

# Contents

<b>1</b>	<b>Introduction</b>	<b>9</b>
1.1	Motivation . . . . .	9
1.2	Overview . . . . .	10
<b>2</b>	<b>Unified Terminologies</b>	<b>13</b>
2.1	Motivation . . . . .	13
2.2	Specific Swarm-based Algorithms . . . . .	14
2.2.1	Bat-inspired Algorithm (BA) . . . . .	14
2.2.2	Grasshopper Optimization Algorithm (GOA) . . . . .	16
2.2.3	Crow Search Algorithm (CSA) . . . . .	17
2.2.4	Moth-flame Optimization Algorithm (MFO) . . . . .	19
2.2.5	Monarch Butterfly Optimization (MBO) . . . . .	20
2.2.6	Butterfly Optimization Algorithm (BOA) . . . . .	22
2.2.7	Particle Swarm Optimization (PSO) . . . . .	24
2.3	Unified Terminologies . . . . .	26
2.3.1	Unified Terminologies Catalog . . . . .	26
2.3.2	Extension in Other Swarm-based Algorithms. . . . .	28
2.4	Summary . . . . .	29
<b>3</b>	<b>Unified Procedure</b>	<b>35</b>
3.1	Introduction . . . . .	35
3.2	Unified Procedure . . . . .	38
3.3	Summary . . . . .	41
<b>4</b>	<b>Unified Framework</b>	<b>42</b>
4.1	Motivation . . . . .	42
4.2	Re-framed Nature-inspired Algorithms . . . . .	43
4.2.1	Re-framed BA . . . . .	43

4.2.2	Re-framed GOA . . . . .	45
4.2.3	Re-framed CSA . . . . .	47
4.2.4	Re-framed MFO . . . . .	48
4.2.5	Re-framed MBO . . . . .	50
4.2.6	Re-framed BOA . . . . .	52
4.2.7	Re-framed PSO . . . . .	54
4.3	Unified Nature-Inspired Optimization Algorithm (UNIOA) . . . . .	56
4.3.1	Unified Framework — UNIOA . . . . .	56
4.3.2	UNIOA Package . . . . .	60
4.4	Summary . . . . .	66
<b>5</b>	<b>Experimental Setup of Benchmark Study</b>	<b>67</b>
5.1	Motivation for experiments . . . . .	67
5.2	Summary . . . . .	70
<b>6</b>	<b>Experimental Results</b>	<b>71</b>
6.1	Introduction . . . . .	71
6.2	Experiments for avoiding side effects . . . . .	72
6.2.1	Results . . . . .	72
6.2.2	Conclusion . . . . .	77
6.3	Comparing UNIOA to the original framework . . . . .	78
6.3.1	Results . . . . .	78
6.3.2	Conclusion . . . . .	81
6.4	Comparing algorithms in UNIOA . . . . .	82
6.5	Summary . . . . .	84
<b>7</b>	<b>Summary</b>	<b>86</b>
7.1	Conclusions in our studies . . . . .	86
7.2	Limitations in our analysis . . . . .	87
7.3	Future work . . . . .	87
	<b>References</b>	<b>90</b>
	<b>Appendices</b>	<b>91</b>
A	Original positions of eight tuples in selected algorithms . . . . .	91
B	Re-framed Algorithm in the Unified Framework . . . . .	94
C	Experiments used in this work . . . . .	115

# List of Figures

1.1	Overview. . . . .	11
2.1	Entire compulsory components and some of selective components can comprise any of these seven algorithms. . . . .	31
4.1	Unified framework for Nature-Inspired Optimization Algorithm —UNIOA, only for seven selected algorithms . . . . .	61
6.1	Distribution of 24 AUC values for paired algorithm performances in syncE or asyncE, when $n = 5$ or $n = 20$ . . . . .	74
6.2	Distribution of 24 AUC values for paired algorithm performances in syncG or asyncG, when $n = 5$ or $n = 20$ . . . . .	76
6.3	Distribution of 24 AUC values for paired algorithm performances in orig framework or UNIOA framework, when $n = 5$ or $n = 20$ . . . . .	79
6.4	ERT values of the MFO in UNIOA and in ORIGINAL, when $n = 5$ . . . .	82
6.5	Distribution of 24 AUC values for UNIOA algorithms' performance, when $n = 5$ (top) or $n = 20$ (bottom). . . . .	83
6.6	ERT values of seven algorithms in UNIOA, when $n = 5$ . . . . .	84

# List of Tables

1	Some symbols in this work. . . . .	8
2.1	Unified Terminologies: all information appearing in these seven algorithms can be analyzed and then categorized into 20 components. The none in gray color means this information doesn't appear in this algorithm. . . .	32
3.1	These 20 unified components can be categorized into eight tuples, according to their functions. Numbers in the 'Component' column represents the index of information discussed in Table 2.1 . . . . .	37
3.2	Original positions of these eight tuples in these seven algorithms . . . . .	39
4.1	Representations of eight tuples in each algorithm. . . . .	64
4.2	In UNIOA, It is possible to customize <i>Tuple.</i> , but it is impossible to customize	65
5.1	Setups and usages in IOHproflier. . . . .	69
6.1	P-values of Wilcoxon signed-rank test for difference on average between performances of paired algorithms in syncE or asyncE for a random optimization function, when $n = 5$ or $n = 20$ . . . . .	74
6.2	P-values of Wilcoxon signed-rank test for difference on average between performances of paired algorithms in syncG or asyncG for a random optimization function, when $n = 5$ or $n = 20$ . . . . .	77
6.3	How each algorithm uses $\mathbf{x}_g$ in ORIGINAL framework and UNIOA framework.	77
6.4	When reproducing their original framework, if the way of calculating the Evaluation and the $\mathbf{x}_g$ changes. . . . .	78

6.5	(1) P-values of Wilcoxon signed-rank test for difference on average between performances of paired algorithms in ORIGINAL framework or UNIOA framework for a random optimization function, when $n = 5$ or $n = 20$ . (2) The number of wins orig had or UNIOA had when comparing their own AUC value throughout 24 BBOB functions. . . . .	80
7.15	28 pre-experiments for testing if syncE and asyncE will affect the algorithm performance differently. . . . .	115
7.16	Five pre-experiments for testing if syncG and asyncG will affect the algorithm performance differently. . . . .	116

# Symbols

Table 1: Some symbols in this work.

Symbol	Meaning
$\mathcal{U}(a, b)$	A single random uniform distributed number in $[a, b)$
$r$	A single random uniform distributed number in $[0, 1)$ . It is always randomly generated when utilized.
$Exp(a)$	A single random exponential distributed number with the scale $\lambda = a$
$\mathbf{Min}(\{a_i\}) \rightarrow a_i$	Minimization calculator that will generate one $a_i$ that has the minimum fitness value amongst the set $\{a_i\}$ .
$\mathbf{Sort}(\{a_i\}) \rightarrow \langle a_i \rangle$	Sort calculator that will generate an ordered sequence $\langle a_i \rangle$ of the ascending fitness values of $a_i$ amongst the set $\{a_i\}$ .
$\mathbf{Round}(a) \rightarrow a'$	Rounding calculator that will round a value $a$ to its nearest integer $a'$ .
$\mathbf{Dist}(a, b) \rightarrow D_{a,b}$	Distance calculator that will calculate the Euclidean distance by $D_{a,b} = \sqrt{(a - b)^2}$ .
$  $	Only denotes absolute value.

# Chapter 1

## Introduction

### 1.1 Motivation

The pace of designing optimization algorithms has never been stopped even since antiquity [1]. In the late stage of the 1900s, heuristic algorithms began to stand out with their impressive performance on more complicated modern optimization problems [26]. In these heuristic research fields, due to the great success of the simulated annealing optimization algorithm [26], the analogy connection between nature and optimization methods attracts increasing attention.

As the most successful analogy [26], the natural evolution process has inspired most metaphor-based optimization algorithms until now. From classical genetic algorithms, particle swarm optimization to modern 'animal'-inspired optimization algorithms, from pure algorithms to their various variants, from independent algorithms to combination algorithms, these algorithms that have come or are coming are welcome to solve problems and enrich the research community, however, one issue in this research field becomes increasingly obvious: **a great number of new nature-inspired optimization algorithms are just old heuristic algorithms in new clothes, which might already mislead the development of heuristic algorithms** [11, 12, 24, 26].

This problem frequently appears in population-based algorithms that are also well-known as swarm-based algorithms. One paper mentioned that these new modern nature-inspired algorithms are actually similar to swarm intelligence [11], in which they theoretically displayed the common components between the swarm intelligence and the nature-inspired algorithm. Some researchers also mentioned that several core components comprised

most meta-heuristic algorithms [12], in which they provided a generalized flow chart to display these common components. Moreover, some papers pointed out that a unified representation for organizing these algorithms together, especially a mathematical representation, will help prevent a worse development of heuristic algorithms [27, 33].

**Therefore, in this work, we aim to propose a unified framework that can combine swarm-based optimization algorithms together. This work will start with several selected swarm-based optimization algorithms.** Before starting this work, we want to clarify:

- *Variants of algorithms are not considered in this work.*  
Variants are important improvements to the algorithms, however, these variants will be ignored when building up the unified framework in this work. We will only focus on the base algorithms.
- *Hybrid algorithms are not considered in this work.*  
We must accept that it is very challenging to discuss independent and hybrid algorithms together. Therefore, we will start with independent algorithms and ignore hybrid algorithms here.

To build up a unified framework, we get inspired by gene terminologies and operations in the Evolutionary algorithm group. Therefore, in this work, we firstly study the possibility of unifying different terminologies in selected algorithms (see [Chapter 2](#)), which is the same as the gene terminology that is the only way to express evolutionary algorithms. Next, we study the possibility of unifying different procedures in these selected algorithms (see [Chapter 3](#)), in which we discuss the possibility that these algorithms have the same model structure. Finally, we give our unified framework for several selected algorithms (see [Chapter 4](#)), in which selected algorithms can be re-written in this unified framework without losing their quality of optimization.

Furthermore, we will also prove the feasibility of this unified framework by benchmark experiments, and meanwhile compare the performance of these selected algorithms in our unified framework (see [Chapter 5](#), [Chapter 6](#)). Lastly, we conclude our observations, point out our limitations in this work, and give future research suggestions (see [Chapter 7](#)).

## 1.2 Overview

As [Figure 1.1](#) displayed, since unifying the entire swarm-based algorithms is impossible in one work, we will start with working on seven examples, including new modern and old classical examples, after proving the feasibility of the unified framework we will build, we



will discuss the possibility that this framework is also able to be used in other swarm-based algorithms.

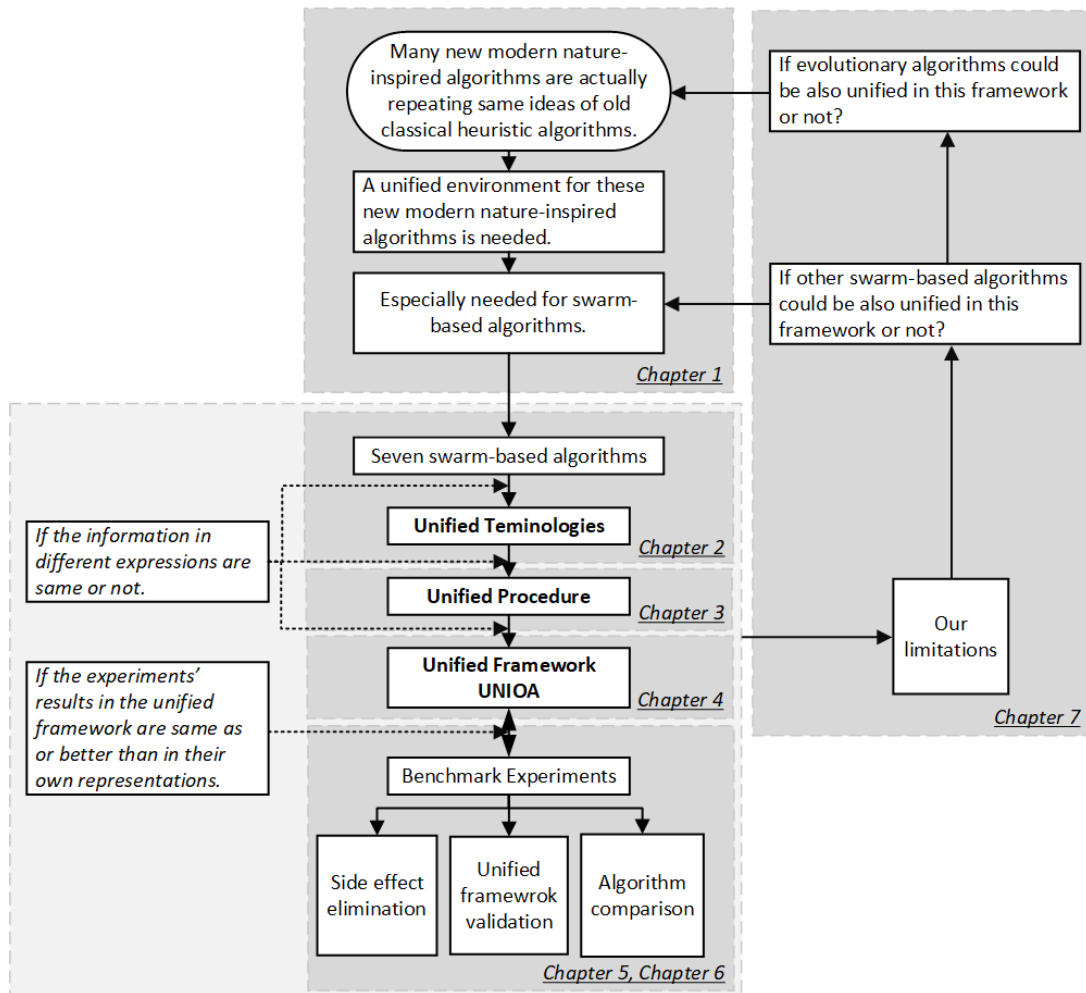


Figure 1.1: Overview.

These seven examples will be selected from the Bestiary of evolutionary computation<sup>1</sup>. They are :

- New modern swarm-based algorithms:  
Bat-inspired algorithm [32], Grasshopper optimization algorithm [25], Crow search

<sup>1</sup><https://github.com/fcampelo/EC-Bestiary>

algorithm [2], Moth-flame optimization algorithm, Monarch butterfly optimization algorithm [29], Butterfly optimization algorithm [3].

- Old classical swarm-based algorithms:  
Particle swarm optimization [17]

We hope our unified framework is unified enough to cover both traditional and modern nature-inspired algorithms.

# Chapter 2

## Unified Terminologies

### 2.1 Motivation

**This chapter discusses the commonalities among seven selected swarm-based algorithms on the level of their components.** Specifically, in these algorithms, we find that various terminologies of components with their meanings are able to be categorized into same terminologies with same meanings.

Firstly, we want to introduce several common symbols in swarm-based algorithms. In the swarm-based algorithm research field<sup>1</sup>, the given objective optimization problem is formulated as a math formula  $f(\vec{x})$  in which all decision variables are represented as a vector  $\vec{x}$ . Moreover, the number of decision variables<sup>2</sup> is usually defined as the dimension of search space to the objective optimization problem. When seeking the optimal solution to the optimization problem, the swarm-based algorithm will first initialize an arbitrary set of  $\vec{x}$  and then optimize this set by using iterative strategies. Same as in most studies, we also denote  $\vec{x}$  as  $\mathbf{x}$  and call it the possible objective solution. Inspired by this representation, we define that any term  $\mathbf{a}$  in the bold style denotes that this item is also a vector with the same dimension as the objective solution  $\mathbf{x}$ . Furthermore, we want to clarify that if the bold style item  $\mathbf{a}$  has a subscript  $i$ , it means this item has a one-to-one correspondence with each  $\mathbf{x}_i$ . However, one item without this subscript will act on the whole set of  $\mathbf{x}_i$ , whether this item is in bold style or not in bold style.

---

<sup>1</sup>In our work, we only consider single-objective continuous problem optimization.

<sup>2</sup>The number of decision variables is also the number of independent elements in  $\vec{x}$ .

Moreover, usually, the set of  $\mathbf{x}_i$  is called a population, and the number of  $\mathbf{x}_i$  in this set is called the size of this population. Each  $\mathbf{x}_i$  is a possible optimal solution to the target optimization problem  $f()$ , therefore the value of  $f(\mathbf{x}_i)$  is called the fitness value of  $\mathbf{x}_i$ .

In the following sections, Section 2.2 introduces these seven algorithms in detail but doesn't include how natures are simulated to create mathematical models. Readers can find such information in their published papers. In this work, the description of each algorithm is straightforward, only including the final model and its default hyper-parameter settings. Meanwhile, we also summarize their special points. Section 2.3 will give and discuss our unified terminologies based on these seven selected algorithms. A summary about this chapter is displayed in Section 2.4.

## 2.2 Specific Swarm-based Algorithms

### 2.2.1 Bat-inspired Algorithm (BA)

**Description** When simulating the echolocation behavior of micro-bats, for simplicity, the Bat-inspired algorithm only considered the velocity  $\mathbf{v}_i$  of each bat  $\mathbf{x}_i$ , the echo frequency  $freq$  emitted by each bat  $\mathbf{x}_i$ , the rate  $r$  of echo pulse, and the loudness  $A$  of echo [32]. The algorithm described here follows the published implementation code<sup>3</sup>, and explanations with math formulas are from their published paper [32].

This algorithm starts from is (1) initializing a bat  $\mathbf{x}_i$  population with size  $n = 20$  in the  $d$  dimension environment and the initialization method is Eq.2.1 in which  $Lb/Ub$  is the lower/upper boundary of each element in  $\mathbf{x}_i$ . Meanwhile, (2) velocity  $\mathbf{v}_i$  of each bat  $\mathbf{x}_i$  is initialized as Eq.2.2.

$$\mathbf{x}_i(t=0) = \mathcal{U}(Lb, Ub), i = 1, 2, \dots, n \quad 2.1$$

$$\mathbf{v}_i(t=0) = \mathcal{U}(0, 0), i = 1, 2, \dots, n \quad 2.2$$

After (3) evaluating fitness of each  $\mathbf{x}_i$ , it (4) goes to find the best individual  $\mathbf{x}_*$  that is the best  $\mathbf{x}_i$  the whole population has found so far. Next is (5) the iterative optimization process in which the maximum number of iterations is  $t_{max} = 1000$ . Under the iterative process, it firstly (6) updates  $A$  as Eq.2.3 in which  $A^0 = 1$  is its initial value, and  $\alpha = 0.97$  is a decreasing hyper-parameter. Next, (7)  $r$  is updated as Eq.2.4 in which  $r_0 = 1$  is the initial value of pulse rate and  $\gamma = 0.1$  is a decreasing hyper-parameter. In this algorithm,

<sup>3</sup>[https://uk.mathworks.com/matlabcentral/fileexchange/74768-the-standard-bat-algorithm-ba?s\\_tid=prof\\_contriblnk](https://uk.mathworks.com/matlabcentral/fileexchange/74768-the-standard-bat-algorithm-ba?s_tid=prof_contriblnk)

$t$  is the iteration counter.

$$\begin{cases} A(t) = \alpha \times A^0 & , \quad t = 0 \\ A(t+1) = \alpha \times A(t) & , \quad \text{o.w} \end{cases} \quad 2.3$$

$$r(t) = r_0 \times (1 - e^{-\gamma \times t}) \quad 2.4$$

Next is (8) updating each  $\mathbf{x}_i$  as Eq.2.5 in which  $freq\_min = 0/freq\_max = 2$  is the lower/upper boundary of  $freq_i$ .

$$\begin{aligned} freq_i &= \mathcal{U}(freq\_min, freq\_max) \\ \mathbf{v}_i(t+1) &= \mathbf{v}_i(t) + (\mathbf{x}_i(t) - \mathbf{x}_*) \times freq_i \\ \mathbf{x}_i(t+1) &= \mathbf{x}_i(t) + \mathbf{v}_i(t+1) \end{aligned} \quad 2.5$$

If (9)  $rand < r(t)$ ,  $\mathbf{x}_i$  will be updated differently as Eq.2.6 in which  $\epsilon = 0.1$ .

$$\mathbf{x}_i(t+1) = \mathbf{x}_*(t) + \epsilon \times \mathbf{rand} \times A(t+1) \quad 2.6$$

After (10) fixing outliers in  $\mathbf{x}_i(t+1)$  by replacing them with the boundary value, the BA model evaluates fitness of each  $\mathbf{x}_i(t+1)$  again. Then (11) if the fitness of  $\mathbf{x}_i(t+1)$  is better than  $\mathbf{x}_i(t)$  or the situation is  $rand > A$ ,  $\mathbf{x}_i(t+1)$  will be accepted. Lastly (12) the  $\mathbf{x}_*$  is updated.

Until now, one optimization round (from step (6) to step(12)) is finished, and it will iteratively execute until the stop condition meets.

**Summary** When modeling a Bat algorithm, there are two main processes: Initialization and Optimization. In the initialization process, besides initializing an initial population and all default parameters (including hyper-parameters and initial values), an initial velocity population is also initialized. Meanwhile, the initialization process also calculates the initial best bat. In the optimization process, two dynamically changing parameters are updated before updating the population. It is worth mentioning that the whole process of updating the population is asynchronous, including asynchronously evaluating the fitness of each bat, asynchronously judging whether the updated bat is improved, and asynchronously updating the current best bat.<sup>4</sup>

Moreover, we want to clarify that in their implemented codes, it is unreasonable to compare the previous best bat with a current unaccepted bat when updating the current

<sup>4</sup>The asynchronism here means the next individual will update only after the previous individual finishes updating, evaluating and all other optimization steps.

best bat. The definition of the current best bat shall be finding the best bat in the current population. The unaccepted bat is not in the current population anymore. Therefore, it is more reasonable to calculate the current best bat by comparing the previous best bat with the current bat who is at this same position in the current population.

Furthermore, we want to point out that in the Initialization process, the evaluation of the population happens immediately after initializing the initial population. However, in the Optimization process, the evaluation happens immediately after generating a temporary population. In other words, it happens in the middle of obtaining the temporary population and judging if accepting the generated individuals as the updated individuals, which means the evaluation does not happen on the final new updated population, but on a temporary updated population.

### 2.2.2 Grasshopper Optimization Algorithm (GOA)

**Description** The Grasshopper Optimization Algorithm was proposed in 2017 [25]. They designed this algorithm by studying grasshoppers' main swarming behaviors in larval and adulthood phases, such as slow movement and small steps in the larval phase, abrupt movement and long-range steps in the adulthood phase, and food source seeking in both phases. Their final model described here is from their published code <sup>5</sup>, and explanations with math formulas are from their published paper [25].

This model starts at (1) initializing grasshoppers  $\mathbf{x}_i$  with size  $N = 100$  in the  $d$  dimension environment. The initialization method in their published code can be simply re-framed as Eq.2.7 in which *up/down* is the upper/down boundary of  $\mathbf{x}_i$  <sup>6</sup>.

$$\mathbf{x}_i = \mathcal{U}(\text{down}, \text{up}), i = 1, 2, \dots, N \quad 2.7$$

Next is (2) calculating the fitness of each individual  $\mathbf{x}_i$  and (3) finding the best individual  $\mathbf{T}$  that the whole population has found so far. The followings are iterative optimizing process in which the stop condition is current number of iteration  $l$  is smaller than the maximum number of iterations  $\text{max\_iteration} = 100$  <sup>7</sup>. Under the iterative process, (4) the hyper-parameter  $c$  is firstly updated as Eq.2.8 in which  $cMin = 0.00004/cMax = 1$

<sup>5</sup><https://seyedalimirjalili.com/goa>

<sup>6</sup>Their published implementation code considered two boundary cases: if elements in  $\mathbf{x}_i$  share the same boundary or if elements in  $\mathbf{x}_i$  have different boundaries. According to the further experiments in the Chapter 5, we only consider the first case.

<sup>7</sup>Their published code thought the current iteration  $l = 2$  when starting optimizing, however, we define the current iteration  $l = 1$  when starting optimization in order to be same as other algorithms.

is the lower/upper boundary of  $c$ .

$$c = cMax - l \times \frac{cMax - cMin}{max\_iteration} \quad 2.8$$

Then, the GOA model (5) normalizes the distances  $\tilde{d}$  between individuals into  $[1, 4]$ . The normalization method in their published code is Eq.2.9 in which  $d$  is the Euclidean distance<sup>8</sup> between individual pairs.

$$\tilde{d} = 2 + d \bmod 2 \quad 2.9$$

Next is (6) updating  $\mathbf{x}_i$  as Eq.2.10 that is a simplified version of their published original formula. Because we only consider one case that elements in  $\mathbf{x}_i$  share the same boundary, we replace  $ub_d/lb_d$  by  $ub/lb$ . We also replace  $x_i^d$  by  $\mathbf{x}_i$  and  $\hat{T}_d$  by  $\mathbf{T}$  in the view of vector<sup>9</sup>.

$$\mathbf{x}_i = c \times \left( \sum_{j=1, j \neq i}^N c \times \frac{ub - lb}{2} \times S(\tilde{d}_{i,j}) \times \frac{\mathbf{x}_j - \mathbf{x}_i}{d_{ij}} \right) + \mathbf{T} \quad 2.10$$

The method of updating Eq.2.10 has a function named  $S$  in which  $\tilde{d}_{i,j}$  is the input. The  $S$  function is formulated as Eq.2.11 in which  $f = 0.5$  and  $l = 1.5$  are two hyper-parameters<sup>10</sup>.

$$S(\tilde{d}_{i,j}) = f e^{\frac{-\tilde{d}_{i,j}}{l}} - e^{-\tilde{d}_{i,j}} \quad 2.11$$

After updating each  $\mathbf{x}_i$ , (7) outliers in  $\mathbf{x}_i$  are replaced by the boundary values before (8) calculating its fitness. The last step is (9) updating the best individual  $\mathbf{T}$ .

Until now, one optimization round (from step(4) to step(9)) is finished, and it will iteratively execute until the stop condition meets.

**Summary** When modeling a GOA algorithm, there are two main processes: Initialization process and Optimization process. In the Initialization process, the algorithm has only two main tasks: initializing the initial population and finding the global best individual in the initial population. In the Optimization process, the update strategy depends on neighbors' positions.

### 2.2.3 Crow Search Algorithm (CSA)

**Description** The crow search algorithm was designed [2] in 2016. The core metaphor is crows' behaviour of hiding excess food and retrieving this stored food when needed.

<sup>8</sup>It refers to their faster version code.

<sup>9</sup>Their published paper didn't point out whether  $\hat{\cdot}$  is a special operation or not.

<sup>10</sup>Please note, here the  $l$  is not the current iteration mentioned above.

The author believes that these behaviours, including stealing others' food by tailing after them and fighting with theft by moving to another place instead of their real hiding place, are similar to an optimization process [2]. Their final algorithm model described here is from their published implementation code <sup>11</sup>, and explanations with math formulas are from their published paper [2].

The algorithm starts at (1) initializing the crow population of size  $N = 20$  in the  $d$  dimension environment. The initialization method they used in their implementation code can be formulated as Eq.2.12 in which  $l/u$  is the lower/upper boundary.

$$\mathbf{x}_i = \mathcal{U}(l, u), i = 1, 2, \dots, N \quad 2.12$$

The next is (2) calculating the fitness of each individual  $\mathbf{x}_i$ , and (3) initializing the memory  $\mathbf{m}_i$  for each individual  $\mathbf{x}_i$  as Eq.2.13.

$$\mathbf{m}_i = \mathbf{x}_i \quad 2.13$$

The following steps are about the iterative optimization process in which the stop condition is the maximum number of iterations  $tmax = 5000$ . Under the iterative process, (4) each individual  $\mathbf{x}_i$  is updated as Eq.2.14 in which  $fl = 2$  and  $AP = 0.1$  are two hyper-parameters.

$$\mathbf{x}_i(t+1) = \begin{cases} \mathbf{x}_i(t) + rand \times fl \times (\mathbf{m}_j(t) - \mathbf{x}_i(t)) & , \quad rand > AP \\ \mathcal{U}(u, l) & , \quad o.w \end{cases} \quad 2.14$$

After (5) calculating the fitness of the new individual  $\mathbf{x}_i(t+1)$ , if (6) the new individual  $\mathbf{x}_i(t+1)$  is in the case of  $l \leq \mathbf{x}_{i,d}(t+1) \leq u$ , the new individual  $\mathbf{x}_i(t+1)$  will be accepted, if not, the individual  $\mathbf{x}_i(t)$  will keep unchanged and reject to update. Lastly, (7) the memory  $\mathbf{m}_i(t+1)$  will be updated as Eq.2.15.

$$\mathbf{m}_i(t+1) = \begin{cases} \mathbf{x}_i(t+1) & , \quad f(\mathbf{x}_i(t+1)) < f(\mathbf{m}_i(t)) \\ \mathbf{m}_i(t) & , \quad o.w \end{cases} \quad 2.15$$

Until now, one optimization round (step (4) to step (7)) is finished, and it will iteratively execute until the stop condition meets.

---

<sup>11</sup><https://nl.mathworks.com/matlabcentral/fileexchange/56127-crow-search-algorithm>



**Summary** When modeling a CSA algorithm, there are two main processes: Initialization process and Optimization process. In the Initialization process, the algorithm has only two main tasks: initializing the initial population and the initial memory population that is the same as the initial population. In the Optimization process, the memory position of neighbors determines how far the individual moves.

When meeting outliers, the CSA algorithm chooses to give up the new position. In other words, as far as there is an outlier, the individual rejects to be updated.

### 2.2.4 Moth-flame Optimization Algorithm (MFO)

**Description** In 2015, the Moth-flame optimization algorithm simulated the process that moths will fly to the center of artificial light (that is also seen as flames) in the spiral path [22]. The core characteristic in MFO is that the author defined the positions of flames as the optimization orientation of moths, and each moth has its own dynamically changing flame. The final algorithm model described here is from their published implementation code <sup>12</sup>, and explanations with math formulas are from their published paper [22].

This algorithm starts from (1)  $N = 30$  initial moths population  $Moth\_pos$  in a  $dim$  dimension environment. The initialization method in their published code can be formulated as Eq.2.16 in which  $\mathbf{m}_i$  is each moth and  $ub/lb$  is the upper/lower boundary of elements in each moth <sup>13</sup>.

$$\mathbf{m}_i = \mathcal{U}(lb, ub), i = 1, 2, \dots, N \quad 2.16$$

Then the algorithm model starts iteratively optimizing each  $\mathbf{m}_i$  in which the stop condition is there is a maximum number of iterations  $T$ . Under the iterative optimization process, first of all, (2) the number of flames is calculated as Eq.2.17 in which  $l$  is the current number of iteration. The  $flame\_no$  will impact the updating positions of moths in the following steps.

$$flame\_no = \mathbf{Round}(N - l \times \frac{N - 1}{T}) \quad 2.17$$

After (3) replacing outliers in  $\mathbf{x}_i$  by the boundary values, fitness of each moth is calculated. Next is (4) sorting the combination of the current population and the previous population <sup>14</sup> based on the ascending of their fitness, and the first  $N = 30$  moths in the combination

<sup>12</sup>[https://nl.mathworks.com/matlabcentral/fileexchange/52270-moth-flame-optimization-mfo-algorithm-toolbox?s\\_tid=srchtitle](https://nl.mathworks.com/matlabcentral/fileexchange/52270-moth-flame-optimization-mfo-algorithm-toolbox?s_tid=srchtitle)

<sup>13</sup>In this work, we only consider that the boundaries of all elements in each moth are same, although their published code also provided a method to deal with different boundaries.

<sup>14</sup>The previous population is  $\emptyset$  (there is no previous population), when the current population is the initial population.

population are defined as *sorted\_population*. Then, a linearly decreasing parameter  $a$  is calculated by Eq.2.18 in which  $l$  is the current number of iteration.

$$a = -1 + l \times \frac{-1}{T} \quad 2.18$$

Next, (5) each element  $j$  of each moth  $\mathbf{m}_i$  will be updated as follows: if  $i \leq \text{flame\_no}$ , Eq.2.19; if not, Eq.2.20 in which  $\mathbf{m}'_{i,j}$  denotes each newly updated element  $j$  of each moth  $\mathbf{m}_i$ ;  $b = 1$  is a constant parameter used to define the shape of spiral [22];  $t = (a - 1) \times \text{rand} + 1$  is a random number related to  $a$ .

$$\begin{aligned} \text{distance\_to\_flame} &= |\text{sorted\_population}(i, j) - \mathbf{m}_{i,j}| \\ \mathbf{m}'_{i,j} &= \text{distance\_to\_flame} \times e^{b \times t} \times \cos(t \times 2 \times \pi) + \text{sorted\_population}(i, j) \end{aligned} \quad 2.19$$

$$\begin{aligned} \text{distance\_to\_flame} &= |\text{sorted\_population}(i, j) - \mathbf{m}_{i,j}| \\ \mathbf{m}'_{i,j} &= \text{distance\_to\_flame} \times e^{b \times t} \times \cos(t \times 2 \times \pi) + \text{sorted\_population}(\text{flame\_no}, j) \end{aligned} \quad 2.20$$

Until now, one optimization round (from step(2) to step(5)) is finished, and it will iteratively execute until the stop condition meets.

**Summary** When modeling a Moth-flame algorithm, there are two processes: Initialization process and Optimization process. In the initialization process, this algorithm only initializes an initial population with all default parameters. In the optimization process, the main convergency mechanism is to keep moving to better solutions by narrowing better choices down. Here, the better choices are selected from the top several best solutions, and the method of narrowing down is to narrow the number of best solutions. The main convergency mechanism happens at Eq.2.20.

Compared with other algorithms, this algorithm has another two different places. The first one is that there is no evaluation operator in the initialization process. The second one is that fixing outliers and evaluating are happening at the beginning of the optimization process.

### 2.2.5 Monarch Butterfly Optimization (MBO)

**Description** The Monarch Butterfly Optimization simulated the migration behavior of monarch butterflies in 2015 [29]. The core metaphor is that monarch butterflies in two different habitats evolve differently. These two habitats are mimicked by dividing the whole population into two subpopulations. The author defined that the whole population

will firstly be sorted based on their fitness before dividing them into two with a ratio [29]. Then, each habitat experiences its particular evolution approach (migration operator or adjusting operator [29]). We describe their final algorithm model implemented in their published code <sup>15</sup>, and explanations with math expressions from their published paper [29].

This algorithm starts from (1) an initial population *Population* with size  $M = 50$ , and the initialization method <sup>16</sup> in their published code can be formulated as Eq.2.21 in which *MinParValue*/*MaxParValue* is the lower/upper boundary of elements in  $\mathbf{x}_i$ .

$$\mathbf{x}_i = \mathcal{U}(\text{MinParValue}, \text{MaxParValue}), i = 1, 2, \dots, M \quad 2.21$$

After (2) evaluating each  $\mathbf{x}_i$ , the entire population is sorted from the best to the worst. Then (3) it starts iteratively optimizing each  $\mathbf{x}_i$  in which the stop condition is the maximum number of iterations  $\text{Maxgen} = 50$ . Under the iterative process, (4)  $\text{Keep} = 2$  best  $\mathbf{x}_i$  are first kept in *chromKeep* <sup>17</sup>. Next is (5) dividing the entire *Population* into two sub-populations with a ratio  $\text{partition} = \frac{5}{12}$ . The two sub-populations (*Population1*, *Population2*) will be updated by two operators ('Migration operator', 'Adjusting operator') respectively. (6) In the *Population1*, if  $\text{rand} \times \text{period} \leq \text{partition}$  in which  $\text{period} = 1.2$ ,  $\mathbf{x}_{i,k}^{t+1} = \mathbf{x}_{\text{population1},k}^t$ ; else  $\mathbf{x}_{i,k}^{t+1} = \mathbf{x}_{\text{population2},k}^t$ , in which  $k$  is the  $k$ -th element of  $\mathbf{x}_i$ , and *population1*/*population2* is one arbitrary moth from previous *Population1*/*Population2*. (7) Next is fixing outliers in this newly generated *Population1* by replacing them with the boundary value, before evaluating this newly generated *Population1*.

(8) In *Population2*, if  $\text{rand} \leq \text{partition}$ ,  $\mathbf{x}_{i,k}^{t+1} = \mathbf{x}_{\text{best},k}^t$  in which  $k$  is the  $k$ -th element and *best* is the best moth the entire population has found so far; else  $\mathbf{x}_{i,k}^{t+1} = \mathbf{x}_{\text{population2},k}^t$ , in which  $k$  is the  $k$ -th element, and *population2* is one arbitrary moth from previous *Population2*, furthermore, under this situation, if  $\text{rand} > \text{BAR}$  in which  $\text{BAR} = \frac{5}{12}$ ,  $\mathbf{x}_{i,k}^{t+1}$  will be further updated as Eq.2.22 in which  $\alpha$  is a weighting factor as Eq.2.23, and  $d\mathbf{x}$  is a walk step as Eq.2.24. (9) Then, outliers in the newly generated *Population2* are fixed by replacing them with the boundary value, and next is evaluating this newly generated *Population2*.

$$\mathbf{x}_{i,k}^{t+1} = \mathbf{x}_{i,k}^{t+1} + \alpha \times (d\mathbf{x}_k - 0.5) \quad 2.22$$

<sup>15</sup>[https://nl.mathworks.com/matlabcentral/fileexchange/101400-monarch-butterfly-optimization-mbo?s\\_tid=srchtitle](https://nl.mathworks.com/matlabcentral/fileexchange/101400-monarch-butterfly-optimization-mbo?s_tid=srchtitle)

<sup>16</sup>The original paper considered the situation if boundaries of elements are same or not, but in this work, we only consider the situation that boundaries of elements are same

<sup>17</sup>This elitism strategy stated in their published code wasn't stated in their publish paper.

$$\alpha = \frac{S_{max}}{t^2} \text{ where } S_{max} = 1 \text{ and } t \text{ is the current iteration.} \quad 2.23$$

$$dx = \text{Levy}(x_j^t) \quad 2.24$$

where **Levy** function in their published code can be formulated as Eq.2.25.

$$\begin{aligned} Stepsize &\sim Exp(2 \times Maxgen) \\ dx_k &= \underbrace{\sum (\tan(\pi \times rand), \dots, \tan(\pi \times rand))}_{Stepsize} \end{aligned} \quad 2.25$$

Then (10) these two newly generated sub-populations are combined into one population. Next, (11) after the last *Keep* = 2 worst  $x_i$  will be replaced by *chromKeep* that is mentioned in step (4), (12) the final new population will be sorted again.

Until now, one optimization round (from step (4) to step(12)) is finished, and it will iteratively execute until the stop condition meets.

**Summary** When modeling an MBO algorithm, there are two main processes: Initialization and Optimization. In the initialization process, besides initializing an initial population and all default parameters, the initial population is also sorted from best to worst. In the optimization process, the core idea is that the whole population is divided into two sub-populations, and they are updated differently. One sub-population is better than another one because of the previous sort operator. The elitism strategy is also a highlight of this model. It happens at the beginning and the end of the optimization process.

Furthermore, although in the initialization process, the evaluation of the population only happens once immediately after initializing the initial population, there are two evaluations in the optimization process. There are two evaluations because the whole population is updated in two methods for two sub-populations.

### 2.2.6 Butterfly Optimization Algorithm (BOA)

**Description** The Butterfly Optimization Algorithm was modeled in 2018 [3]. Butterflies' behaviors of searching for food and mates are mimicked to solve optimization problems. The algorithm model described here is from their published code <sup>18</sup>, and explanations with math formulas are from their published paper [3].

---

<sup>18</sup><https://www.mathworks.com/matlabcentral/mlc-downloads/downloads/b4a529ac-c709-4752-8ae1-1d172b8968fc/67a434dc-8224-4f4e-a835-bc92c4630a73/previews/BOA.m/index.html>

In this algorithm, they define a  $dim$  dimensional environment in which each individual is  $\mathbf{x}_i$  and the fitness of  $\mathbf{x}_i$  is  $f(\mathbf{x}_i)$ . (1) The algorithm first initializes a population  $n = 50$  of  $\mathbf{x}_i$  by Eq.2.26 in which  $Lb/Ub$  is lower/upper bound <sup>19</sup>.

$$\mathbf{x}_i = \mathcal{U}(Lb, Ub), i = 1, 2, \dots, n \quad 2.26$$

It also (2) setups three parameters: probability switch  $p = 0.8$ ,  $power\_exponent = 0.1$  and  $sensory\_modality = 0.01$ . After evaluating each  $\mathbf{x}_i$ , it (3) calculates the best individual  $g^*$  that the whole population had found so far.

Next steps are the iterative optimization process in which the stop condition is the current iteration  $t$  is smaller than the maximum number of iterations  $N\_iter$ . Under the iterative process, for each  $\mathbf{x}_i$ , (4) its own fragrance  $FP_i$  is first calculated as Eq.2.27 in which  $f(\mathbf{x}_i)$  is the fitness value of  $\mathbf{x}_i$  <sup>20</sup>.

$$FP_i(t) = sensory\_modality \times f(\mathbf{x}_i(t))^{power\_exponent} \quad 2.27$$

Then (5) if  $rand > p$ , the  $\mathbf{x}_i$  will be updated as Eq.2.28, (6) if not, the  $\mathbf{x}_i$  will be updated as Eq.2.29 in which  $\mathbf{x}_j$  and  $\mathbf{x}_k$  are two arbitrary neighbors around  $\mathbf{x}_i$ .

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + FP_i(t) \times (rand^2 \times g^*(t) - \mathbf{x}_i(t)) \quad 2.28$$

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + FP_i(t) \times (rand^2 \times \mathbf{x}_j(t) - \mathbf{x}_k(t)) \quad 2.29$$

Then after (7) replacing the outliers by the boundary values, (8) only if the updated  $\mathbf{x}_i(t+1)$  is better than the previous  $\mathbf{x}_i(t)$ , the update  $\mathbf{x}_i(t+1)$  will be delivered into the next optimization round. Next steps are (9) updating  $g^*$  and (10) updating  $sensory\_modality$  using Eq.2.30.

$$sensory\_modality(t+1) = sensory\_modality(t) + \frac{0.025}{sensory\_modality(t) \times N\_iter} \quad 2.30$$

Until now, one optimization round (step(4) to step(10)) is finished, it will iteratively execute until the stop condition meets.

---

<sup>19</sup>The initialization method isn't shown in their published paper or code, but according to their published code that only said there was an 'initialization' function receiving  $n$ ,  $dim$ ,  $Ub$  and  $Lb$  as inputs. We can reasonably assume their initialization method is Eq.2.26.

<sup>20</sup>Their published code calculated the fitness twice here. However, the published paper did not mention whether this is a necessary operator. Therefore, we ignore this duplicated evaluation when studying this algorithm.

**Summary** When modeling a BOA algorithm, there are two processes: Initialization process and Optimization process. In the Initialization process, there are four tasks: initialize the initial population and evaluate them, find the global best one individual, and set up several hyper-parameters. In the optimization process, the most interesting part is the fitness joins in optimizing the population. Moreover, the square value of the random number is also a special place.

Meanwhile, the original implemented code also has the same issue as the BA algorithm (see [Subsection 2.2.1](#)) when updating the global best one individual. We think it is unreasonable to compare the previous best individual with an unaccepted one when updating the current global best individual. The definition of the current best global individual shall be the best one that the current population has found so far. Here, the current population shall be the population that has been accepted and will be delivered into the next optimization round. Therefore, it is more reasonable to calculate the current global best one by comparing the previous best one with the current accepted best one at this same position in the current population.

## 2.2.7 Particle Swarm Optimization (PSO)

**Description** Particle Swarm Optimization was firstly introduced in 1995 [17]. The main idea comes from the movements of the animal swarm. By simulating their behaviors that the swarm dynamically moves to a 'roost' [17], the extremely simple algorithm, even with a small swarm size (15 to 30 agents), boasted impressive performance on solving continuous optimization functions. In the very first paper, the position of each agent in its swarm was controlled by two velocities  $\mathbf{X}$  and  $\mathbf{Y}$ . Moreover, each agent can remember its best position and know the global best position in its swarm. Although the very first paper also provided the algorithm model with formulas, we preferred to reference clearer published pseudocode whose algorithm model is as same as the very first paper possible. The final PSO model described in this work is from a tutorial of this method [21], and the hyper-parameter settings are from an application of this method [5].

This algorithm starts (1) from an initial swarm  $\mathbf{x}_i$  with the size  $N = 25$ . The initialization method is formulated as [Eq.2.31](#) in which  $min/max$  is the lower/upper boundary of  $\mathbf{x}_i$ .

$$\mathbf{x}_i(t = 0) = \mathcal{U}(min, max), i = 1, 2, \dots, N \quad 2.31$$

Then (2) each velocity  $\mathbf{v}_i$  is initialized as [Eq.2.32](#) in which  $lb_{\mathbf{v}}/ub_{\mathbf{v}}$  is the boundary of velocity. Next (3) each  $pbest_i$  is calculated as [Eq.2.33](#).

$$\mathbf{v}_i = \mathcal{U}(lb_{\mathbf{v}}, ub_{\mathbf{v}}), i = 1, 2, \dots, N \quad 2.32$$

$$pbest_i = \mathbf{x}_i, i = 1, 2, \dots, N \quad 2.33$$

After (4) calculating the fitness of each agent  $\mathbf{x}_i$ , the algorithm obtains (5) the best agent  $gbest$  that is the whole swarm has found so far in the initial swarm. The following steps are about (6) iterative optimization process with a stopping condition. Under the iterative process, the published pseudo-code first updates each velocity  $\mathbf{v}_i$  by Eq.2.34 in which  $w = 0.73$  is an adjusting parameter,  $c_1 = 1.49$  and  $c_2 = 1.49$  is respectively the cognitive and social coefficient [21].

$$\mathbf{v}_i(t+1) = w \times \mathbf{v}_i(t) + c_1 \times rand \times (pbest_i - \mathbf{x}_i) + c_2 \times rand \times (gbest - \mathbf{x}_i) \quad 2.34$$

After updating (7) each agent as Eq.2.35, all elements in each new agent are checked if they are feasible. The last step in one optimization round is (8) updating  $pbest$  and (9)  $gbest$ .

$$\mathbf{x}_i(t+1) = \mathbf{x}_i(t) + \mathbf{v}_i(t+1) \quad 2.35$$

Until now, one optimization round (from step(6) to step(9)) is finished, and it will iteratively execute until the stop condition meets.

As a side note, the very first published paper did not directly point out the notation for search space dimension, the stop condition, and if outliers should be dealt with or not. However, as discussed in the Section 2.1, the dimension depends on the number of elements in  $\mathbf{x}_i$ , therefore, it is reasonable to accept that the PSO model also needs a concept of dimension. Meanwhile, the stop condition and the method to deal with the outliers also must exist. In this work, the method of fixing outliers is the most common way that replaces outliers by boundary values of  $\mathbf{x}_i$ .

**Summary** When modeling a PSO algorithm, there are two processes: Initialization process and Optimization process. In the Initialization process, four components need to be initialized: the initial population, their velocity population, their personal best position group, the global best one position in this population. In the Optimization process, individuals are optimized one by one, which means that after updating one individual, its fitness, its  $pbest$ , and the current  $gbest$  are then updated.

Moreover, in the Initialization process, the evaluation happens immediately after generating the initial population. In the Optimization process, the evaluation happens after generating a temporary population and before updating  $pbest$  and  $gbest$ .

## 2.3 Unified Terminologies

### 2.3.1 Unified Terminologies Catalog

Although these selected algorithms use different expressions, their core meanings are the same. For example, the memory position in CSA is the personal best position in PSO if we only consider the information carried by the memory position and the personal best position. Therefore, **the information carried by these various expressions is the only principle to categorize these original terminologies**. As shown in Table 2.1, we find **20 unified components can cover the entire components in these seven selected algorithms**. Specifically, only 20 kinds of information are used when modeling these seven algorithms. At the same time, **these 20 unified components can be further divided into two groups: compulsory components and selective components**, as shown in Figure 2.1.

In Table 2.1, **Information 1 to 5, 13, 15, 17-20 are compulsory components**, because they must exist when modeling a swarm-based algorithm. In other words, these components can theoretically make up a most basic swarm-based algorithm, if we don't consider whether the algorithm performance is excellent. Meanwhile, we consider the **Information 18 is also a compulsory component**, although not all algorithms use it in this work. Because firstly, selection process is an important process in natural evolution. Secondly, if one algorithm doesn't use selection method, it means that all individuals are delivered into the next generation round, which is also a kind of selection method. More details are as follows:

- *Information 1, 3, 4*: the objective function  $f()$  must exist to determine where the algorithm will happen. When the  $f()$  is determined, the dimension  $n$  and the boundary  $[lb_x, ub_x]$  are also determined which is because they exist in the  $f()$ .
- *Information 2, 5, 13, 15*: when searching possible solutions to  $f()$ , possible solutions  $x_i$  with the number  $M$  of  $x_i$  also must exist, because it determines the size of the search pool in which the algorithm could find the final optimal solution to  $f$ . Furthermore, the method  $Init_x$  to initialize the initial possible solutions also must exist, because it determines where the algorithm starts to find the final optimal solution, which means there also must be a method  $Opt_x$  that can iteratively optimize the initial possible solutions.
- *Information 17, 18, 19, 20*: because these swarm-based optimization algorithms are a kind of iterative optimization heuristic algorithms in which sampling and repetition play an important role in finding optimal solutions [7], the method  $S$  related to



sampling and the stop condition  $T$  must exist. Moreover, these algorithms will solve the problem in a limited search space, there must also be a method  $C$  to deal with outliers outside the search space.

Therefore, we conclude that  $f$ ,  $\mathbf{x}_i$ ,  $M$ ,  $T$  and methods  $Init_{\mathbf{x}}$ ,  $Opt_{\mathbf{x}}$ ,  $C$ ,  $S$  must exist in a swarm-based optimization algorithm. As shown in the left side of Figure 2.1, these compulsory components exist in all of seven selected algorithms.

In Table 2.1, **Information 6 to 12, 14, 16 are selective components**, because not every algorithm needs these information. We find these information plays a similar role in modeling swarm-based algorithms, and the role is that these information acts on compulsory components to improve the performance of the algorithm. **Considering the role of these information is to influence how far the newly generated individual will move, we define these selective components step-size with notation  $\Delta$** . Meanwhile, we consider the **Information 11 is also a selective component**, because its main role is also to the movement of individuals, although it exists in every algorithm. Furthermore, we further categorize these selective components as follows:

- **Information 11**: Static numerical  $w$ -relative step-size  $\Delta : w, z^0, [lb_{\mathbf{y}}, ub_{\mathbf{y}}]$ .  
Most static numerical step-size, such as  $w$ , is able to be understood as common hyper-parameters in most algorithms. They are numbers; are set before running algorithms and are unchanged during optimization. In Figure 2.1, we find there are two groups of static numerical step-size. The first group is  $w$  that is commonly seen as hyper-parameters. The second group is the initial value of dynamic step-size (see next item ▪) in which the dynamic numerical step-size  $z$  needs an initial value  $z^0$ , the  $\mathbf{y}$ -relative dynamic vector step-size  $\mathbf{y}_i$  needs an initial value  $[lb_{\mathbf{y}}, ub_{\mathbf{y}}]$ .
- **Information 6, 7, 8, 9, 12**: Dynamic step-size  $\Delta : z, \mathbf{x}_{i_p}, \mathbf{x}_g, \mathbf{x}_s, \mathbf{y}_i$ .
  - **Information 12**: Dynamic numerical  $z$ -relative step-size  $\Delta : z$ .  
This kind of step-size is a single numeric that is changing over the iteration  $t$  during the optimization process.
  - Dynamic vector step-size  $\Delta : \mathbf{x}_{i_p}, \mathbf{x}_g, \mathbf{x}_s, \mathbf{y}_i$ .  
This kind of step-size is a vector with the same dimension as  $\mathbf{x}$ .
  - \* **Information 6, 7, 8**:  $\mathbf{x}$ -relative step-size  $\Delta : \mathbf{x}_{i_p}, \mathbf{x}_g, \mathbf{x}_s$ .  
This kind of step-size can be converted from the target  $\mathbf{x}$ . Specifically, there could be a straightforward formula between  $\mathbf{x}$  and  $\mathbf{x}$ -relative step-size. For example, personal best position  $\mathbf{x}_{i_p}$ , global best position  $\mathbf{x}_g$ . The  $\mathbf{x}_s$  denotes a special  $\mathbf{x}$ -relative step-size related to  $\mathbf{x}$ , for example, the sorted population in MFO.

\* *Information 9*: y-relative step-size  $\Delta : \mathbf{y}_i$ .

This kind of step-size is not related to the target  $\mathbf{x}_i$ . In other words, it is impossible to represent this kind of step-size by translating other existing variables in a straightforward method. Normally, this kind of step-size has same size as the population of individuals. For example, velocity  $\mathbf{v}_i$ .

For other selective components *Information 10, 14, 16*, because the dynamic step-size  $\Delta : \mathbf{x}_{ip}, \mathbf{x}_g, \mathbf{x}_s, \mathbf{y}_i$  is changing during the optimization process, there must be an initialization status and a changing status. Therefore, we also define  $Init_\Delta$  and  $Opt_\Delta$  to achieve the initialization status and the changing status for dynamic step-size, in which the  $Init_\Delta$  method can also include the set-up of static  $w$ -relative step-size  $\Delta : w, z^0, [lb_y, ub_y]$ . As shown in the right side of [Figure 2.1](#), all of these step-size  $\Delta$  have an initial status determined by  $Init_\Delta$ , and each dynamic step-size  $\Delta : z, \mathbf{x}_{ip}, \mathbf{x}_g, \mathbf{x}_s, \mathbf{y}_i$  will change by an optimization operator  $Opt_\Delta$ .

### 2.3.2 Extension in Other Swarm-based Algorithms.

This section discusses the possibility of applying these unified terminologies in any other swarm-based algorithm. First of all, all of compulsory components (see *Information 1 to 5, 13, 15, 17-20* in [Table 2.1](#)) and several selective components (see *Information 6, 7, 11, 12* in [Table 2.1](#)) are easily detected in any one swarm-based algorithm, because they are straightforward. For example,  $\mathbf{x}$  is the animal,  $w$  is the hyper-parameter and  $z$  is also the hyper-parameter but  $z$  is changing. Mostly any one swarm-based algorithm will use very clear noun or terminology to point out these information.

The place where it is most likely to confuse readers is **how to detect the dynamic vector step-size: Why we think one vector step-size is a x-relative step-size not a y-relative step-size?** This question can be answered according to the observations found in these seven algorithms.

In most of swarm-based algorithms, besides the animal  $\mathbf{x}$ , another noun is also always existing. This noun could be the velocity (in BA, PSO), the memory (in CSA) or the flame (in MFO). We are not allowed to directly assign them the y-relative step-size, although this is the most straightforward way. **The reason is that the information carried by the noun is the only principle used to assign them different unified terminology name.** For example:

- The memory  $\mathbf{m}_i$  in CSA is initialized by  $\mathbf{x}_i$  itself as  $\mathbf{m}_i(t=0) = \mathbf{x}_i(t=0)$ , and then will be updated by  $\mathbf{m}_i(t+1) = \text{Min}\{\mathbf{m}_i(t), \mathbf{x}_i(t+1)\}$ . Therefore in the first iteration, the update will happen as [Eq.2.36](#). In other words, the  $\mathbf{m}_i$  always

can be represented by  $\mathbf{x}_i$  all the time. Moreover, the update method to the  $\mathbf{m}_i$  is same to find the  $\mathbf{x}_{ip}$ , therefore, it is totally safe to replace  $\mathbf{m}_i$  by x-relative vector  $\mathbf{x}_{ip}$  rather than the y-relative vector  $\mathbf{y}_i$ .

$$\begin{aligned}\mathbf{m}_i(t=1) &= \text{Min}\{\mathbf{m}_i(t=0), \mathbf{x}_i(t=1)\} \\ &= \text{Min}\{\mathbf{x}_i(t=0), \mathbf{x}_i(t=1)\} \\ &= \mathbf{x}_i(t=0) \text{ or } \mathbf{x}_i(t=1)\end{aligned}\tag{2.36}$$

- The flame  $\langle \text{flame}_i \rangle$  in MFO is initialized by  $\langle \mathbf{x}_i \rangle$  itself as  $\langle \text{flame}_i \rangle(t=0) = \langle \mathbf{x}_i \rangle(t=0)$ , and then will be updated by  $\langle \text{flame}_i \rangle(t+1) = \text{Sort}(\{\mathbf{x}_i(t)\} \cup \{\mathbf{x}_i(t+1)\})$ ,  $i = 1 \dots M$ . Therefore, in the first iteration, the update will happen as Eq.2.37. In other words, the  $\langle \text{flame}_i \rangle$  always can be represented by x-relative vectors  $\langle \mathbf{x}_i \rangle$  all the time, rather than the y-relative vectors  $\langle \mathbf{y}_i \rangle$ .

$$\begin{aligned}\langle \text{flame}_i \rangle(t=1) &= \text{Sort}(\langle \text{flame}_i \rangle(t=0) \cup \{\mathbf{x}_i(t=1)\}), i = 1 \dots M \\ &= \text{Sort}(\langle \mathbf{x}_i \rangle(t=0) \cup \{\mathbf{x}_i(t=1)\}), i = 1 \dots M\end{aligned}\tag{2.37}$$

- The velocity in BA and PSO uses a different way to initialize itself, such as  $\mathcal{U}(lb, ub)$ . Moreover, the most important point is there is no way to represent the velocity by x-relative vector  $\mathbf{x}_i$ . Therefore, we prefer assigning y-relative vector  $\mathbf{y}_i$  to this kind of step-size.

After resolving the confusion between x-relative vector step size and y-relative vector step size, other *Information 8 to 10, 14, 16* also become clear to detect in any other swarm-based algorithms according to the previous observations.

## 2.4 Summary

In this chapter, we discuss the commonalities among terminologies in seven selected algorithms. We summarize their commonalities in Table 2.1 in which we conclude 20 unified components that can cover entire various components in these algorithms. Meanwhile, we discuss the classification of these 20 general components in Figure 2.1.

In Subsection 2.3.1, we detail the classification of these 20 general components. Every algorithm must have the target problem  $f$ , the search space  $n$  and  $[lb_x, ub_x]$ , the possible target solution  $\mathbf{x}_i$ , the population size  $M$ , the stop condition  $T$ , the initialization method  $Init_x$  to  $\mathbf{x}_i$ , the optimization method  $Opt_x$  to  $\mathbf{x}_i$ , the method  $C$  to deal with outliers and the method  $S$  to select which individuals could be delivered into the next generation.

However, selective components are more flexible, and these selective components are all related to determine how far the newly generated individual will move. Therefore, we name selective components step-size  $\Delta$ . These selective components can be further categorized into static numerical  $w$ -relative step-size  $\Delta : w, z^0, [lb_y, ub_y]$  and dynamic step-size  $\Delta : z, \mathbf{x}_{i_p}, \mathbf{x}_g, \mathbf{x}_s, \mathbf{y}_i$ . What is interesting is the dynamic step-size also has two categories: dynamic numeric  $z$ -relative step-size  $\Delta : z$  and dynamic vector step-size  $\Delta : \mathbf{x}_{i_p}, \mathbf{x}_g, \mathbf{x}_s$  that are  $\mathbf{x}$ -relative step-size and  $\Delta : \mathbf{y}_i$  that is  $\mathbf{y}$ -relative step-size.

In [Subsection 2.3.2](#), we discuss how to extend these unified terminologies into any other swarm-based algorithm. Furthermore, we resolve a confusion of the difference between  $\mathbf{x}$ -relative step-size and  $\mathbf{y}$ -relative step-size.

In conclusion, we provide a catalog of unified terminologies for seven selected algorithms.

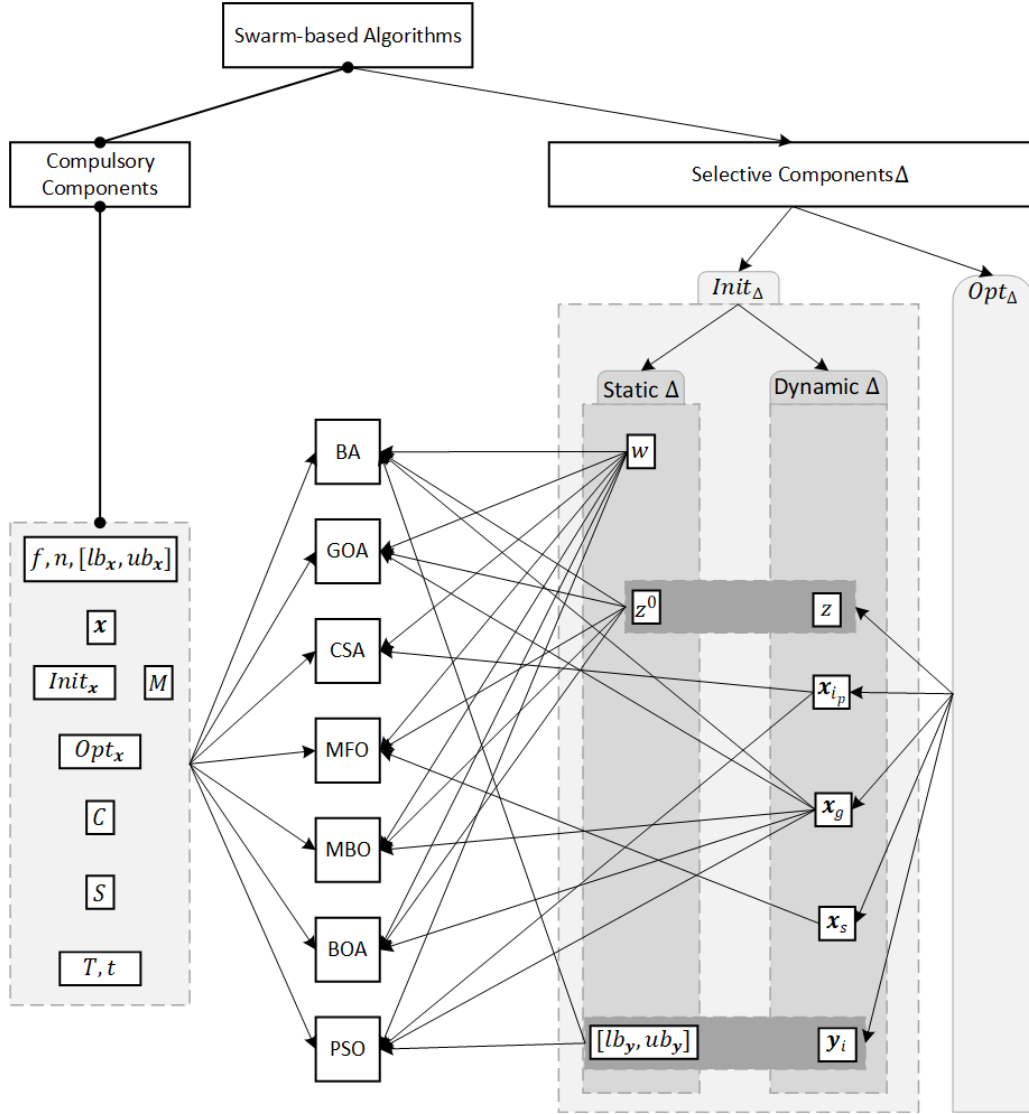


Figure 2.1: Entire compulsory components and some of selective components can comprise any of these seven algorithms.

Table 2.1: Unified Terminologies: all information appearing in these seven algorithms can be analyzed and then categorized into 20 components. The none in gray color means this information doesn't appear in this algorithm.

Information	Different Terminologies	Unified Terminology
1.The objective optimization problem.	The representations in seven algorithms are same.	$f()$ : optimization function.
2.One possible objective solution.	• <i>BA</i> : each bat $\mathbf{x}_i$ . • <i>GOA</i> : each grasshopper $\mathbf{x}_i$ . • <i>CSA</i> : crow $\mathbf{x}_i$ . • <i>MFO</i> : moth $\mathbf{m}_i$ . • <i>MBO</i> : monarch butterfly $\mathbf{x}_i$ . • <i>BOA</i> : butterfly $\mathbf{x}_i$ . • <i>PSO</i> : particle $\mathbf{x}_i$ .	$\mathbf{x}_i$ : one objective solution.
3.The number of independent elements in $\mathbf{x}_i$ .	• <i>BA</i> : $d$ . • <i>GOA</i> : $dim$ . • <i>CSA</i> : $d$ . • <i>MFO</i> : $dim$ . • <i>MBO</i> : $k$ . • <i>BOA</i> : $dim$ . • <i>PSO</i> : dimension.	$n$ : dimension of the $f()$ search space.
4.The boundary of each independent element in $\mathbf{x}_i$ <sup>21</sup> .	• <i>BA</i> : $Lb/Ub$ . • <i>GOA</i> : $down/up$ . • <i>CSA</i> : $l/u$ . • <i>MFO</i> : $lb/ub$ . • <i>MBO</i> : $MinParValue/MaxParValue$ . • <i>BOA</i> : $Lb/Ub$ . • <i>PSO</i> : $lb/ub$ .	$lb_{\mathbf{x}}/ub_{\mathbf{x}}$ : the lower/upper boundary of all elements in $\mathbf{x}_i$ .
5.The number of $\mathbf{x}_i$ .	• <i>BA</i> : $n$ . • <i>GOA</i> : $N$ . • <i>CSA</i> : $N$ . • <i>MFO</i> : $N$ . • <i>MBO</i> : $M$ . • <i>BOA</i> : $n$ . • <i>PSO</i> : $N$ .	$M$ : population size.
6.The best $\mathbf{x}_i$ that it itself has found so far. It will change during the optimization process.	• <i>BA</i> : none. • <i>BA</i> : none. • <i>GOA</i> : none. • <i>CSA</i> : memory $\mathbf{m}_i$ . • <i>MFO</i> : none. • <i>MBO</i> : none. • <i>BOA</i> : none. • <i>PSO</i> : $pbest$ .	$\mathbf{x}_{i_p}$ : x-relative step-size $\Delta$ .
7.The best $\mathbf{x}_i$ that the entire population has found so far. It will change during the optimization process.	• <i>BA</i> : $\mathbf{x}_*$ . • <i>GOA</i> : $\mathbf{T}$ . • <i>CSA</i> : none. • <i>MFO</i> : none. • <i>MBO</i> : $\mathbf{x}_{best}$ . • <i>BOA</i> : $g^*$ . • <i>PSO</i> : $gbest$ .	$\mathbf{x}_g$ : x-relative step-size $\Delta$ .

<sup>21</sup>Some algorithms, such as GOA, MFO, consider the boundary of every element is different, however, we only consider the boundary of every element is same in this work.

8.Special vector step-size that is able to be represented by $\mathbf{x}_i$ . It will change during the optimization process.	<ul style="list-style-type: none"> <li>•BA: none.</li> <li>•GOA: none.</li> <li>•CSA: none.</li> <li>•MFO: sorted population.</li> <li>•MBO: none.</li> <li>•BOA: none.</li> <li>•PSO: none.</li> </ul>	$\mathbf{x}_s$ : $\mathbf{x}$ -relative step size $\Delta$ .
9.Special vector step-size that is not able to be represented by $\mathbf{x}_i$ . It will change during the optimization process.	<ul style="list-style-type: none"> <li>•BA: velocity <math>\mathbf{v}_i</math>.</li> <li>•GOA: none.</li> <li>•CSA: none.</li> <li>•MFO: none.</li> <li>•MBO: none.</li> <li>•BOA: none.</li> <li>•PSO: velocity <math>\mathbf{v}_i</math>.</li> </ul>	$\mathbf{y}_i$ : $\mathbf{x}$ -relative step-size $\Delta$ .
10.The boundary of each independent element in $\mathbf{y}_i$ .	<ul style="list-style-type: none"> <li>•BA: Eq.2.2.</li> <li>•GOA: none.</li> <li>•CSA: none.</li> <li>•MFO: none.</li> <li>•MBO: none.</li> <li>•BOA: none.</li> <li>•PSO: Eq.2.32.</li> </ul>	$lb_{\mathbf{y}}/ub_{\mathbf{y}}$ : the lower/upper boundary of all elements in $\mathbf{y}_i$ .
11.step-size that is set before running algorithm. They are unchanged during the optimization process.	<ul style="list-style-type: none"> <li>•BA: echo frequency interval <math>[freq\_min, freq\_max]</math>, 3 decreasing factors <math>\epsilon, \alpha, \gamma</math>.</li> <li>•GOA: attractive intensity <math>f</math>, attractive length <math>l</math>.</li> <li>•CSA: awareness probability <math>AP</math>, flight length <math>fl</math>.</li> <li>•MFO: spiral shape <math>b</math>.</li> <li>•MBO: <i>Keep, partition, period, BAR, <math>S_{max}</math></i>.</li> <li>•BOA: switch probability <math>p</math>, <i>power_exponent, sensory_modality</i>.</li> <li>•PSO: adjusting parameter <math>w</math>, cognitive coefficient <math>c_1</math>, social coefficient <math>c_2</math>.</li> </ul>	$w$ : $w$ -relative step-size $\Delta$ .
12.step-size whose initial values are set before running algorithm. They will change by math formulas during the optimization process.	<ul style="list-style-type: none"> <li>•BA: loudness <math>A</math>, pulse rate <i>rate</i>.</li> <li>•GOA: coefficient <math>c</math>.</li> <li>•CSA: none.</li> <li>•MFO: decreasing weight <math>tt</math>, threshold <i>Fame_no</i>.</li> <li>•MBO: weight <math>\alpha</math>.</li> <li>•BOA: <i>sensory_modality</i>.</li> <li>•PSO: none.</li> </ul>	$z$ : $z$ -relative step-size $\Delta$ .
13.How to initialize $\mathbf{x}_i$ with size $M$ in the $n$ dimension environment.	<ul style="list-style-type: none"> <li>•BA: Eq.2.1.</li> <li>•GOA: Eq.2.7.</li> <li>•CSA: Eq.2.12.</li> <li>•MFO: Eq.2.16.</li> <li>•MBO: Eq.2.21.</li> <li>•BOA: Eq.2.26.</li> <li>•PSO: Eq.2.31.</li> </ul>	$Init_{\mathbf{x}}$ : initialization method on $\mathbf{x}$ .

14.How to initialize changing step-size.	•BA: Eq.2.2,step(4)(6)(7). •GOA: step(3),Eq.2.8. •CSA: Eq.2.13. •MFO: Eq.2.17,Eq.2.18,step(4). •MBO: step(6),Eq.2.23. •BOA: step(3),Eq.2.27. •PSO: Eq.2.32,Eq.2.33,step(5).	$Init_{\Delta}$ : initialization method on $\Delta$ .
15.How to update $x_i$ .	•BA: Eq.2.5,Eq.2.6. •GOA: Eq.2.10. •CSA: Eq.2.14. •MFO: Eq.2.19,Eq.2.20. •MBO: step(5)(6),Eq.2.22. •BOA: Eq.2.28,Eq.2.29. •PSO: Eq.2.35.	$Opt_x$ : optimization method on $x$ .
16.How to update changing step-size.	•BA: Eq.2.3,Eq.2.4,Eq.2.5. •GOA: Eq.2.8,step(9). •CSA: Eq.2.15. •MFO: Eq.2.17,Eq.2.18, step(4). •MBO: Eq.2.23,step(5). •BOA: step(3),Eq.2.30. •PSO: Eq.2.34,sep(7)(8).	$Opt_{\Delta}$ : optimization method on $\Delta$ .
17.How to deal with outliers in $x_i$ .	•BA: step(10). •GOA: step(7). •CSA: step(6). •MFO: step(3). •MBO: step(7). •BOA: step(7). •PSO: common method.	$C$ : outliers treatment in $x_i$ .
18.How to decide if a updated $x_i$ would be accepted.	•BA: step(11). •GOA: none. •CSA: none. •MFO: none. •MBO: step(11). •BOA: step(8). •PSO: none.	$S$ : selection method to select which individuals will be delivered into the next generation.
19.The iteration counter.	•BA: $t$ . •GOA: $l$ . •CSA: $t$ . •MFO: $t$ . •MBO: $t$ . •BOA: $t$ . •PSO: current iteration.	$t$ : current iteration.
20.The maximum number of iterations.	•BA: $t_{max}$ . •GOA: $max\_iteration$ . •CSA: $tmax$ . •MFO: $T$ . •MBO: $Naxgen$ . •BOA: $N\_iter$ . •PSO: stop condition.	$T$ : stop condition.



# Chapter 3

## Unified Procedure

### 3.1 Introduction

**This chapter discusses the commonality among these seven selected algorithms on the level of their procedures for achieving optimization. In other words, we discuss how these 20 unified components make up a valid algorithm.** Specifically, we are interested in whether these seven selected algorithms can arrange these 20 components in the same order.

In [Chapter 2](#), we have discussed that these algorithms are able to be composed of 20 unified components, and these components can be categorized into compulsory and selective components according to their information. In this chapter, we also discuss the category of these components, but the principle of categorizing becomes their functions. This is because the function of each component determines their positions when constructing a valid nature-inspired algorithm. As shown in [Table 3.1](#), we find that **these 20 unified components can be categorized into eight tuples, and these eight tuples are enough to construct any one of these seven selected algorithms**. Therefore, these selected swarm-based optimization algorithms  $NIOA$ <sup>1</sup> can be defined in 8-tuple as [Eq.3.1](#).

$$NIOA = (f, Init_x, Opt_x, C, T, S, Init_{\Delta}, Opt_{\Delta}) \quad 3.1$$

where

---

<sup>1</sup>We expect our framework can be applied in all nature-inspired algorithms, therefore, the  $NIOA$  is short for Nature-Inspired Optimization Algorithm

- *Tuple.1*: 'Evaluation' function  $f$ :  
'Evaluation' is needed to measure the quality of solutions that are found by the algorithm to the problem, when modeling a valid swarm-based algorithm.
- *Tuple.2*: 'Initialize population' function  $Init_x$ :  
'Initialize population' provides a set of possible solutions. A swarm-based algorithm will improve this set and find the optimal one in this set.
- *Tuple.3*: 'Initialize step-size' function  $Init_\Delta$ :  
'Initialize step-size' helps the swarm-based algorithm to improve the quality of possible solutions efficiently. In [Chapter 2](#), we have discussed that there are four kinds of step-size  $\Delta$ :  $w$ -relative  $\Delta$ ,  $z$ -relative  $\Delta$ ,  $x$ -relative  $\Delta$  and  $y$ -relative  $\Delta$ .
- *Tuple.4*: 'Update population' function  $Opt_x$ :  
'Update population' is the strategy that the algorithm used to improve the quality of possible solutions to the problem.
- *Tuple.5*: 'Update step-size' function  $Opt_\Delta$ :  
If the step-size  $\Delta$  used by the algorithm is dynamically changing, such as  $z$ -relative  $\Delta$ ,  $x$ -relative  $\Delta$  and  $y$ -relative  $\Delta$ , 'Update step-size' function is needed to implement this mechanism of changing.
- *Tuple.6*: 'Outliers treatment' function  $C$ :  
'Outliers treatment' deals with every outlier when the swarm-based algorithm updates the population.
- *Tuple.7*: 'Selection' function  $S$ :  
'Selection' determines how many updated individuals will be delivered into the next generation round, when the nature-inspired algorithm iteratively updates the population.
- *Tuple.8*: 'Stop strategy' function  $T$ :  
'Stop strategy' determines when the iterative optimization process stops.

Table 3.1: These 20 unified components can be categorized into eight tuples, according to their functions. Numbers in the 'Component' column represents the index of information discussed in [Table 2.1](#)

Tuple	Component		Function	
<i>Tuple.1</i>	1. <i>f</i>		evaluation	
<i>Tuple.2</i>	2. <i>x<sub>i</sub></i> 3. <i>n</i> 4. <i>lb<sub>x</sub>/ub<sub>x</sub></i> 5. <i>M</i> 13. <i>Init<sub>x</sub></i>		initialize the initial population.	
<i>Tuple.3</i>	14. <i>Init<sub>Δ</sub></i>	6. <i>x<sub>i<sub>p</sub></sub></i> 7. <i>x<sub>g</sub></i> 8. <i>x<sub>s</sub></i>	initialize <i>x</i> -relative step-size <i>Δ</i> .	initialize step-size <i>Δ</i> .
		9. <i>y<sub>i</sub></i> 10. <i>lb<sub>y</sub>/ub<sub>y</sub></i>	initialize <i>y</i> -relative step-size <i>Δ</i> .	
		11. <i>w</i>	initialize <i>w</i> -relative step-size <i>Δ</i> .	
		12. <i>z</i>	initialize <i>z</i> -relative step-size <i>Δ</i> .	
<i>Tuple.4</i>	15. <i>Opt<sub>x</sub></i>		update the population.	
<i>Tuple.5</i>	16. <i>Opt<sub>Δ</sub></i>		update step-size <i>Δ</i> .	
<i>Tuple.6</i>	17. <i>C</i>		outliers treatment.	
<i>Tuple.7</i>	18. <i>S</i>		deliver selected individuals into next generation.	
<i>Tuple.8</i>	19. <i>t</i> 20. <i>T</i>		iteration counter, and the maximum number of generation	

In the following sections, [Section 3.2](#) discusses whether these eight tuples can be ordered at the same position in these seven selected algorithms. The overview summary about this chapter is displayed in [Section 3.3](#).

## 3.2 Unified Procedure

As observed in Table 3.2, we find that each algorithm only has two processes — Initialization process and Optimization process — separated by a stop strategy *Tuple.8 T*.

In the Initialization process, *Tuple.2 Init<sub>x</sub>*, *Tuple.1 f* and *Tuple.3 Init<sub>Δ:w,z,y,x</sub>* exist in every algorithm except the MFO. The MFO does not evaluate the initial population in the initialization process. However, in MFO's optimization process, the *Tuple.1 f* happens before updating the population, which means the initial population also will be evaluated, but it will be evaluated in the optimization process. Therefore, considering the swarm-based algorithm is a kind of an iterative optimization heuristic algorithm, the feature of iteration leads the Initialization process can have one *Tuple.1 f* function, and then the *Tuple.1 f* function in the optimization process has to be moved after updating the population.

After making this change, in the Optimization process, we can conclude that *Tuple.4 Opt<sub>x</sub>*, *Tuple.6 C* and *Tuple.1 f* also exist in every algorithm, and the order of them must be that *Tuple.4 Opt<sub>x</sub>* is before *Tuple.6 C* followed by *Tuple.1 f*. Because outliers only might appear after updating the population. Furthermore, three places are also meaningful to be unified:

- (1) The position of *Tuple.3 Init<sub>Δ:z</sub>* and *Tuple.5 Opt<sub>Δ:z</sub>* in BA, GOA, MFO and MBO. As shown in Table 3.2, the position of *Tuple.3 Init<sub>Δ:z</sub>* and *Tuple.5 Opt<sub>Δ:z</sub>* in these three algorithms are together assigned at the beginning of the optimization process. However, we can find that the generalized math equation of *Tuple.3 Init<sub>Δ:z</sub>* and *Tuple.5 Opt<sub>Δ:z</sub>* has two forms. For example,
  - In BA, the generalized math equation  $G_1$  of one  $z$  is  $z(t+1) = G_1(z(t))$  which means the *Tuple.3 Init<sub>Δ:z</sub>* is assigning an initial value  $z^0$  to  $z(t=0)$  and can move from the beginning of optimization process to the initialization process. Meanwhile, the *Tuple.5 Opt<sub>Δ:z</sub>* has to move to the end of the optimization process. These two assignments are same as in the BOA, where *Tuple.3 Init<sub>Δ:z</sub>* will go to the initialization process, and *Tuple.5 Opt<sub>Δ:z</sub>* will go to the end of the optimization process.
  - In BA, GOA, MFO and MBO, there is another generalized math equation  $G_2$  of  $z$  is  $z(t) = G_2(z(t), t)$  which means both *Tuple.3 Init<sub>Δ:z</sub>* and *Tuple.5 Opt<sub>Δ:z</sub>* are a function of iteration counter  $t$ . It means it doesn't matter where *Tuple.3 Init<sub>Δ:z</sub>* and *Tuple.5 Opt<sub>Δ:z</sub>* are, as far as the iteration counter  $t$  is input correctly. Therefore, we can move *Tuple.3 Init<sub>Δ:z</sub>* to the initialization process, and move *Tuple.5 Opt<sub>Δ:z</sub>* to the end of the optimization process.

- (2) The position of *Tuple.3*  $Init_{\Delta:x}$  and *Tuple.5*  $Opt_{\Delta:x}$  in MFO.  
 The original position of *Tuple.3*  $Init_{\Delta:x}$  and *Tuple.5*  $Opt_{\Delta:x}$  are together at the beginning of the optimization process. After looking at the implement code in MFO, we can make sure the position of them can be separated. The *Tuple.3*  $Init_{\Delta:x}$  can be in the initialization process, and the *Tuple.5*  $Opt_{\Delta:x}$  can move to the end of the optimization process.
- (3) The position of *Tuple.5*  $Opt_{\Delta:y}$  and *Tuple.5*  $Opt_{\Delta:x}$  in all of these selected algorithms.  
 Generally speaking, it would be more reasonable if these two tuples can be at the same position. However, updated y-relative  $\Delta$  is used in the current generation round, but updated x-relative  $\Delta$  is used in the next generation round. Therefore, the position of *Tuple.5*  $Opt_{\Delta:y}$  must be before *Tuple.4*  $Opt_x$ , and the position of *Tuple.5*  $Opt_{\Delta:x}$  must be after *Tuple.1*  $f$  in the optimization.

Meanwhile, we also find that in the Optimization process, *Tuple.7*  $S$  does not exist all of these selected algorithms, such as GOA, CSA, MFO and PSO. However, considering the function of *Tuple.7*  $S$  is to select which individuals can be delivered into the next generation round which is a meaningful part in swam-based algorithms, we prefer assigning *Tuple.7*  $S$  to all of these selected algorithms. For these algorithms who does not use *Tuple.7*  $S$ , we define a kind of selection method in which all updated individuals will be delivered into the next generation round. (see [Chapter 4](#) for more details about this special selection method).

Table 3.2: Original positions of these eight tuples in these seven algorithms

Process	BA	GOA	CSA	MFO	MBO	BOA	PSO
Initialization	<i>Tuple.2</i> $Init_x$	<i>Tuple.2</i> $Init_x$	<i>Tuple.2</i> $Init_x$	<i>Tuple.2</i> $Init_x$	<i>Tuple.2</i> $Init_x$	<i>Tuple.2</i> $Init_x$	<i>Tuple.2</i> $Init_x$
	<i>Tuple.1</i> $f$	<i>Tuple.1</i> $f$	<i>Tuple.1</i> $f$	<i>Tuple.3</i> $Init_{\Delta:w}$	<i>Tuple.1</i> $f$	<i>Tuple.1</i> $f$	<i>Tuple.1</i> $f$
	<i>Tuple.3</i> $Init_{\Delta:w}$	<i>Tuple.3</i> $Init_{\Delta:w}$	<i>Tuple.3</i> $Init_{\Delta:w}$		<i>Tuple.3</i> $Init_{\Delta:w}$	<i>Tuple.3</i> $Init_{\Delta:z}$	<i>Tuple.3</i> $Init_{\Delta:w}$
	<i>Tuple.3</i> $Init_{\Delta:y}$	<i>Tuple.3</i> $Init_{\Delta:x}$	<i>Tuple.3</i> $Init_{\Delta:x}$		<i>Tuple.3</i> $Init_{\Delta:x}$	<i>Tuple.3</i> $Init_{\Delta:x}$	<i>Tuple.3</i> $Init_{\Delta:x}$

continued ...

... continued

Process	BA	GOA	CSA	MFO	MBO	BOA	PSO
	<i>Tuple.3</i> <i>Init<sub>Δ:x</sub></i>						<i>Tuple.3</i> <i>Init<sub>Δ:y</sub></i>
Stop strategy <i>Tuple.8 T</i>							
Optimization	<i>Tuple.3</i> <i>Init<sub>Δ:z</sub></i> and <i>Tuple.5</i> <i>Opt<sub>Δ:z</sub></i>	<i>Tuple.3</i> <i>Init<sub>Δ:z</sub></i> and <i>Tuple.5</i> <i>Opt<sub>Δ:z</sub></i>	<i>Tuple.4</i> <i>Opt<sub>x</sub></i>	<i>Tuple.3</i> <i>Init<sub>Δ:z</sub></i> and <i>Tuple.5</i> <i>Opt<sub>Δ:z</sub></i>	<i>Tuple.3</i> <i>Init<sub>Δ:z</sub></i> and <i>Tuple.5</i> <i>Opt<sub>Δ:z</sub></i>	<i>Tuple.4</i> <i>Opt<sub>x</sub></i>	<i>Tuple.5</i> <i>Opt<sub>Δ:y</sub></i>
	<i>Tuple.5</i> <i>Opt<sub>Δ:y</sub></i>	<i>Tuple.4</i> <i>Opt<sub>x</sub></i>	<i>Tuple.6</i> <i>C</i>	<i>Tuple.6</i> <i>C</i>	<i>Tuple.4</i> <i>Opt<sub>x</sub></i>	<i>Tuple.6</i> <i>C</i>	<i>Tuple.4</i> <i>Opt<sub>x</sub></i>
	<i>Tuple.4</i> <i>Opt<sub>x</sub></i>	<i>Tuple.6</i> <i>C</i>	<i>Tuple.1</i> <i>f</i>	<i>Tuple.1</i> <i>f</i>	<i>Tuple.1</i> <i>f</i>	<i>Tuple.1</i> <i>f</i>	<i>Tuple.6</i> <i>C</i>
	<i>Tuple.6</i> <i>C</i>	<i>Tuple.1</i> <i>f</i>	<i>Tuple.5</i> <i>Opt<sub>Δ:x</sub></i>	<i>Tuple.3</i> <i>Init<sub>Δ:x</sub></i> and <i>Tuple.5</i> <i>Opt<sub>Δ:x</sub></i>	<i>Tuple.4</i> <i>Opt<sub>x</sub></i>	<i>Tuple.7</i> <i>S</i>	<i>Tuple.1</i> <i>f</i>
	<i>Tuple.1</i> <i>f</i>	<i>Tuple.5</i> <i>Opt<sub>Δ:x</sub></i>		<i>Tuple.4</i> <i>Opt<sub>x</sub></i>	<i>Tuple.1</i> <i>f</i>	<i>Tuple.5</i> <i>Opt<sub>Δ:x</sub></i>	<i>Tuple.5</i> <i>Opt<sub>Δ:x</sub></i>
	<i>Tuple.7</i> <i>S</i>				<i>Tuple.7</i> <i>S</i>	<i>Tuple.5</i> <i>Opt<sub>Δ:z</sub></i>	
	<i>Tuple.5</i> <i>Opt<sub>Δ:x</sub></i>				<i>Tuple.5</i> <i>Opt<sub>Δ:x</sub></i>		

Based on these previous discussions, the unified procedure for these seven selected algorithms can be formulated as [Algorithm.1](#). As the [Algorithm.1](#) described, every algorithm will start at the Initialization process in which the population must be evaluated after being initialized, moreover, different algorithms initialize  $Init_{\Delta}$  differently. For example, CSA only uses  $w$ -relative and  $x$ -relative step-size  $\Delta$ , but PSO uses one more  $y$ -relative step-size  $\Delta$ .

In the optimization process, if  $y$ -relative step-size  $\Delta$  exists in the initialization process, it will be firstly updated before updating the population. Evaluation happens after dealing with outliers in the new generated population. Selection method is possible existing in all of these seven algorithms, and it must exit after evaluation action. The last step in the optimization process is to update  $z$ -relative and  $x$ -relative step-size  $\Delta$ , if they exist in the initialization process.

---

**Algorithm 1** Unified Procedure for seven selected algorithms (the unified positions of eight tuples in [Table 3.1](#))

---

```

1:  $t \leftarrow 0$                                 ▷ iteration counter
2: Tuple.2  $Init_x$                             ▷ initialize population
3: Tuple.1  $f$                                 ▷ evaluation
4: Tuple.3  $Init_{\Delta:w,z,x,y}$                 ▷ initialize  $w$ -relative,  $z$ -relative,  $x$ -relative and  $y$ -relative
   step-size  $\Delta$ 
5: while Tuple.8  $T$  do                    ▷ stop strategy
6:   Tuple.5  $Opt_{\Delta:y}$  ▷ update  $y$ -relative step-size  $\Delta$ , if  $y$ -relative step-size exists in the
   initialization process
7:   Tuple.4  $Opt_x$                             ▷ update population
8:   Tuple.6  $C$                                 ▷ outliers treatment
9:   Tuple.1  $f$                                 ▷ evaluation
10:  Tuple.7  $S$                                 ▷ selection
11:  Tuple.5  $Opt_{\Delta:z,x}$                     ▷ update  $z$ -relative and  $x$ -relative step-size  $\Delta$ 
12:   $t \leftarrow t + 1$ 
13: end while

```

---

### 3.3 Summary

In this chapter, we discuss the possibility to unify the various procedures among these seven selected algorithm. We firstly categorize the 20 unified components (see [Chapter 2](#)) again, but the principle becomes their functions in this chapter. As summarized in [Table 3.1](#), all of these seven algorithms can be represented by eight tuples, and the unified representation is defined as [Equation 3.1](#). Secondly, we detail each algorithm's original position of each tuple, and [Table 3.2](#) displays all of these original positions. Then, we discuss the possibility of arranging these eight tuples in the same positions in these seven algorithms. Lastly, the final unified procedures is performed in [Algorithm.1](#).

In conclusion, all of these seven algorithms can be represented by eight unified tuples, and these eight tuples can be at the same position when modeling these seven algorithms.

# Chapter 4

## Unified Framework

### 4.1 Motivation

In this chapter, we aim to complete the purpose in this work — **building up a unified framework for selected algorithms** — by combining the Unified Terminologies (see [Chapter 2](#)) and the Unified Procedure (see [Chapter 3](#)) together. Specifically, **we focus on studying how these eight tuples (see [Chapter 3](#)) are represented in each algorithm by using unified terminologies (see [Chapter 2](#)), because the order of these eight tuples is already fixed in [Chapter 3](#).**

Firstly, we want to introduce several special symbols used in this work. As we discussed in [Chapter 2](#), the swarm-based algorithm happens on a set of individuals  $\mathbf{x}_i$ , therefore, we use  $\mathbf{X}$  and  $\mathbf{x}_i$  to distinguish the population and the individual. For example, the set of personal best individual  $\mathbf{x}_{i_p}$  is  $\mathbf{X}_p$ , but there is only one global best individual  $\mathbf{x}_g$ . Meanwhile, the capital letter  $F$  is used to represent the set of the fitness of the whole population  $\mathbf{X}$ .<sup>1</sup>

Moreover, we use  $\hat{\mathbf{X}}$  and  $\mathbf{X}$  to distinguish the population updated before selection operation and the population updated after selection operation. We also want to clarify that the population updated after the selection operation is the finally updated population after one generation round, on the contrary, the population updated before the selection operation is the temporarily updated population.

---

<sup>1</sup>The capital letter in bold style denotes a set of vector. The capital letter in plain style denotes a set of numerics.



In the following sections, [Section 4.2](#) introduces how these original representations of each algorithm is represented with only eight tuples. We also mention difficulties when re-framing the original model into these eight tuples. Next, in [Section 4.3](#), we display the Unified Nature-Inspired Optimization Algorithm (UNIOA) and give discussions. The summary about this chapter is shown in [Section 4.4](#).

## 4.2 Re-framed Nature-inspired Algorithms

### 4.2.1 Re-framed BA

The *Tuple.1*  $f$  can be represented as  $f(\mathbf{x}_i)$ . The *Tuple.2*  $Init_{\mathbf{x}}$  can be extracted from [Eq.2.1](#), and re-framed as [Eq.4.1](#).

$$\mathbf{x}_i(t=0) = \mathcal{U}(lb_{\mathbf{x}}, ub_{\mathbf{x}}), i = 1, 2, \dots, M \quad 4.1$$

For the *Tuple.3*  $Init_{\Delta}$ , as discussed in [Table 2.1](#), we find the velocity  $\mathbf{v}_i$  is a kind of  $y$ -relative step-size, and the global best individual  $\mathbf{x}_*$  is a kind of  $x$ -relative step-size. Meanwhile, the loudness  $A$  and the rate  $r$  are a kind of  $z$ -relative step-size. Therefore, we re-frame  $\mathbf{v}_i$  (see step (2) in [Subsection 2.2.1](#)) into [Eq.4.2](#). The  $\mathbf{x}_*$  (see step (4) in [Subsection 2.2.1](#)) can be re-framed as [Eq.4.3](#). The two  $z$ -relative step-size  $A$  and  $r$  can be respectively re-framed as [Eq.4.4](#) and [Eq.4.5](#).

$$\mathbf{y}_i(t=0) = \mathcal{U}(lb_{\mathbf{y}}, ub_{\mathbf{y}}), i = 1 \dots M \quad 4.2$$

$$\mathbf{x}_g(t=0) = \mathbf{Min}(\{\mathbf{x}_i(t=0)\}), i = 1 \dots M \quad 4.3$$

$$z_1(t=0) = z_1^0 \times w_2 \quad 4.4$$

$$z_2(t=0) = z_2^0 \times (1 - e^{-w_1 \times t}) \quad 4.5$$

For the *Tuple.4*  $Opt_{\mathbf{x}}$ , we find step (8) is able to be re-framed into one equation as [Eq.4.6](#).

$$\hat{\mathbf{x}}_i(t+1) = \begin{cases} \mathbf{x}_g(t) + w_3 \times \mathbf{rand} \times z_1(t), & rand < z_2(t) \\ \mathbf{x}_i(t) + \mathbf{y}_i(t+1), & \text{o.w} \end{cases} \quad 4.6$$

For the *Tuple.5*  $Opt_{\Delta}$  that corresponds to *Tuple.3*  $Init_{\Delta}$ , we find the update way of  $\mathbf{v}_i$  (see one step in step (8) in [Subsection 2.2.1](#)) is able to be re-framed into as [Eq.4.7](#). The update way of  $\mathbf{x}_*$  (see step (9) in [Subsection 2.2.1](#)) is able to be re-framed as [Eq.4.8](#). The update way of  $A$  (see [Eq.2.3](#)) and the update way of  $r$  (see [Eq.2.4](#)) can be respectively re-framed as [Eq.4.9](#) and [Eq.4.10](#).

$$\mathbf{y}_i(t+1) = \mathbf{y}_i(t) + \mathcal{U}(lb_{w_4}, ub_{w_4}) \times (\mathbf{x}_i(t) - \mathbf{x}_g(t)) \quad 4.7$$

$$\mathbf{x}_g(t+1) = \mathbf{Min}(\mathbf{x}_g(t) \cup \{\mathbf{x}_i(t+1)\}), i = 1 \dots M \quad 4.8$$

$$z_1(t+1) = z_1(t) \times w_2 \quad 4.9$$

$$z_2(t+1) = z_2^0 \times (1 - e^{-w_1 \times (t+1)}) \quad 4.10$$

For the *Tuple.6 C*, the step (10) can be re-framed as [Eq.4.11](#).

$$\mathbf{x}_{i,n}^{\text{fixed}}(t+1) = \begin{cases} ub_{\mathbf{x}} & , \mathbf{x}_{i,n}(t+1) > ub_{\mathbf{x}} \\ \mathbf{x}_{i,n}(t+1) & , \text{o.w} \\ lb_{\mathbf{x}} & , \mathbf{x}_{i,n}(t+1) < lb_{\mathbf{x}} \end{cases} \quad 4.11$$

For the *Tuple.7 S*, the step (11) can be re-framed as [Eq.4.12](#).

$$\mathbf{x}_i(t+1) = \begin{cases} \hat{\mathbf{x}}_i(t+1) & , f(\hat{\mathbf{x}}_i(t+1)) < f(\mathbf{x}_i(t)) \text{ or } rand > z_1(t) \\ \mathbf{x}_i(t) & , \text{o.w} \end{cases} \quad 4.12$$

For the *Tuple.8 T* which is a stop condition related to the maximum number of iteration.

Till now, all of other unknown parameters can be defined as  $w$ -relative step-size, and they belong to the *Tuple.3 Init<sub>Δ</sub>*. It has common  $w$ -relative step-size  $w_1 = 0.1$ ,  $w_2 = 0.97$ ,  $[lb_{w_4}, ub_{w_4}] = [0, 2]$ ,  $w_3 = 0.1$  and  $[lb_{\mathbf{y}}, ub_{\mathbf{y}}] = [0, 0]$ . It also has special  $w$ -relative step-size that is related to the initial value of  $z$ -relative step-size  $z_1^0 = 1$ ,  $z_2^0 = 1$ .

Therefore, the complete pseudo-code of re-framed BA model can be organized as [Algorithm.2](#). When we re-frame the BA algorithm, the main difficulty is how to formulate the  $z$ -relative step-size. For example, the  $z_1$  (that is the loudness  $A$  in [Subsection 2.2.1](#)) is located at the beginning of optimization process. Therefore, when we move it to the initialization process, the updated  $z_1$  has to be moved after updating the population.

**Algorithm 2** Re-framed BA

---

```

1:  $t \leftarrow 0$ 
2:  $\mathbf{X}(t) \leftarrow \text{Init}_{\mathbf{x}}(n, M, [lb_{\mathbf{x}}, ub_{\mathbf{x}}])$  as Eq.4.1 ▷ initialize initial population
3:  $F(t) \leftarrow f(\mathbf{X}(t))$  ▷ evaluate
4:  $w, [lb_{\mathbf{y}}, ub_{\mathbf{y}}], z^0 \leftarrow \text{Init}_{\Delta:w}(w, [lb_{\mathbf{y}}, ub_{\mathbf{y}}], z^0)$  ▷ initialize  $w$ -relative step-size
5:  $\mathbf{Y}(t) \leftarrow \text{Init}_{\Delta:\mathbf{y}}(n, M, [lb_{\mathbf{y}}, ub_{\mathbf{y}}])$  as Eq.4.2 ▷ initialize  $\mathbf{y}$ -relative step-size
6:  $\mathbf{x}_g(t) \leftarrow \text{Init}_{\Delta:\mathbf{x}}(\mathbf{X}(t))$  as Eq.4.3 ▷ initialize  $\mathbf{x}$ -relative step-size
7:  $z(t) \leftarrow \text{Init}_{\Delta:z}(t, z^0, w)$  as Eq.4.4, Eq.4.5 ▷ initialize  $z$ -relative step-size
8: while stop condition  $T$  do
9:    $\mathbf{Y}(t+1) \leftarrow \text{Opt}_{\Delta:\mathbf{y}}(\mathbf{Y}(t), \mathbf{x}_g, w)$  as Eq.4.7 ▷ update  $\mathbf{y}$ -relative step-size
10:   $\hat{\mathbf{X}}(t+1) \leftarrow \text{Opt}_{\mathbf{x}}(\mathbf{X}(t), \mathbf{Y}(t), \mathbf{x}_g(t), z(t), w)$  as Eq.4.6 ▷ generate temporarily
    updated population
11:   $\hat{\mathbf{X}}(t+1) \leftarrow C(\hat{\mathbf{X}}(t+1))$  as Eq.4.11 ▷ treatment to outliers
12:   $F(t+1) \leftarrow f(\hat{\mathbf{X}}(t+1))$  ▷ evaluate
13:   $\mathbf{X}(t+1) \leftarrow S(\mathbf{X}(t), \hat{\mathbf{X}}(t+1), z(t))$  as Eq.4.12 ▷ select and generate finally
    updated population
14:   $\mathbf{x}_g(t+1) \leftarrow \text{Init}_{\Delta:\mathbf{x}}(\mathbf{X}(t), \mathbf{X}(t+1))$  as Eq.4.8 ▷ update  $\mathbf{x}$ -relative step-size
15:   $z(t+1) \leftarrow \text{Opt}_{\Delta:z}(t+1, z(t), w)$  as Eq.4.9, Eq.4.10 ▷ update  $z$ -relative step-size
16:   $t \leftarrow t+1$ 
17: end while

```

---

**4.2.2 Re-framed GOA**

The *Tuple.1*  $f$  can be represented as  $F(t) = f(\mathbf{X}(t))$ . The *Tuple.2*  $\text{Init}_{\mathbf{x}}$  can be extracted from Eq.2.7, and re-framed as Eq.4.13.

$$\mathbf{x}_i(t=0) = \mathcal{U}(lb_{\mathbf{x}}, ub_{\mathbf{x}}), i = 1, 2, \dots, M \quad 4.13$$

For the *Tuple.3*  $\text{Init}_{\Delta}$ , as discussed in Table 2.1, we find the global best individual  $\mathbf{T}$  is a kind of  $\mathbf{x}$ -relative step-size. Meanwhile, the hyper-parameter  $c$  is a kind of  $z$ -relative step-size. Therefore, we re-frame  $\mathbf{T}$  (see step (3)) into Eq.4.14. The  $z$ -relative step-size  $c$  can be re-framed as Eq.4.15.

$$\mathbf{x}_g(t=0) = \text{Min}(\{\mathbf{x}_i(t=0)\}), i = 1 \dots M \quad 4.14$$

$$z(t=0) = ub_z - t \times \left( \frac{ub_z - lb_z}{T} \right) \quad 4.15$$

For the *Tuple.4*  $Opt_{\mathbf{x}}$ , we find [Eq.2.9](#), [Eq.2.10](#) and [Eq.2.11](#) are able to be re-framed into one equation as [Eq.4.16](#).

$$\begin{aligned} \tilde{D}_{i,j}(t) &= 2 + \mathbf{Dist}(\mathbf{x}_i(t), \mathbf{x}_j(t)) \bmod 2 \\ \hat{\mathbf{x}}_i(t+1) &= z(t) \times \left( \sum_{j=1, j \neq i}^M z(t) \times \frac{ub_{\mathbf{x}} - lb_{\mathbf{x}}}{2} \times (w_1 \times e^{\frac{-\tilde{D}_{i,j}(t)}{w_2}} - e^{-\tilde{D}_{i,j}(t)}) \times \frac{\mathbf{x}_i(t) - \mathbf{x}_j(t)}{\mathbf{Dist}(\mathbf{x}_i(t), \mathbf{x}_j(t))} \right) + \mathbf{x}_g \end{aligned} \quad 4.16$$

For the *Tuple.5*  $Opt_{\Delta}$  that corresponds to *Tuple.3*  $Init_{\Delta}$ , we find the update way of  $\mathbf{T}$  (see step (9)) is able to be re-framed into as [Eq.4.17](#). The update way of  $z$  (see [Eq.2.8](#)) can be re-framed as [Eq.4.18](#).

$$\mathbf{x}_g(t+1) = \mathbf{Min}(\mathbf{x}_g(t) \cup \{\mathbf{x}_i(t+1)\}), i = 1 \dots M \quad 4.17$$

$$z(t+1) = ub_z - (t+1) \times \left( \frac{ub_z - lb_z}{T} \right) \quad 4.18$$

For the *Tuple.6*  $C$ , the step (7) can be re-framed as [Eq.4.19](#).

$$\mathbf{x}_{i,n}^{\text{fixed}}(t+1) = \begin{cases} ub_x & , \mathbf{x}_{i,n}(t+1) > ub_{\mathbf{x}} \\ \mathbf{x}_{i,n}(t+1) & , \text{o.w} \\ lb_x & , \mathbf{x}_{i,n}(t+1) < lb_{\mathbf{x}} \end{cases} \quad 4.19$$

For the *Tuple.7*  $S$ , the GOA model doesn't have a process to select which individuals will be delivered into the next generation, therefore, we re-frame it as [Eq.4.20](#) that means all newly generated individuals will be delivered into the next generation.

$$\mathbf{x}_i(t+1) = \hat{\mathbf{x}}_i(t+1) \quad 4.20$$

For the *Tuple.8*  $T$  which is a stop condition related to the maximum number of iteration.

Till now, all of other unknown parameters can be defined as  $w$ -relative step-size, and they belong to the *Tuple.3*  $Init_{\Delta}$ . It has common  $w$ -relative step-size  $[ub_w, lb_w] = [0.00004, 1]$ ,  $w_1 = 0.5$ , and  $w_2 = 1.5$ .

Therefore, the complete pseudo-code of re-framed BA model can be organized as [Algorithm.3](#).

**Algorithm 3** Re-framed GOA

---

```

1:  $t \leftarrow 0$ 
2:  $\mathbf{X}(t) \leftarrow \text{Init}_{\mathbf{x}}(n, M, [lb_{\mathbf{x}}, ub_{\mathbf{x}}])$  as Eq.4.13 ▷ initialize initial population
3:  $F(t) \leftarrow f(\mathbf{X}(t))$  ▷ evaluate
4:  $w \leftarrow \text{Init}_{\Delta:w}(w)$  ▷ initialize  $w$ -relative step-size
5:  $\mathbf{x}_g(t) \leftarrow \text{Init}_{\Delta:\mathbf{x}}(\mathbf{X}(t))$  as Eq.4.14 ▷ initialize  $\mathbf{x}$ -relative step-size
6:  $z(t) \leftarrow \text{Init}_{\Delta:z}(z^0)$  as Eq.4.15 ▷ initialize  $z$ -relative step-size
7: while stop condition  $T$  do
8:    $\hat{\mathbf{X}}(t+1) \leftarrow \text{Opt}_{\mathbf{x}}(\mathbf{X}(t), \mathbf{x}_g(t), z(t), w)$  as Eq.4.16 ▷ generate temporarily updated population
9:    $\hat{\mathbf{X}}(t+1) \leftarrow C(\hat{\mathbf{X}}(t+1))$  as Eq.4.19 ▷ treatment to outliers
10:   $F(t+1) \leftarrow f(\hat{\mathbf{X}}(t+1))$  ▷ evaluate
11:   $\mathbf{X}(t+1) \leftarrow S(\mathbf{X}(t), \hat{\mathbf{X}}(t+1))$  as Eq.4.20 ▷ select and generate finally updated population
12:   $\mathbf{x}_g(t+1) \leftarrow \text{Init}_{\Delta:\mathbf{x}}(\mathbf{X}(t), \mathbf{X}(t+1))$  as Eq.4.17 ▷ update  $\mathbf{x}$ -relative step-size
13:   $z(t+1) \leftarrow \text{Opt}_{\Delta:z}(z(t), t+1)$  as Eq.4.18 ▷ update  $z$ -relative step-size
14:   $t \leftarrow t+1$ 
15: end while

```

---

**4.2.3 Re-framed CSA**

The *Tuple.1*  $f$  can be represented as  $F(t) = f(\mathbf{X}(t))$ . The *Tuple.2*  $\text{Init}_{\mathbf{x}}$  can be extracted from Eq.2.12, and re-framed as Eq.4.21.

$$\mathbf{x}_i(t=0) = \mathcal{U}(lb_{\mathbf{x}}, ub_{\mathbf{x}}), i = 1, 2, \dots, M \quad 4.21$$

For the *Tuple.3*  $\text{Init}_{\Delta}$ , as discussed in Table 2.1, we find the personal best  $\mathbf{m}_i$  is a kind of  $\mathbf{x}$ -relative step-size. Therefore, we re-frame  $\mathbf{m}_i$  (see step (3)) into Eq.4.22.

$$\mathbf{x}_{i_p}(t=0) = \mathbf{x}_i(t=0), i = 1 \dots M \quad 4.22$$

For the *Tuple.4*  $\text{Opt}_{\mathbf{x}}$ , we find Eq.2.14 is able to be re-framed as Eq.4.23 in which  $\mathbf{x}_{j_p}$  is any one neighbor around the  $\mathbf{x}_{i_p}$ .

$$\hat{\mathbf{x}}_i(t+1) = \begin{cases} \mathbf{x}_i(t) + rand \times w_2 \times (\mathbf{x}_{j_p}(t) - \mathbf{x}_i(t)) & , \quad r > w_1 \\ \mathcal{U}(lb_{\mathbf{x}}, ub_{\mathbf{x}}) & , \quad \text{o.w} \end{cases} \quad 4.23$$

For the *Tuple.5*  $\text{Opt}_{\Delta}$  that corresponds to *Tuple.3*  $\text{Init}_{\Delta}$ , we find the update way of  $\mathbf{m}_i$  (see Eq.2.15) is able to be re-framed into as Eq.4.24.

$$\mathbf{x}_{i_p}(t+1) = \text{Min}(\{\mathbf{x}_{i_p}(t), \mathbf{x}_i(t+1)\}) \quad 4.24$$

For the *Tuple.6*  $C$ , the step (6) can be re-framed as [Eq.4.25](#).

$$\mathbf{x}_{i,n}^{\text{fixed}}(t+1) = \begin{cases} \mathbf{x}_{i,n}(t) & , \quad \mathbf{x}_{i,n}(t+1) < lb_{\mathbf{x}} \text{ or } \mathbf{x}_{i,n}(t+1) > ub_{\mathbf{x}} \\ \mathbf{x}_{i,n}(t+1) & , \quad \text{o.w} \end{cases} \quad 4.25$$

For the *Tuple.7*  $S$ , the CSA model doesn't have a process to select which individuals will be delivered into the next generation, therefore, we re-frame it as [Eq.4.26](#) that means all newly generated individuals will be delivered into the next generation.

$$\mathbf{x}_i(t+1) = \hat{\mathbf{x}}_i(t+1) \quad 4.26$$

For the *Tuple.8*  $T$  which is a stop condition related to the maximum number of iteration.

Till now, all of other unknown parameters can be defined as  $w$ -relative step-size, and they belong to the *Tuple.3*  $Init_{\Delta}$ . It has common  $w$ -relative step-size  $w_1 = 0.1$  and  $w_2 = 2$ .

Therefore, the complete pseudo-code of re-framed BA model can be organized as [Algorithm.3](#).

---

**Algorithm 4** Re-framed CSA

---

```

1:  $t \leftarrow 0$ 
2:  $\mathbf{X}(t) \leftarrow Init_{\mathbf{x}}(n, M, [lb_{\mathbf{x}}, ub_{\mathbf{x}}])$  as Eq.4.21           ▷ initialize initial population
3:  $F(t) \leftarrow f(\mathbf{X}(t))$                                            ▷ evaluate
4:  $w \leftarrow Init_{\Delta:w}(w)$                                          ▷ initialize  $w$ -relative step-size
5:  $\mathbf{X}_p(t) \leftarrow Init_{\Delta:\mathbf{x}}(\mathbf{X}(t))$  as Eq.4.22                 ▷ initialize  $\mathbf{x}$ -relative step-size
6: while stop condition  $T$  do
7:    $\hat{\mathbf{X}}(t+1) \leftarrow Opt_{\mathbf{x}}(\mathbf{X}(t), \mathbf{X}_p(t), w)$  as Eq.4.23   ▷ generate temporarily updated population
8:    $\hat{\mathbf{X}}(t+1) \leftarrow C(\hat{\mathbf{X}}(t+1))$  as Eq.4.25                 ▷ treatment to outliers
9:    $F(t+1) \leftarrow f(\hat{\mathbf{X}}(t+1))$                                    ▷ evaluate
10:   $\mathbf{X}(t+1) \leftarrow S(\mathbf{X}(t), \hat{\mathbf{X}}(t+1))$  as Eq.4.26 ▷ select and generate finally updated population
11:   $\mathbf{X}_p(t+1) \leftarrow Init_{\Delta:\mathbf{x}}(\mathbf{X}(t), \mathbf{X}(t+1))$  as Eq.4.24 ▷ update  $\mathbf{x}$ -relative step-size
12:   $t \leftarrow t + 1$ 
13: end while
    
```

---

#### 4.2.4 Re-framed MFO

The *Tuple.1*  $f$  can be represented as  $F(t) = f(\mathbf{X}(t))$ . The *Tuple.2*  $Init_{\mathbf{x}}$  can be extracted from [Eq.2.16](#), and re-framed as [Eq.4.27](#).

$$\mathbf{x}_i(t=0) = \mathcal{U}(lb_{\mathbf{x}}, ub_{\mathbf{x}}), i = 1, 2, \dots, M \quad 4.27$$

For the *Tuple.3*  $Init_{\Delta}$ , as discussed in Table 2.1, we find the *sort\_population* is a kind of  $\mathbf{x}$ -relative step-size. Meanwhile, the *flame\_no* and the linearly decreasing parameter  $a$  with another parameter  $t$  are a kind of  $z$ -relative step-size. Therefore, we re-frame *sort\_population* (see step (4)) into Eq.4.28. The two  $z$ -relative step-size *flame\_no* (see Eq.2.17) and  $a$  with  $t$  (see Eq.2.18, step (5)) can be respectively re-framed as Eq.4.28 and Eq.4.29.

$$\langle \mathbf{x}_i(t=0) \rangle = \mathbf{Sort}(\{\mathbf{x}_i(t=0)\}), i = 1 \dots M \quad 4.28$$

$$z_2(t=0) = \mathbf{Round}(M - t \times \frac{M-1}{T}) \quad 4.29$$

$$z_{1_i}(t=0) = rand \times (-2 - \frac{t}{T}) + 1 \quad 4.30$$

For the *Tuple.4*  $Opt_{\mathbf{x}}$ , we find Eq.2.19 and Eq.2.20 are able to be re-framed as Eq.4.31.

$$\hat{\mathbf{x}}_i(t+1) = \begin{cases} (\mathbf{x}_{s_i}(t) - \mathbf{x}_i(t)) \times e^{w \times z_{1_i}(t)} \times \cos(2\pi \times z_{1_i}(t)) + \mathbf{x}_{s_i}(t), & i \leq z_2(t) \\ (\mathbf{x}_{s_{z_2(t)}}(t) - \mathbf{x}_i(t)) \times e^{w \times z_{1_i}(t)} \times \cos(2\pi \times z_{1_i}(t)) + \mathbf{x}_{s_{z_2(t)}}(t), & \text{o.w.} \end{cases} \quad 4.31$$

For the *Tuple.5*  $Opt_{\Delta}$  that corresponds to *Tuple.3*  $Init_{\Delta}$ , we find the update way of *sort\_population* into Eq.4.32. The two  $z$ -relative step-size *flame\_no* and  $a$  with  $t$  can be respectively re-framed as Eq.4.33 and Eq.4.34.

$$\langle \mathbf{x}_i(t+1) \rangle = \mathbf{Sort}(\{\mathbf{x}_i(t)\} \cup \{\mathbf{x}_i(t+1)\}), i = 1 \dots M \quad 4.32$$

$$z_2(t+1) = \mathbf{Round}(M - (t+1) \times \frac{M-1}{T}) \quad 4.33$$

$$z_{1_i}(t+1) = rand \times (-2 - \frac{t+1}{T}) + 1 \quad 4.34$$

For the *Tuple.6*  $C$ , the MFO model doesn't have a process to deal with outliers, but we think the method  $C$  is necessary. Therefore, we use the most common way here to deal with outliers and re-frame it as Eq.4.35.

$$\mathbf{x}_{i,n}^{\text{fixed}}(t+1) = \begin{cases} ub_x & , \mathbf{x}_{i,n}(t+1) > ub_x \\ \mathbf{x}_{i,n}(t+1) & , \text{o.w} \\ lb_x & , \mathbf{x}_{i,n}(t+1) < lb_x \end{cases} \quad 4.35$$

For the *Tuple.7*  $S$ , the MFO model doesn't have a process to select which individuals will be delivered into the next generation, therefore, we re-frame it as Eq.4.36 that means all newly generated individuals will be delivered into the next generation.

$$\mathbf{x}_i(t+1) = \hat{\mathbf{x}}_i(t+1) \quad 4.36$$

For the *Tuple.8*  $T$  which is a stop condition related to the maximum number of iteration.

Till now, all of other unknown parameters can be defined as  $w$ -relative step-size, and they belong to the *Tuple.3*  $Init_{\Delta}$ . It has common  $w$ -relative step-size  $w = 1$ .

Therefore, the complete pseudo-code of re-framed BA model can be organized as [Algorithm.3](#).

**Algorithm 5** Re-framed MFO with population size  $M$ ; search space  $n, [lb_x, ub_x]$ ; stop condition  $T$ ; initialization method  $Init_x$ , optimization method  $Opt_x$ , treatment  $C$  of outliers, and selection  $S$  to objective solutions; initialization method  $Init_{\Delta}$  and optimization method  $Opt_{\Delta}$  to step-size  $\Delta$ .

---

```

1:  $t \leftarrow 0$ 
2:  $\mathbf{X}(t) \leftarrow Init_x(n, M, [lb_x, ub_x])$  as Eq.4.27           ▷ initialize initial population
3:  $F(t) \leftarrow f(\mathbf{X}(t))$                                            ▷ evaluate
4:  $w \leftarrow Init_{\Delta:w}(w)$                                        ▷ initialize  $w$ -relative step-size
5:  $\mathbf{X}_s(t) \leftarrow Init_{\Delta:x_s}(\mathbf{X}(t))$  as Eq.4.28                 ▷ initialize  $x$ -relative step-size
6:  $z(t) \leftarrow Init_{\Delta:z}(t, T)$  as Eq.4.29, Eq.4.30             ▷ initialize  $z$ -relative step-size
7: while stop condition  $T$  do
8:    $\hat{\mathbf{X}}(t+1) \leftarrow Opt_x(\mathbf{X}(t), \mathbf{X}_s(t), z(t), w)$  as Eq.4.31   ▷ generate temporarily
     updated population
9:    $\hat{\mathbf{X}}(t+1) \leftarrow C(\hat{\mathbf{X}}(t+1))$  as Eq.4.35                   ▷ treatment to outliers
10:   $F(t+1) \leftarrow f(\hat{\mathbf{X}}_i(t+1))$                                    ▷ evaluate
11:   $\mathbf{X}(t+1) \leftarrow S(\mathbf{X}_i(t), \hat{\mathbf{X}}_i(t+1))$  as Eq.4.36       ▷ select and generate finally
     updated population
12:   $\mathbf{X}_s(t+1) \leftarrow Init_{\Delta:x_s}(\mathbf{X}(t), \mathbf{X}(t+1))$  as Eq.4.32 ▷ update dynamic  $x$ -relative
     vector step-size  $\mathbf{x}_s$ 
13:   $z_1(t+1) \leftarrow Opt_{\Delta:z_1}(t+1, T)$  as Eq.4.33, Eq.4.34 ▷ update  $z$ -relative step-size
14:   $t \leftarrow t + 1$ 
15: end while
    
```

---

### 4.2.5 Re-framed MBO

The *Tuple.1*  $f$  can be represented as  $F(t) = f(\mathbf{X}(t))$ . The *Tuple.2*  $Init_x$  can be extracted from [Eq.2.21](#), and re-framed as [Eq.4.37](#).

$$\mathbf{x}_i(t=0) = \mathcal{U}(lb_x, ub_x), i = 1, 2, \dots, M \quad 4.37$$

For the *Tuple.3*  $Init_{\Delta}$ , as discussed in [Table 2.1](#), we find the global best individual  $\mathbf{x}_{best}$  is a kind of  $x$ -relative step-size. Meanwhile, the weighting factor  $\alpha$  is a kind of  $z$ -relative



step-size. Therefore, we re-frame  $\mathbf{v}_i$  (see step (2)) into Eq.4.2. The  $\mathbf{x}_{best}$  (see step (8)) can be re-framed as Eq.4.38. The  $z$ -relative step-size  $\alpha$  can be re-framed as Eq.4.39.

$$\mathbf{x}_g(t=0) = \mathbf{Min}(\{\mathbf{x}_i(t=0)\}), i = 1 \dots M \quad 4.38$$

$$z(t=0) = \frac{w_4}{t^2} \quad 4.39$$

For the *Tuple.4 Opt<sub>x</sub>*, we find its update way (from step (3) to step (8)) is able to be re-framed into one equation as Eq.4.40, in which  $\mathbf{L\acute{e}vy}_{i,n} = \mathbf{L\acute{e}vy}(d, n, T)$  with  $d \sim \text{Exp}(2 \times T)$ .

$$\begin{aligned} \langle \mathbf{x}_i(t) \rangle &= \mathbf{Sort}(\{\mathbf{x}_i(t)\}), i = 1 \dots M \\ \text{strong} \hat{\mathbf{x}}_{i,n}(t+1) &= \begin{cases} \mathbf{x}_{j,n}(t) \in \langle \mathbf{x}_i(t) \rangle, j \in [1, M'] & , r \times w_2 \leq w_1 \\ \mathbf{x}_{j,n}(t) \in \langle \mathbf{x}_i(t) \rangle, j \in (M', M] & , \text{o.w.} \end{cases} \\ M' &= \lceil w_1 \times M \rceil \\ \text{weak} \hat{\mathbf{x}}_{i,n}(t+1) &= \begin{cases} \mathbf{x}_{g,n}(t), r \geq w_1 \\ \begin{cases} \mathbf{x}_{j,n}(t) + z(t) \times (\mathbf{L\acute{e}vy}_{i,n} - 0.5), j \in (M', M], r > w_3 \\ \mathbf{x}_{j,n}(t) \in \langle \mathbf{x}_i(t) \rangle, j \in (M', M], \text{o.w.} \end{cases} & , \text{o.w.} \end{cases} \\ \{\hat{\mathbf{x}}_i(t+1)\} &= \{\text{strong} \hat{\mathbf{x}}_i(t+1)\} \cup \{\text{weak} \hat{\mathbf{x}}_i(t+1)\} \end{aligned} \quad 4.40$$

For the *Tuple.5 Opt<sub>Δ</sub>* that corresponds to *Tuple.3 Init<sub>Δ</sub>*, we find the update way of  $\mathbf{x}_{best}$  (see step (8)) is able to be re-framed into as Eq.4.41. The update way of  $\alpha$  (see Eq.2.23) can be re-framed as Eq.4.42.

$$\mathbf{x}_g(t+1) = \mathbf{Min}(\mathbf{x}_g(t) \cup \{\mathbf{x}_i(t+1)\}), i = 1 \dots M \quad 4.41$$

$$z(t+1) = \frac{w_4}{(t+1)^2} \quad 4.42$$

For the *Tuple.6 C*, the step (9) can be re-framed as Eq.4.43.

$$\mathbf{x}_{i,n}^{\text{fixed}}(t+1) = \begin{cases} ub_x & , \mathbf{x}_{i,n}(t+1) > ub_x \\ \mathbf{x}_{i,n}(t+1) & , \text{o.w} \\ lb_x & , \mathbf{x}_{i,n}(t+1) < lb_x \end{cases} \quad 4.43$$

For the *Tuple.7 S*, the step (11) can be re-framed as Eq.4.44.

$$\begin{aligned} \langle \hat{\mathbf{x}}_i(t+1) \rangle &= \mathbf{Sort}(\{\hat{\mathbf{x}}_i(t+1)\}), i = 1 \dots M \\ \mathbf{x}_i(t+1) &\in \{\langle \hat{\mathbf{x}}_i(t+1) \rangle, i = 1 \dots M - w_5\} \cup \{\langle \mathbf{x}_i(t) \rangle, i = 1 \dots w_5\} \end{aligned} \quad 4.44$$

For the *Tuple.8*  $T$  which is a stop condition related to the maximum number of iteration.

Till now, all of other unknown parameters can be defined as  $w$ -relative step-size, and they belong to the *Tuple.3*  $Init_{\Delta}$ . It has common  $w$ -relative step-size  $w_1 = \frac{5}{12}$ ,  $w_2 = 1.2$ ,  $w_3 = \frac{5}{12}$ ,  $w_4 = 1$  and  $w_5 = 2$

Therefore, the complete pseudo-code of re-framed BA model can be organized as [Algorithm.21](#).

---

**Algorithm 6** Re-framed MBO

---

```

1:  $t \leftarrow 0$ 
2:  $\mathbf{X}(t) \leftarrow Init_{\mathbf{x}}(n, M, [lb_{\mathbf{x}}, ub_{\mathbf{x}}])$  as Eq.4.37 ▷ initialize initial population
3:  $F(t) \leftarrow f(\mathbf{X}(t))$  ▷ evaluate
4:  $w \leftarrow Init_{\Delta:w}(w)$  ▷ initialize  $w$ -relative step-size
5:  $\mathbf{x}_g(t) \leftarrow Init_{\Delta:\mathbf{x}}(\mathbf{X}(t))$  as Eq.4.38 ▷ initialize  $\mathbf{x}$ -relative step-size
6:  $z \leftarrow Init_{\Delta:z}(w)$  as Eq.4.39 ▷ initialize  $z$ -relative step-size
7: while stop condition  $T$  do
8:    $\hat{\mathbf{X}}(t+1) \leftarrow Opt_{\mathbf{x}}(\mathbf{X}(t), \mathbf{x}_g(t), z(t), w)$  as Eq.4.40 ▷ generate temporarily updated population
9:    $\hat{\mathbf{X}}(t+1) \leftarrow C(\hat{\mathbf{X}}(t+1))$  as Eq.4.43 ▷ treatment to outliers
10:   $F(t+1) \leftarrow f(\hat{\mathbf{X}}(t+1))$  ▷ evaluate
11:   $\mathbf{X}(t+1) \leftarrow S(\mathbf{X}(t), \hat{\mathbf{X}}(t+1))$  as Eq.4.44 ▷ select and generate finally updated population
12:   $\mathbf{x}_g(t+1) \leftarrow Init_{\Delta:\mathbf{x}}(\mathbf{X}(t), \mathbf{X}(t+1))$  as Eq.4.41 ▷ update  $\mathbf{x}$ -relative step-size
13:   $z(t+1) \leftarrow Opt_{\Delta:z}(z(t), t+1)$  as Eq.4.42 ▷ update  $z$ -relative step-size
14:   $t \leftarrow t+1$ 
15: end while
    
```

---

### 4.2.6 Re-framed BOA

The *Tuple.1*  $f$  can be represented as  $F(t) = f(\mathbf{X}(t))$ . The *Tuple.2*  $Init_{\mathbf{x}}$  can be extracted from [Eq.2.26](#), and re-framed as [Eq.4.45](#).

$$\mathbf{x}_i(t=0) = \mathcal{U}(lb_{\mathbf{x}}, ub_{\mathbf{x}}), i = 1, 2, \dots, M \quad 4.45$$

For the *Tuple.3*  $Init_{\Delta}$ , as discussed in [Table 2.1](#), we find the global best individual  $g_*$  is a kind of  $\mathbf{x}$ -relative step-size. Meanwhile, the *sensory\_modality* is a kind of  $z$ -relative step-size. Therefore, we re-frame the  $g_*$  (see step (3)) can be re-framed as [Eq.4.46](#). The  $z$ -relative step-size *sensory\_modality* can be re-framed as [Eq.4.47](#).

$$\mathbf{x}_g(t=0) = \mathbf{Min}(\{\mathbf{x}_i(t=0)\}), i = 1 \dots M \quad 4.46$$

$$z(t = 0) = z^0 \quad 4.47$$

For the *Tuple.4 Opt<sub>x</sub>*, we find step (4), step (5) and step (6) are able to be re-framed into one equation as [Eq.4.48](#), in which  $\mathbf{x}_j$  and  $\mathbf{x}_k$  are any two neighbors around  $\mathbf{x}_i$ .

$$\hat{\mathbf{x}}_i(t + 1) = \begin{cases} \mathbf{x}_i(t) + (r^2 \times \mathbf{x}_g(t) - \mathbf{x}_i(t)) \times z_1(t) \times f(\mathbf{x}_i(t))^{w_1} & , \quad r > w_2 \\ \mathbf{x}_i(t) + (r^2 \times \mathbf{x}_j(t) - \mathbf{x}_k(t)) \times z_1(t) \times f(\mathbf{x}_i(t))^{w_1} & , \quad \text{o.w} \end{cases} \quad 4.48$$

For the *Tuple.5 Opt<sub>Δ</sub>* that corresponds to *Tuple.3 Init<sub>Δ</sub>*, we find the update way of  $g_*$  (see step (9)) is able to be re-framed into as [Eq.4.49](#). The update way of *sensory\_modality* (see [Eq.2.30](#)) can be re-framed as [Eq.4.50](#).

$$\mathbf{x}_g(t + 1) = \text{Min}(\mathbf{x}_g(t) \cup \{\mathbf{x}_i(t + 1)\}), i = 1 \dots M \quad 4.49$$

$$z(t + 1) = z(t) + \frac{0.025}{z(t) \times T} \quad 4.50$$

For the *Tuple.6 C*, the step (7) can be re-framed as [Eq.4.51](#).

$$\mathbf{x}_{i,n}^{\text{fixed}}(t + 1) = \begin{cases} ub_x & , \quad \mathbf{x}_{i,n}(t + 1) > ub_x \\ \mathbf{x}_{i,n}(t + 1) & , \quad \text{o.w} \\ lb_x & , \quad \mathbf{x}_{i,n}(t + 1) < lb_x \end{cases} \quad 4.51$$

For the *Tuple.7 S*, the step (8) can be re-framed as [Eq.4.52](#).

$$\mathbf{x}_i(t + 1) = \begin{cases} \hat{\mathbf{x}}_i(t + 1) & , \quad f(\hat{\mathbf{x}}_i(t + 1)) < f(\mathbf{x}_i(t)) \\ \mathbf{x}_i(t) & , \quad \text{o.w} \end{cases} \quad 4.52$$

For the *Tuple.8 T* which is a stop condition related to the maximum number of iteration.

Till now, all of other unknown parameters can be defined as  $w$ -relative step-size, and they belong to the *Tuple.3 Init<sub>Δ</sub>*. It has common  $w$ -relative step-size  $w_1 = 0.1$  and  $w_2 = 0.8$ . It also has special  $w$ -relative step-size that is related to the initial value of  $z$ -relative step-size  $z^0 = 0.01$ .

Therefore, the complete pseudo-code of re-framed BA model can be organized as [Algorithm.7](#).

**Algorithm 7** Re-framed BOA

---

```

1:  $t \leftarrow 0$ 
2:  $\mathbf{X}(t) \leftarrow \text{Init}_{\mathbf{x}}(n, M, [lb_{\mathbf{x}}, ub_{\mathbf{x}}])$  as Eq.4.45 ▷ initialize initial population
3:  $F(t) \leftarrow f(\mathbf{X}(t))$  ▷ evaluate
4:  $w, z^0 \leftarrow \text{Init}_{\Delta:w}(w, z^0)$  ▷ initialize  $w$ -relative step-size
5:  $\mathbf{x}_g(t) \leftarrow \text{Init}_{\Delta:\mathbf{x}}(\mathbf{X}(t))$  as Eq.4.46 ▷ initialize  $\mathbf{x}$ -relative step-size
6:  $z(t) \leftarrow \text{Init}_{\Delta:z}(z^0)$  as Eq.4.47 ▷ initialize  $z$ -relative step-size
7: while stop condition  $T$  do
8:    $\hat{\mathbf{X}}(t+1) \leftarrow \text{Opt}_{\mathbf{x}}(\mathbf{X}(t), \mathbf{x}_g(t), z(t), w)$  as Eq.4.48 ▷ generate temporarily updated population
9:    $\hat{\mathbf{X}}(t+1) \leftarrow C(\hat{\mathbf{X}}(t+1))$  as Eq.4.51 ▷ treatment to outliers
10:   $F(t+1) \leftarrow f(\hat{\mathbf{X}}(t+1))$  ▷ evaluate
11:   $\mathbf{X}(t+1) \leftarrow S(\mathbf{X}(t), \hat{\mathbf{X}}(t+1))$  as Eq.4.52 ▷ select and generate finally updated population
12:   $\mathbf{x}_g(t+1) \leftarrow \text{Init}_{\Delta:\mathbf{x}}(\mathbf{X}(t), \mathbf{X}(t+1))$  as Eq.4.49 ▷ update  $\mathbf{x}$ -relative step-size
13:   $z(t+1) \leftarrow \text{Opt}_{\Delta:z}(z(t), t+1)$  as Eq.4.50 ▷ update  $z$ -relative step-size
14:   $t \leftarrow t+1$ 
15: end while

```

---

**4.2.7 Re-framed PSO**

The *Tuple.1*  $f$  can be represented as  $F(t) = f(\mathbf{X}(t))$ . The *Tuple.2*  $\text{Init}_{\mathbf{x}}$  can be extracted from Eq.2.31, and re-framed as Eq.4.53.

$$\mathbf{x}_i(t=0) = \mathcal{U}(lb_{\mathbf{x}}, ub_{\mathbf{x}}), i = 1, 2, \dots, M \quad 4.53$$

For the *Tuple.3*  $\text{Init}_{\Delta}$ , as discussed in Table 2.1, we find the velocity  $\mathbf{v}_i$  is a kind of  $y$ -relative step-size. The personal best individual  $pbest$  and the global best individual  $gbest$  are a kind of  $\mathbf{x}$ -relative step-size. Therefore, we re-frame  $\mathbf{v}_i$  (see step (2)) into Eq.4.54. The  $pbest$  (see Eq.2.33) and the  $gbest$  (see step (5)) can be respectively re-framed as Eq.4.55 and Eq.4.56.

$$\mathbf{y}_i(t=0) = \mathcal{U}(lb_{\mathbf{y}}, ub_{\mathbf{y}}), i = 1 \dots M \quad 4.54$$

$$\mathbf{x}_{i_p}(t=0) = \mathbf{x}_i(t=0), i = 1 \dots M \quad 4.55$$

$$\mathbf{x}_g(t=0) = \mathbf{Min}(\{\mathbf{x}_i(t=0)\}), i = 1 \dots M \quad 4.56$$

For the *Tuple.4*  $\text{Opt}_{\mathbf{x}}$ , we find Eq.2.35 is able to be re-framed as Eq.4.57.

$$\hat{\mathbf{x}}_i(t+1) = \mathbf{x}_i(t) + \mathbf{y}_i(t+1) \quad 4.57$$

For the *Tuple.5*  $Opt_{\Delta}$  that corresponds to *Tuple.3*  $Init_{\Delta}$ , we find the update way of  $\mathbf{v}_i$  (see Eq.2.34) is able to be re-framed into as Eq.4.58. The update way of  $pbest$  (see step (9)) is able to be re-framed into as Eq.4.59. The update way of  $gbest$  (see Eq.2.3) can be re-framed as Eq.4.60.

$$\mathbf{y}_i(t+1) = w_1 \times \mathbf{y}_i(t) + \mathcal{U}(0, w_2) \times (\mathbf{x}_{i_p}(t) - \mathbf{x}_i(t)) + \mathcal{U}(0, w_3) \times (\mathbf{x}_g(t) - \mathbf{x}_i(t)) \quad 4.58$$

$$\mathbf{x}_{i_p}(t+1) = \mathbf{Min}(\{\mathbf{x}_{i_p}(t), \mathbf{x}_i(t+1)\}) \quad 4.59$$

$$\mathbf{x}_g(t+1) = \mathbf{Min}(\mathbf{x}_g(t) \cup \{\mathbf{x}_i(t+1)\}), i = 1 \dots M \quad 4.60$$

For the *Tuple.6*  $C$ , the step (7) can be re-framed as Eq.4.61.

$$\mathbf{x}_{i,n}^{\text{fixed}}(t+1) = \begin{cases} ub_x & , \mathbf{x}_{i,n}(t+1) > ub_x \\ \mathbf{x}_{i,n}(t+1) & , \text{o.w} \\ lb_x & , \mathbf{x}_{i,n}(t+1) < lb_x \end{cases} \quad 4.61$$

For the *Tuple.7*  $S$ , the PSO model doesn't have a process to select which individuals will be delivered into the next generation, therefore, we re-frame it as Eq.4.62 that means all newly generated individuals will be delivered into the next generation.

$$\mathbf{x}_i(t+1) = \hat{\mathbf{x}}_i(t+1) \quad 4.62$$

For the *Tuple.8*  $T$  which is a stop condition related to the maximum number of iteration.

Till now, all of other unknown parameters can be defined as  $w$ -relative step-size, and they belong to the *Tuple.3*  $Init_{\Delta}$ . It has common  $w$ -relative step-size  $w_1 = 0.73$ ,  $w_2 = 1.49$ ,  $w_3 = 1.49$ .

Therefore, the complete pseudo-code of re-framed BA model can be organized as Algorithm.8.

---

**Algorithm 8** Re-framed PSO

---

```

1:  $t \leftarrow 0$ 
2:  $\mathbf{X}(t) \leftarrow \text{Init}_{\mathbf{x}}(n, M, [lb_{\mathbf{x}}, ub_{\mathbf{x}}])$  as Eq.4.53 ▷ initialize initial population
3:  $F(t) \leftarrow f(\mathbf{X}(t))$  ▷ evaluate
4:  $w, [lb_{\mathbf{y}}, ub_{\mathbf{y}}] \leftarrow \text{Init}_{\Delta:w}(w, [lb_{\mathbf{y}}, ub_{\mathbf{y}}])$  ▷ initialize  $w$ -relative step-size
5:  $\mathbf{Y}(t) \leftarrow \text{Init}_{\Delta:\mathbf{y}}(n, M, [lb_{\mathbf{y}}, ub_{\mathbf{y}}])$  as Eq.4.54 ▷ initialize  $\mathbf{y}$ -relative step-size
6:  $\mathbf{X}_p(t), \mathbf{x}_g(t) \leftarrow \text{Init}_{\Delta:\mathbf{x}}(\mathbf{X}(t))$  as Eq.4.55, Eq.4.56 ▷ initialize  $\mathbf{x}$ -relative step-size
7: while stop condition  $T$  do
8:    $\mathbf{Y}(t+1) \leftarrow \text{Opt}_{\Delta:\mathbf{y}}(\mathbf{Y}(t))$  as eeqq:ps05 ▷ update  $\mathbf{y}$ -relative step-size
9:    $\hat{\mathbf{X}}(t+1) \leftarrow \text{Opt}_{\mathbf{x}}(\mathbf{X}(t), \mathbf{Y}(t), \mathbf{x}_g(t), \mathbf{X}_p(t), w)$  as Eq.4.57 ▷ generate temporarily updated population
10:   $\hat{\mathbf{X}}(t+1) \leftarrow C(\hat{\mathbf{X}}(t+1))$  as Eq.4.60 ▷ treatment to outliers
11:   $F(t+1) \leftarrow f(\hat{\mathbf{X}}(t+1))$  ▷ evaluate
12:   $\mathbf{X}(t+1) \leftarrow S(\mathbf{X}(t), \hat{\mathbf{X}}(t+1))$  as Eq.4.61 ▷ select and generate finally updated population
13:   $\mathbf{X}_p(t+1), \mathbf{x}_g(t+1) \leftarrow \text{Init}_{\Delta:\mathbf{x}}(\mathbf{X}(t), \mathbf{X}(t+1))$  as Eq.4.60, Eq.4.61 ▷ update  $\mathbf{x}$ -relative step-size
14:   $t \leftarrow t + 1$ 
15: end while

```

---

## 4.3 Unified Nature-Inspired Optimization Algorithm (UNIOA)

### 4.3.1 Unified Framework — UNIOA

In this section, we introduce the Unified framework for Nature-inspired Optimization Algorithm — UNIOA that is the end goal of this work. As discussed in [Chapter 2](#), [Chapter 3](#) and [Section 4.2](#), we can already be sure that these seven selected algorithms can be re-framed from their original different terminologies in different procedures to our unified terminologies in eight same tuples whose positions can also be same.

We summarize our observations in [Algorithm.9](#). In the Initialization Process, these seven algorithms can be completely same. For example, these algorithms need an initial population with the fitness of this population and also need several step-size  $\Delta$  to trigger the optimization mechanism. Different algorithms can have different step-size  $\Delta$ . In the Optimization Process, there are two options — the algorithm uses  $\mathbf{y}$ -relative step-size  $\Delta$  or the algorithm doesn't use  $\mathbf{y}$ -relative step-size  $\Delta$ . If the algorithm uses the  $\mathbf{y}$ -relative

step-size  $\Delta$ , the  $y$ -relative  $\Delta$  shall update before updating the population  $\mathbf{X}$ . If the algorithm doesn't use the  $y$ -relative step-size  $\Delta$ , all dynamic  $\Delta$  can update together after the selection operation. Meanwhile,  $z$ -relative  $\Delta$  has flexible positions when modeling a valid algorithm. For example,  $z$ -relative  $\Delta$  can be used in  $S$  selection method in BA.

---

**Algorithm 9** Unified Nature-Inspired Optimization Algorithm — UNIOA
 

---

```

1:  $t \leftarrow 0$ 
2:  $\mathbf{X}(t) \leftarrow \text{Init}_{\mathbf{x}}(n, M, [lb_{\mathbf{x}}, ub_{\mathbf{x}}])$  ▷ initialize initial pop
3:  $F(t) \leftarrow f(\mathbf{X}(t))$  ▷ evaluate
4:  $\Delta_{w,z,y,x}(t) \leftarrow \text{Init}_{\Delta:w,z,y,x}(\mathbf{X}(t), w, z^0, [lb_{\mathbf{y}}, ub_{\mathbf{y}}], t)$  ▷ initialize step-size
5: while  $T$  do
6:   if  $y \in \Delta(t=0)$  then
7:      $\mathbf{Y}(t+1) \leftarrow \text{Opt}_{\Delta:y}(\mathbf{Y}(t), w)$  ▷ update  $y$ -relative step-size
8:      $\hat{\mathbf{X}}(t+1) \leftarrow \text{Opt}_{\mathbf{x}}(\mathbf{X}(t), \mathbf{Y}(t+1), \Delta_{w,z,x}(t))$  ▷ temporarily updated pop
9:      $\hat{\mathbf{X}}(t+1) \leftarrow C(\hat{\mathbf{X}}(t+1))$  ▷ outliers treatment
10:     $F(t+1) \leftarrow f(\hat{\mathbf{X}}(t+1))$  ▷ evaluate
11:     $\mathbf{X}(t+1) \leftarrow S(\mathbf{X}(t), \hat{\mathbf{X}}(t+1), \Delta_{w,z}(t))$  ▷ select and generate finally updated
    pop
12:     $\Delta_{z,x}(t+1) \leftarrow \text{Opt}_{\Delta:z,x}(\mathbf{X}(t), \mathbf{X}(t+1), z(t), t+1)$  ▷ update  $z, x$ -relative
    step-size
13:   else
14:      $\hat{\mathbf{X}}(t+1) \leftarrow \text{Opt}_{\mathbf{x}}(\mathbf{X}(t), \Delta(t))$  ▷ temporarily updated pop
15:      $\hat{\mathbf{X}}(t+1) \leftarrow C(\hat{\mathbf{X}}(t+1))$  ▷ outliers treatment
16:      $F(t+1) \leftarrow f(\hat{\mathbf{X}}(t+1))$  ▷ evaluate
17:      $\mathbf{X}(t+1) \leftarrow S(\mathbf{X}(t), \hat{\mathbf{X}}(t+1), \Delta_{w,z}(t))$  ▷ select and generate finally updated
    pop
18:      $\Delta(t+1) \leftarrow \text{Opt}_{\Delta}(\mathbf{X}(t), \mathbf{X}(t+1), \Delta(t), t+1)$  ▷ update step-size
19:   end if
20:    $t \leftarrow t + 1$ 
21: end while
    
```

---

We recall the 8-tuple function (see Eq.3.1) defined in Chapter 3, and extend it as Eq.4.63 with the new name — *UNIOA*.

$$UNIOA = (f, \text{Init}_{\mathbf{x}}, \text{Opt}_{\mathbf{x}}, C, T, S, \text{Init}_{\Delta}, \text{Opt}_{\Delta}) \quad 4.63$$

Furthermore, according to observations in Section 4.2, these eight tuples have following different options:

(1) *Tuple.1 f*:

The *Tuple.1 f* exists in all these seven algorithms. Its function in these algorithms is also same as  $f(\mathbf{X})$  that calculates the fitness of the whole population  $X$ .

(2) *Tuple.2 Init<sub>x</sub>*:

The *Tuple.2 Init<sub>x</sub>* exists in all these seven algorithms. Its function is also same as *Init<sub>x</sub>*.

$$\mathbf{x}_i = \mathcal{U}(lb_{\mathbf{x}}, ub_{\mathbf{x}}), i = 1, 2, \dots, M \quad \text{Init}_{\mathbf{x}}$$

(3) *Tuple.3 Init<sub>Δ</sub>*:

The *Tuple.3 Init<sub>Δ</sub>* exists in all these seven algorithms. But not all kinds of step-size  $\Delta$  are present in all algorithms at the same time. There are four kinds of step-size  $\Delta$ :  $w$ -relative  $\Delta$ ,  $z$ -relative  $\Delta$ ,  $x$ -relative  $\Delta$  and  $y$ -relative  $\Delta$ .

(a)  $w$ -relative  $\Delta$ :

This kind of step-size doesn't change during the optimization process. There are three kinds of  $w$ -relative  $\Delta$ :  $w, z^0, [lb_{\mathbf{y}}, ub_{\mathbf{y}}]$ . The  $w$  is commonly used in most algorithms and its common name is hyper-parameter. The  $z^0$  and  $[lb_{\mathbf{y}}, ub_{\mathbf{y}}]$  are respectively initial values of  $z$ -relative  $\Delta$  and  $y$ -relative  $\Delta$ .

(b)  $z$ -relative  $\Delta$ :

The  $z$  is a kind of hyper-parameter that will be changing over the iteration  $t$  during the optimization process. Its initialization method could be varying (see Index.4 in Table 4.1).

(c)  $x$ -relative  $\Delta$ :

- $\mathbf{x}_{i_p}$  is the best  $\mathbf{x}_i$  that each  $\mathbf{x}_i$  has found so far. Its initialization method is fixed as *Init<sub>Δ: x<sub>i<sub>p</sub></sub></sub>*.

$$\mathbf{x}_{i_p}(t) = \mathbf{x}_i(t), i = 1 \dots M \quad \text{Init}_{\Delta: \mathbf{x}_{i_p}}$$

- $\mathbf{x}_g$  is the best  $\mathbf{x}_i$  that the whole population has found so far. Its initialization method is fixed as *Init<sub>Δ: x<sub>g</sub></sub>*.

$$\mathbf{x}_g(t) = \mathbf{Min}(\{\mathbf{x}_i(t)\}), i = 1 \dots M \quad \text{Init}_{\Delta: \mathbf{x}_g}$$

- $\mathbf{x}_s$  denotes one kind of special  $x$ -relative  $\Delta$  that has a strong connection with the  $\mathbf{x}_i$ . It depends on different algorithms, so it is varying. For example, Eq.4.28 in MFO.



(d)  $y$ -relative  $\Delta$ :

In this work, there is only one kind of initialization method as  $Init_{\Delta:y}$ . In this work, its initialization method is fixed, however, it could be customized in the future.

$$y_i(t) = \mathcal{U}(lb_y, ub_y), i = 1 \dots M \quad Init_{\Delta:y}$$

(4) *Tuple.4*  $Opt_x$ :

The *Tuple.4*  $Opt_x$  exists in all these seven algorithms. Its method is varying (see Index.7 in Table 4.1).

(5) *Tuple.5*  $Opt_{\Delta}$ :

The *Tuple.5*  $Opt_{\Delta}$  that corresponds to *Tuple.3*  $Init_{\Delta}$  exists in all these seven algorithms. *Tuple.5*  $Opt_{\Delta}$  can locate at two positions: Index.6 or Index.11 in Table 4.1.

(a)  $y$ -relative  $\Delta$  at Index.6 position:

The update method to  $y$ -relative  $\Delta$  is varying in different algorithms (see Index.6 in Table 4.1)

(b)  $z$ -relative and  $x$ -relative  $\Delta$  at Index.11 position:

The update method of  $z$ -relative  $\Delta$  is varying in different algorithms (see Index.11 in Table 4.1). However, the update method of two  $x$ -relative  $\Delta$  is fixed as  $Opt_{\Delta:x_{ip}}$  and  $Opt_{\Delta:x_g}$ . The update method of  $x$ -relative  $\Delta$  can also be customized, for example, the  $Opt_{\Delta:x_s}$ : Eq.4.28 in MFO.

$$x_{ip}(t+1) = \text{Min}(\{x_{ip}(t), x_i(t+1)\}) \quad Opt_{\Delta:x_{ip}}$$

$$x_g(t+1) = \text{Min}(x_g(t) \cup \{x_i(t+1)\}), i = 1 \dots M \quad Opt_{\Delta:x_g}$$

(6) *Tuple.6*  $C$ :

The *Tuple.6*  $C$  exists in all algorithms. There are two methods to deal with outliers in this work. For example,  $C1$  is applied in BA, GOA, MFO, MBO, BOA and PSO,  $C2$  is only applied in CSA.<sup>2</sup>

$$x_{i,n}^{\text{fixed}}(t+1) = \begin{cases} ub_x & , \quad x_{i,n}(t+1) > ub_x \\ x_{i,n}(t+1) & , \quad \text{o.w} \\ lb_x & , \quad x_{i,n}(t+1) < lb_x \end{cases} \quad C1$$

<sup>2</sup>Some algorithms didn't mention  $C$  in their published paper or code, however, it is acceptable that outliers are invalid and must be dealt with.

$$\mathbf{x}_{i,n}^{\text{fixed}}(t+1) = \begin{cases} \mathbf{x}_{i,n}(t) & , \mathbf{x}_{i,n}(t+1) < lb_{\mathbf{x}} \text{ or } \mathbf{x}_{i,n}(t+1) > ub_{\mathbf{x}} \\ \mathbf{x}_{i,n}(t+1) & , \text{o.w} \end{cases} \quad \text{C2}$$

(7) *Tuple.7 S.*

The *Tuple.7 S* exists in all algorithms. There are four selection methods in this work. For example, [S1](#) is applied in GOA, CSA, MFO and PSO, [S2](#) is applied in BOA, [S3](#) is applied in BA, [S4](#) is applied in MBO.

$$\mathbf{x}_i(t+1) = \hat{\mathbf{x}}_i(t+1) \quad \text{S1}$$

$$\mathbf{x}_i(t+1) = \begin{cases} \hat{\mathbf{x}}_i(t+1) & , f(\hat{\mathbf{x}}_i(t+1)) < f(\mathbf{x}_i(t)) \\ \mathbf{x}_i(t) & , \text{o.w} \end{cases} \quad \text{S2}$$

$$\mathbf{x}_i(t+1) = \begin{cases} \hat{\mathbf{x}}_i(t+1) & , f(\hat{\mathbf{x}}_i(t+1)) < f(\mathbf{x}_i(t)) \text{ or } rand > z_2(t) \\ \mathbf{x}_i(t) & , \text{o.w} \end{cases} \quad \text{S3}$$

$$\begin{aligned} \langle \hat{\mathbf{x}}_i(t+1) \rangle &= \text{Sort}(\{\hat{\mathbf{x}}_i(t+1)\}), i = 1 \dots M \\ \mathbf{x}_i(t+1) &\in \{\langle \hat{\mathbf{x}}_i(t+1) \rangle, i = 1 \dots M - w_5\} \cup \{\langle \mathbf{x}_i(t) \rangle, i = 1 \dots w_5\} \end{aligned} \quad \text{S4}$$

(8) *Tuple.8 T.*

The *T* determines if the iterative optimization process would stop or not. It must exist.

We also summarize these eight tuples in [Table 4.1](#), in which each algorithm has its own representations of each tuple, but these seven algorithms can also have same representations in some tuples.

### 4.3.2 UNIOA Package

The observations in [Subsection 4.3.1](#) are visualized in [Figure 4.1](#), in which seven selected algorithms can be re-written in one same framework. In this unified framework — UNIOA, tuples for making up algorithms are fixed, and the positions of these tuples are also fixed. We also find these algorithms perform differently mainly because of the different designs of these tuples.

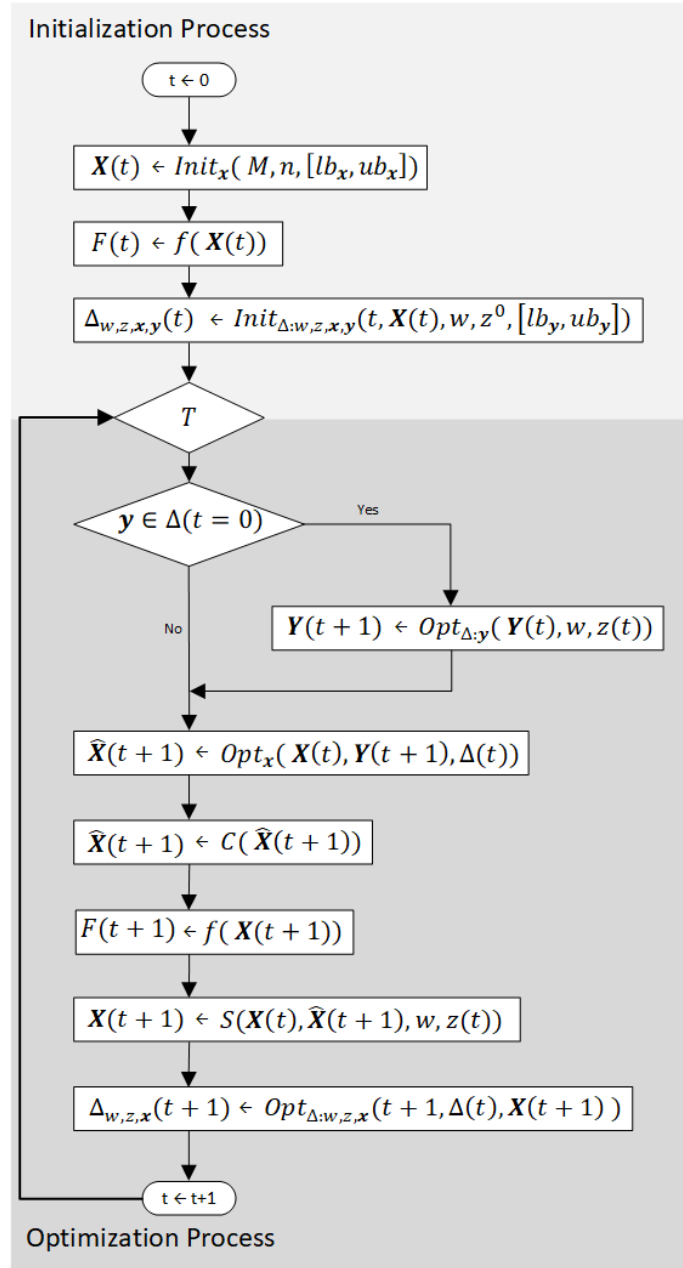


Figure 4.1: Unified framework for Nature-Inspired Optimization Algorithm —UNIOA, only for seven selected algorithms

Therefore, as shown in Table 4.2, we discuss the possibility of each tuple being customized in the future. In a complete optimization algorithm, *Tuple.1*  $f$ , *Tuple.2*  $Init_x$ , *Tuple.3*  $Init_{\Delta:x_{ip},x_g}$  and *Tuple.5*  $Init_{\Delta:x_{ip},x_g}$  are not allowed to be customized. *Tuple.1*  $f$  and *Tuple.2*  $Init_x$  are coming from the objective optimization problem. The way of initializing and updating two special  $x$ -relative step-size  $\Delta$  (*Tuple.3*  $Init_{\Delta:x_{ip},x_g}$  and *Tuple.5*  $Init_{\Delta:x_{ip},x_g}$ ) is also fixed in general.<sup>3</sup> For other tuples (*Tuple.3*  $Init_{\Delta:z,y,x_s}$ , *Tuple.4*  $Opt_x$ , *Tuple.5*  $Init_{\Delta:z,y,x_s}$ , *Tuple.6*  $C$  and *Tuple.7*  $S$ ), they can be freely designed by users.

The general utilization of this observation is to design an auto-designer for algorithms. For example, there are many open Python libraries [4, 6, 9, 13, 23, 28] for developing evolutionary and genetic algorithms (EA). Various 'crossover,' 'mutation,' and 'selection' are freely combined together to build up new algorithms. However, such application might not be suitable in UNIOA for swarm-based algorithms (SI).

The most significant difference between the unified framework for EA and the unified framework for SI is whether customers can freely combine different options of different tuples to build up a new algorithm. In the EA group, it is possible to freely combine different options from different tuples because every tuple operation happens to the entire population. For example, in the most popular unified framework for EA, the entire population is fed into each tuple one by one. The output of each tuple is still the entire population. However, in the SI group, this design idea is impossible. For example, in the Table 4.2, the *Tuple.3*  $Init_{\Delta}$  has six options, and the *Tuple.4*  $Opt_x$  has seven options. However, it doesn't mean any option of *Tuple.3*  $Init_{\Delta}$  can be combined with any option of *Tuple.4*  $Opt_x$  to make up a new algorithm. The reason is *Tuple.3*  $Init_{\Delta}$  and *Tuple.4*  $Opt_x$  are one-to-one correspondence.<sup>4</sup>

The most significant difference between the unified framework for EA and the unified framework for SI is whether customers can freely combine different options of different tuples to build up a new algorithm. In the EA group, it is possible to freely combine different options from different tuples because every tuple operation happens to the entire population. For example, in the most popular unified framework for EA, the entire population is fed into each tuple one by one. The output of each tuple is still the entire population. However, in the SI group, this design idea is impossible. For example, in the

---

<sup>3</sup>Step-size  $\Delta : x_{ip}, x_g$  might also be customized when people have different understanding about what is the personal best solution and what is the global best solution.

<sup>4</sup>This difference can also be understood as: SI has more complicated optimization mechanisms compared to EA. For example, gene operations in EA are limited, but social activities for designing SI in the real world are very various.

Table 4.2, the *Tuple.3*  $Init_{\Delta}$  has six options, and the *Tuple.4*  $Opt_x$  has seven options. However, it doesn't mean any option of *Tuple.3*  $Init_{\Delta}$  can be combined with any option of *Tuple.4*  $Opt_x$  to make up a new algorithm. The reason is *Tuple.3*  $Init_{\Delta}$  and *Tuple.4*  $Opt_x$  are one-to-one correspondence.<sup>5</sup>

Therefore, although in the EA framework, customers are allowed to build up a new algorithm by freely selecting different options from existing methods, customers don't have much freedom in the UNIOA framework. However, the UNIOA gives customers many possibilities to design a totally new algorithm by only considering mathematical equations. For example, users shall start designing *Tuple.4*  $Opt_x$  that determines how many and what kinds of step-size  $\Delta$  will appear in this algorithm. *Tuple.6*  $C$  and *Tuple.7*  $S$  are then customized to help achieve the *Tuple.4*  $Opt_x$ .

The idea of auto-designing SI in UNIOA is basically achieved in UNIOA<sup>6</sup> in which users can design their iterative optimization algorithms only with primitive math knowledge and without any biological knowledge. Moreover, in UNIOA, the customized algorithm can measure its performance or compare its performance with the other seven existing algorithms (studied in this work) with the help of the IOHprofiler environment.

---

<sup>5</sup>This difference can also be understood as: SI has more complicated optimization mechanisms compared to EA. For example, gene operations in EA are limited, but social activities for designing SI in the real world are very various.

<sup>6</sup><https://github.com/Huilin-Li/UNIOA>

Table 4.1: Representations of eight tuples in each algorithm.

Progress	Index	Tuple		BA	GOA	CSA	MFO	MBO	BOA	PSO	
Initialization	1	t=0									
	2	Tuple.2 $Init_{\mathbf{x}}$			$Init_{\mathbf{x}}$						
	3	Tuple.1 $f$			$f(\mathbf{x}(t))$						
	4	Tuple.3 $Init_{\Delta}$	$\mathbf{y}_i$	$Init_{\Delta:\mathbf{y}}$	$\times$	$\times$	$\times$	$\times$	$\times$	$Init_{\Delta:\mathbf{y}}$	
			$\mathbf{x}_{i_p}$	$\times$	$\times$	$Init_{\Delta:\mathbf{x}_{i_p}}$	$\times$	$\times$	$\times$	$Init_{\Delta:\mathbf{x}_{i_p}}$	
			$\mathbf{x}_g$	$Init_{\Delta:\mathbf{x}_g}$			$\times$	$\times$	$Init_{\Delta:\mathbf{x}_g}$		
			$\mathbf{x}_s$	$\times$	$\times$	$\times$	Eq.4.28	$\times$	$\times$	$\times$	
			$z$	Eq.4.3, Eq.4.4	Eq.4.14	$\times$	Eq.4.29, Eq.4.29	Eq.4.38	Eq.4.46	$\times$	
$w$			$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$	$\checkmark$		
Optimization	5	Tuple.8 Stop condition $T$									
	6	Tuple.5 $Opt_{\Delta:\mathbf{y}}$	$\mathbf{y}_i(t+1)$	Eq.4.5	$\times$	$\times$	$\times$	$\times$	$\times$	Eq.4.54	
	7	Tuple.4 $Opt_{\mathbf{x}}$		Eq.4.6	Eq.4.15	Eq.4.22	Eq.4.31	Eq.4.39	Eq.4.47	Eq.4.55	
	8	Tuple.6 $C$		$C1$		$C2$	$C1$				
	9	Tuple.1 $f$		$f(\hat{\mathbf{x}}_i(t+1))$							
	10	Tuple.7 $S$		$S3$	$S1$			$S4$	$S2$	$S1$	
	11	Tuple.5 $Opt_{\Delta:z,\mathbf{x}}$	$z(t+1)$	Eq.4.7, Eq.4.8	Eq.4.16	$\times$	Eq.4.33, Eq.4.34	Eq.4.41	Eq.4.48	$\times$	
			$\mathbf{x}_{i_p}(t+1)$	$\times$	$\times$	$Opt_{\Delta:\mathbf{x}_{i_p}}$	$\times$	$\times$	$\times$	$Opt_{\Delta:\mathbf{x}_{i_p}}$	
			$\mathbf{x}_g(t+1)$	$Opt_{\Delta:\mathbf{x}_g}$			$\times$	$\times$	$Init_{\Delta:\mathbf{x}_g}$		
			$\mathbf{x}_s(t+1)$	$\times$	$\times$	$\times$	Eq.4.32	$\times$	$\times$	$\times$	
	12	t=t+1									

Table 4.2: In UNIOA, It is possible to customize *Tuple.*, but it is impossible to customize

	Index	Tuple		Existing Method	Customization
Initialization	1	t=0			
	2	Tuple.2 $Init_x$		$Init_x$	no
	3	Tuple.1 $f$		$f(\mathbf{x}_i(t)), i = 1, 2, \dots, M$	no
	4	Tuple.3 $Init_\Delta$	$\mathbf{y}_i$	$Init_{\Delta:\mathbf{y}}$	yes
			$\mathbf{x}_{i_p}$	$Init_{\Delta:\mathbf{x}_{i_p}}$	no
			$\mathbf{x}_g$	$Init_{\Delta:\mathbf{x}_g}$	no
			$\mathbf{x}_s$	Eq.4.28(MFO)	yes
			$z$	Eq.4.3, Eq.4.4(BA). Eq.4.14(GOA). Eq.4.29, Eq.4.30(MFO). Eq.4.38(MBO). Eq.4.46(BOA)	yes
$w$	commonly known as hyper-parameters			yes	
Optimization	5	Tuple.8 Stop condition $T$			yes
	6	Tuple.5 $Opt_\Delta$	$\mathbf{y}_i(t+1)$	Eq.4.5(BA). Eq.4.54(PSO)	yes
	7	Tuple.4 $Opt_x$		Eq.4.6(BA). Eq.4.15(GOA). Eq.4.31(MFO). Eq.4.39(MBO). Eq.4.47(BOA). Eq.4.55(PSO) Eq.4.22(CSA)	yes
	8	Tuple.6 $C$		C1, C2	yes
	9	Tuple.1 $f$		$f(\hat{\mathbf{x}}_i(t+1))$	no
	9	$S \rightarrow \mathbf{x}_i(t+1)$		$S1, S2, S3, S4$	yes
	11	Tuple.5 $Opt_\Delta$	$z(t+1)$	Eq.4.7, Eq.4.8(BA). Eq.4.16(GOA). Eq.4.33, Eq.4.34(MFO). Eq.4.41(MBO). Eq.4.48(BOA)	yes
			$\mathbf{x}_{i_p}(t+1)$	$Opt_{\Delta:\mathbf{x}_{i_p}}$	no
			$\mathbf{x}_g(t+1)$	$Opt_{\Delta:\mathbf{x}_g}$	no
			$\mathbf{x}_s(t+1)$	Eq.4.32(MFO)	yes
	12	t=t+1			

## 4.4 Summary

In this chapter, we give our Unified Framework — UNIOA that is the end goal in this work. We firstly re-write these seven algorithms in UNIOA in [Section 4.2](#), in which each algorithm is re-written in UNIOA with eight tuples whose positions are also fixed. Meanwhile, we attempt to use primitive math knowledge to represent each tuple, which means basic math knowledge is enough to understand these swarm-based algorithms. Re-framed pseudo-code of each algorithm is also displayed in this section. The detailed representations of each tuple in each algorithm is shown in [Table 4.1](#).

In [Section 4.3](#), we discuss UNIOA deeply. The pseudo-code of the UNIOA is shown as [Algorithm.9](#), in which we conclude the cooperation among these eight tuples. In [Subsection 4.3.2](#), the visualization [Figure 4.1](#) inspires the auto-designer UNIOA Package for swarm-based algorithms. Furthermore, we discuss the difference between the auto-designer for EA and the auto-designer UNIOA for swarm-based algorithms.

In conclusion, the Unified Framework UNIOA is constructed well in this chapter. based on seven selected algorithms.