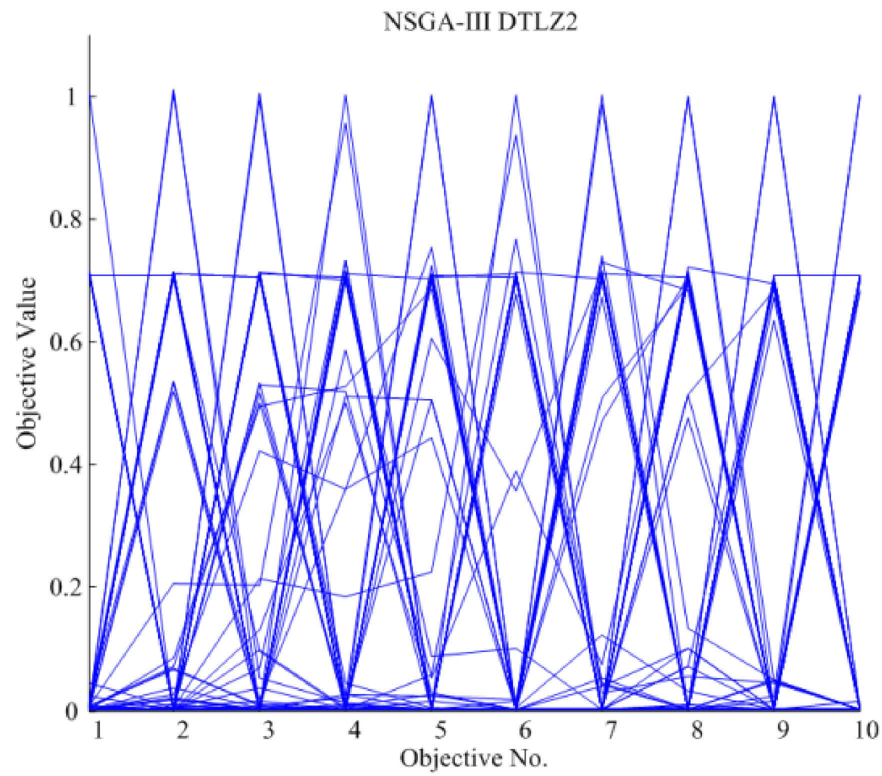
1. **Dynamic number of objectives, e.g.,**
   - R. Chen, K. Li and X. Yao, "Dynamic Multiobjectives Optimization With a Changing Number of Objectives," *IEEE Transactions on Evolutionary Computation*, vol. 22, no. 1, pp. 157-171, Feb. 2018.

2. **Constraint handling, e.g.,**
   - K. Li, R. Chen, G. Fu and X. Yao, "Two-Archive Evolutionary Algorithm for Constrained Multi-Objective Optimization," *IEEE Transactions on Evolutionary Computation*, online on 19/7/2018.
     DOI: 10.1109/TEVC.2018.2855411

Two_Arch2 vs. NSGA-III on DTLZ2 with 10 Objectives

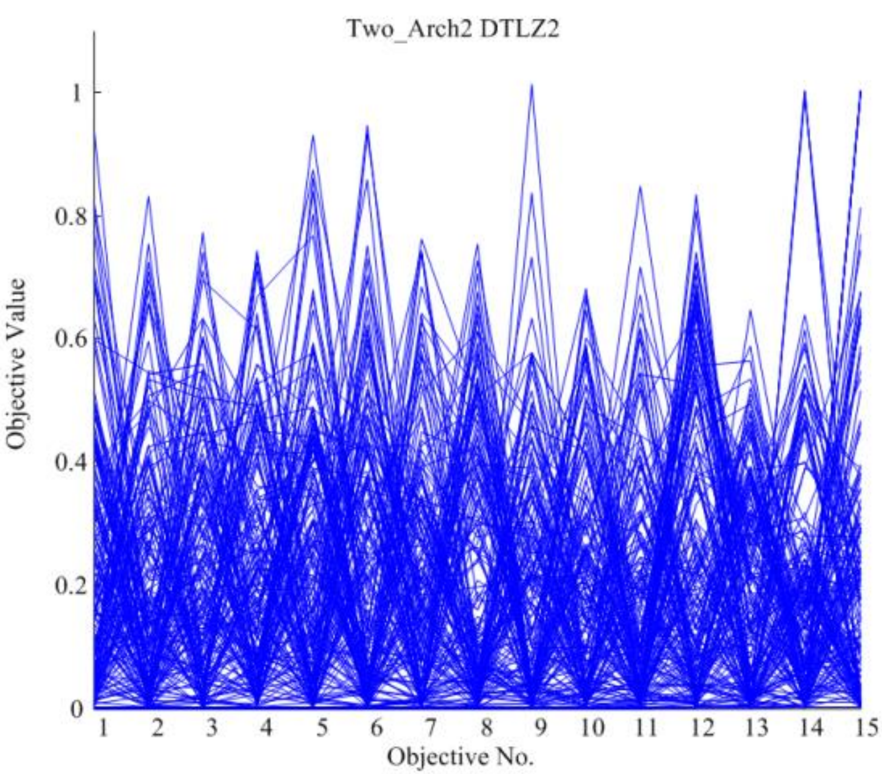
	Convergence	Diversity	Extreme point
Two_Arch2	Good	Good	Fair
NSGA-III	Good	Fair	Good

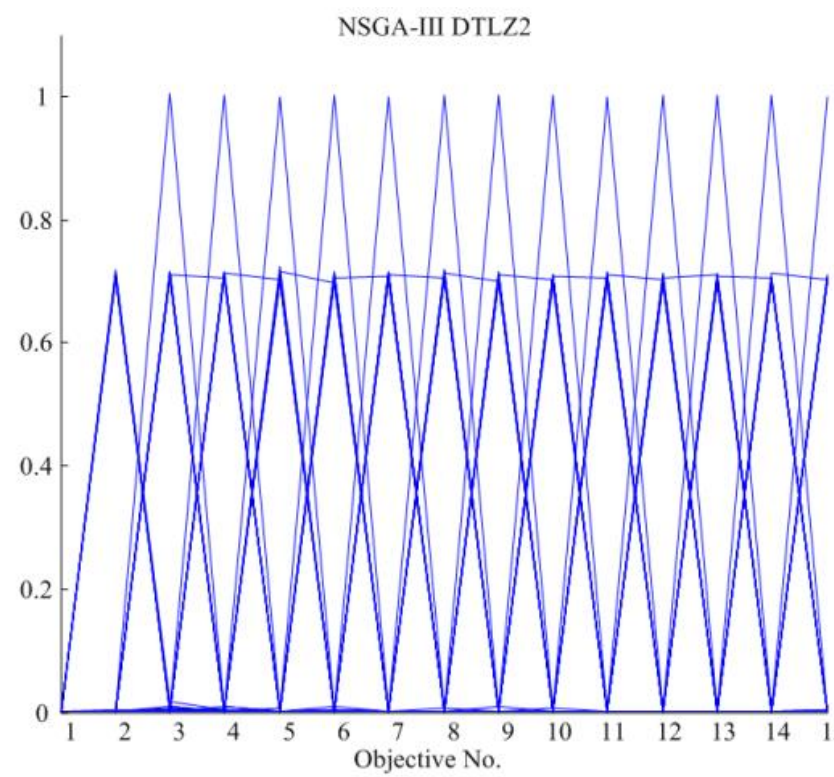




Two_Arch2 vs. NSGA-III on DTLZ2 with 15 Objectives

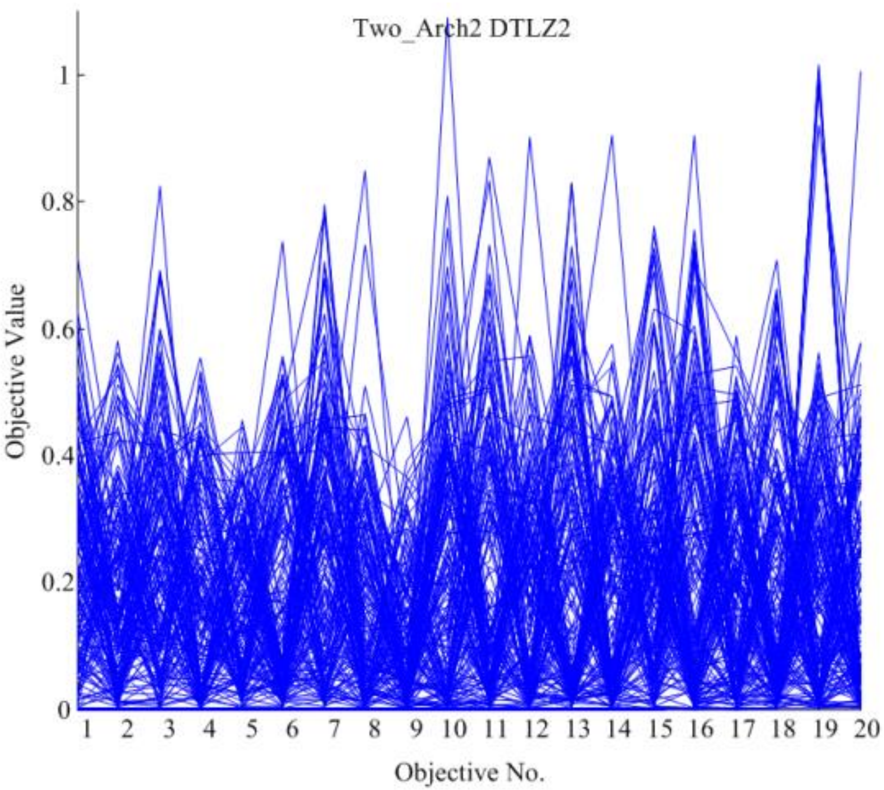
	Convergence	Diversity	Extreme point
Two_Arch2	Good	Good	Poor
NSGA-III	Good	Fair	Good

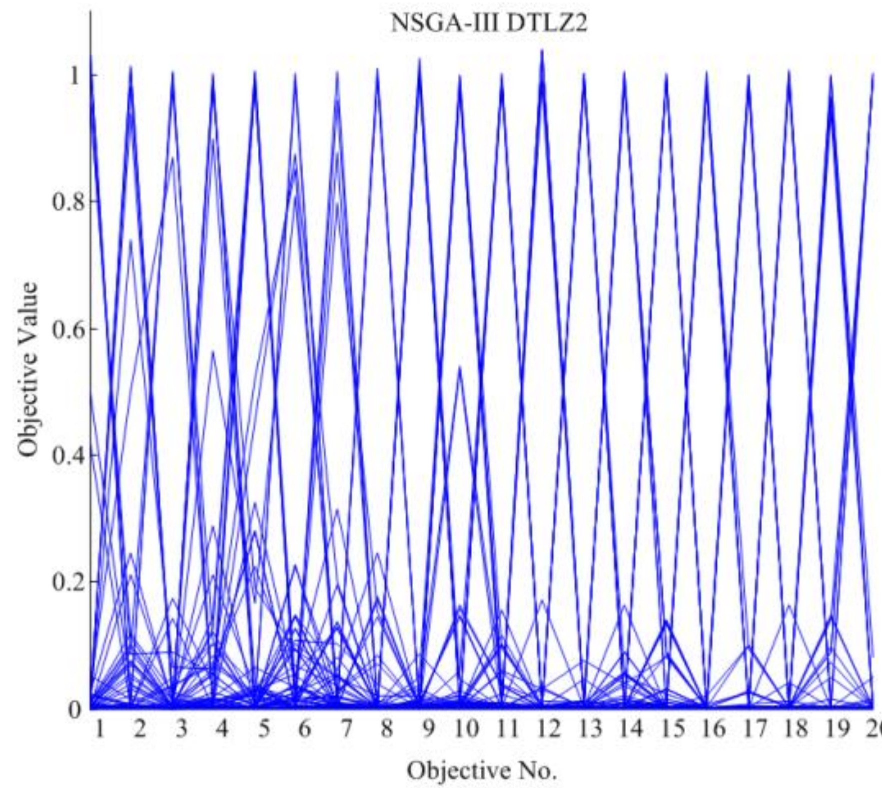




Two_Arch2 vs. NSGA-III on DTLZ2 with 20 Objectives

	Convergence	Diversity	Extreme point
Two_Arch2	Good	Good	Poor
NSGA-III	Good	Fair	Good





Two_Arch2 vs. NSGA-III

	Two_Arch2	NSGA-III
Convergence methodology	$I_{\varepsilon+}$	Pareto dominance
Convergence degeneration	No	No
Diversity maintenance	$L_{1/m}$-norm-based distance	Minimal perpendicular distances to reference points
Diversity degeneration	No	Increase with the dimension of objective space
Manual Settings	None	Reference points