Swarm and Evolutionary Computation

An enhanced Kalman filtering and historical learning mechanism driven estimation of distribution algorithm

--Manuscript Draft--

Manuscript Number:	
Article Type:	Full Length Article
Keywords:	Estimation of distribution algorithm; Kalman filtering; Historical learning mechanism; Adaptive adjustment strategy; The elite strategy
Corresponding Author:	fuqing Zhao, Ph.D Lanzhou University of Technology Lanzhou, CHINA
First Author:	Ningning Zhu
Order of Authors:	Ningning Zhu
	fuqing Zhao, Ph.D
	Ling Wang
	Chenxin Dong
Abstract:	As an evolutionary algorithm based on probabilistic models, estimation of distribution algorithm (EDA) utilizes the information of promising solutions to solve complex continuous optimization problems. However, the traditional EDA is confronted with the limitations that the search direction and scope are not conducive to the evolution process, which lessens the efficiency of the search. An enhanced Kalman filtering and historical learning mechanism driven EDA (KFHLEDA) is proposed to improve the search efficiency and the modeling accuracy of traditional EDA in this study. The enhanced Kalman filtering covers the stages of the prediction, the observation, and the first and the second revision in allusion to specific problems during the search, increasing population diversity and modeling accuracy. The elite strategy is employed to modify the revision improvement matrix, which accelerates the process of evolution. The enhanced revision information is fed back to the probabilistic model through the history learning mechanism to ameliorate the modeling accuracy. The population adaptive adjustment strategy reduces the number of invalid iterations. KFHLEDA improves the search direction and expands the search scope after the theoretical analysis. Compared with several classical metaheuristic algorithms and the state-of-the-art EDA variants, the evaluation results on benchmark functions of CEC 2017 test suit indicate that KFHLEDA is effective and competitive.
Suggested Reviewers:	Zhongshi Shao, Ph.D Shaanxi Normal University shaozhongshi@hotmail.com Dr.Shao shares the same research interests with me on the meta-heustic andcomputation intelligence.
	Weishi Shao, Ph.D Nanjing Normal University shaoweishi@hotmail.com Dr. Shao does the reasearch work on the computation intelligence and its application in the complex systems, and has a good reputation on the optimization algorithms.
	Shuo Qin, Ph.D Nanjing University of Aeronautics and Astronautics 863359486@qq.com Dr. Qin does the reasearch work on the computation intelligence and swarm intelligence optimization in the complex systems, and has a good reputation on the optimization algorithms.
Opposed Reviewers:	

Fuqing Zhao
School of Computer and Communication
Lanzhou University of Technology
Lanzhou, Gansu, 730050
P.R.China
November 05, 2022

Dear editor of Swarm and Evolutionary Computation,

Here within enclosed is our paper for consideration to be published on Swarm and Evolutionary Computation.

The further information about the paper is in the following:

The title: An enhanced Kalman filtering and historical learning mechanism driven estimation of distribution algorithm

The Authors: Ningning Zhu, Fuqing Zhao, Ling Wang, Chenxin Dong

The authors claim that none of the material in the paper has been published or is under consideration for publication elsewhere.

I am the corresponding author and my address and other information is as follows,

School of Computer and Communication
Lanzhou University of Technology
Lanzhou, Gansu,730050
P.R.China

Email: fzhao2000@hotmail.com
Fax: 86-931-2973901
Tel.: 86-931-1391977005

Thank you very much for consideration!

Sincerely Yours,
Dr. Fuqing Zhao

Highlights

An enhanced Kalman filtering and historical learning mechanism driven estimation of distribution algorithm

Ningning Zhu[1], Fuqing Zhao[1], Ling Wang[2], Chenxin Dong[3]

[1]School of Computer and Communication Technology, Lanzhou University of Technology, Lanzhou 730050, China

307516638@qq.com(Ningning Zhu); Fzhao2000@hotmail.com(Fuqing Zhao);

[2]Department of Automation, Tsinghua University, Beijing, 10084, China

wangling@tsinghua.edu.cn

[3]School of Mechanical and Automotive Engineering, Qingdao Hengxing University of Science and Technology, Qingdao 266100, China

deyangxuanyi@hotmail.com

A Kalman filtering driven estimation of distribution algorithm is proposed.

A historical learning mechanism embedded in the probabilistic model is designed.

The enhanced information specific to the problem characteristics is introduced.

The filtering mechanism based on the elite strategy is presented.

The population adaptive adjustment strategy is employed.

An enhanced Kalman filtering and historical learning mechanism driven estimation of distribution algorithm

Ningning Zhu[1], Fuqing Zhao[1], Ling Wang[2], Chenxin Dong[3]
[1]*School of Computer and Communication Technology, Lanzhou University of Technology, Lanzhou 730050, China*
307516638@qq.com(Ningning Zhu); Fzhao2000@hotmail.com(Fuqing Zhao);
[2]*Department of Automation, Tsinghua University, Beijing, 10084, China*
wangling@tsinghua.edu.cn
[3]*School of Mechanical and Automotive Engineering, Qingdao Hengxing University of Science and Technology, Qingdao 266100, China*
deyangxuanyi@hotmail.com

Abstract: As an evolutionary algorithm based on probabilistic models, estimation of distribution algorithm (EDA) utilizes the information of promising solutions to solve complex continuous optimization problems. However, the traditional EDA is confronted with the limitations that the search direction and scope are not conducive to the evolution process, which lessens the efficiency of the search. An enhanced Kalman filtering and historical learning mechanism driven EDA (KFHLEDA) is proposed to improve the search efficiency and the modeling accuracy of traditional EDA in this study. The enhanced Kalman filtering covers the stages of the prediction, the observation, and the first and the second revision in allusion to specific problems during the search, increasing population diversity and modeling accuracy. The elite strategy is employed to modify the revision improvement matrix, which accelerates the process of evolution. The enhanced revision information is fed back to the probabilistic model through the history learning mechanism to ameliorate the modeling accuracy. The population adaptive adjustment strategy reduces the number of invalid iterations. KFHLEDA improves the search direction and expands the search scope after the theoretical analysis. Compared with several classical metaheuristic algorithms and the state-of-the-art EDA variants, the evaluation results on benchmark functions of CEC 2017 test suit indicate that KFHLEDA is effective and competitive.

Keywords: Estimation of distribution algorithm; Kalman filtering; Historical learning mechanism; Adaptive adjustment strategy; The elite strategy

1. Introduction

Complex continuous optimization problems [1] are the common fields applied to optimize various indexes in industrial production, economic forecasting, and manufacturing systems with the purpose of minimizing the objective functions and the optimal solution vectors. The computational complexity increases exponentially with the dimension of the problem scale [2]. In addition, due to the variable non-separability, rotation invariance, nonlinearity and other characteristics, Branch and bound [3], linear programming [4], as well as the traditional mathematical methods are hard to solve such problems effectively, and the mathematical models are difficult to establish accurately in many practical optimization problems.

The swarm intelligence optimization algorithms [5] based on Darwin's law of natural selection achieve the remarkable success in the continuous and combinational optimization problems. An offline learning co-evolutionary algorithm (OLCA) based on machine learning was proposed by Zhao et al. [6], and achieves good effects in CEC 2017 test suit. A population-based iterated greedy algorithm (PBIGA) is presented to address a distributed assembly no-wait flow-shop scheduling problem to minimize the total flowtime [7]. Meta-heuristic [8], a typical population-based intelligent optimization algorithm, overcomes the defects of the traditional mathematical methods due to its robustness and flexibility. The exploration and exploitation capability are very important indexes for meta-heuristic algorithms [9]. The aim of exploration is to find the potential area where the optimal solution is located, while exploitation aims to focus on the precise location. For optimization problems, a better balance between exploration and exploitation is essential to achieve the optimal optimization effect.

EDA [10] is a typical meta-heuristic method based on statistical probability model estimation and sampling, different from other meta-heuristic methods generating new solutions through crossover and mutation operations, such as differential evolution (DE) [11], the invasive weed optimization (IWO) [12] and brain storm optimization algorithm (BSO) [13]. EDA combines genetic algorithm (GA) [14] and statistical learning methods to select the dominant individuals in the population for modeling and estimating the distribution of solutions for the next generation, so it guides the population close to the optimal potential area. Gaussian EDA (GEDA) [15] is one of the classical models, which makes full use of the probability distribution information of effective dominant solutions to converge to obtain the optimal solution. Effective information is extracted and utilized in the evolutionary process to obtain the relationship between variables stealthily.

The traditional GEDA captures the characteristics of the optimization process and utilizes the dominant information to guide the population closer to potential areas, it still suffers from some limitations, attracting the attention of researchers. First, the population diversity is easy to lose, thus falling into the local optimum. Liang et al. preserved a certain number of high-quality solutions generated in the previous generations into an archive and employed them to assist in estimating the covariance matrix of the Gaussian model to promote population diversity [16]. Then, the direction of population evolution is perpendicular to that of fitness improvement, reducing the search efficiency. Ren et al. proposed an anisotropic adaptive variance scaling (AAVS) technique adjusting the variance of different characteristic directions to improve the search efficiency [17]. Moreover, the quality of inferior candidate solutions from the previous generation reduces the modeling accuracy. Yang et al. proposed a multimodal EDA, equipped with a dynamic cluster sizing strategy, an alternative utilization of mixed distributions to generate offspring, and an adaptive local search to improve the modeling accuracy for multimodal problems [18].

Can a method be proposed to solve the problems mentioned above simultaneously? In addition, how to analyze the improvement of the search range and direction of EDA by the proposed method theoretically? This is exactly what this study does. Kalman filtering, a digital processing technology in a communication system, is introduced. Based on the specific problem characteristics, it is combined with other learning mechanisms and corresponding adjustment strategies. A novel EDA variant, named KFHLEDA, is developed to solve the limits of loss of population diversity, small search scope and poor search direction of traditional EDA.

An enhanced Kalman filtering and historical learning mechanism driven EDA is proposed to solve the problems mentioned above to improve the search efficiency and modeling accuracy. The effectiveness is verified through theoretical analysis and experimental results. The main contributions are summarized as follows.

- A novel EDA model is introduced based on an enhanced Kalman filtering mechanism designed for specific problem characteristics. In order to prevent population stagnation, the filtering process is divided into four stages further: prediction, observation, and the first and second revision, which modify the new search center effectively and improves the search range and direction of EDA.

- The elite strategy is joined into the revision improvement matrix to accelerate the process of evolution. The population adaptive adjustment strategy is employed to reduce the invalid search and achieves a good balance between the exploration and exploitation better.
- The enhanced information acquired by Kalman filtering operation is fed back to EDA model through a historical learning mechanism to increase the population diversity and improve the quality of solutions further.

The remainder of this study is organized as follows. A literature review is presented in Section 2. The proposed KFHLEDA is demonstrated in Section 3. The experimental results of KFHLEDA and the comparison algorithms are described in Section 4. The conclusions are drawn in Section 5.

2. Literature review

EDA was first proposed by H.Mühlenbein and G.Paaβ [10] in 1996. EDA builds a probability model based on the dominant solutions generated by the parents, which has a high probability of obtaining the optimal solution. Different probabilistic models lead to different ways of generating solutions. Common models include histogram models [19], Cauchy model [20] and Gaussian model. Among them, GEDA is constructed by Gaussian model based on the Gaussian probability distribution, which has been studied by many scholars in recent years. GEDA is divided into three types. The first is the univariate model represented by the univariate marginal distribution algorithms (UMDA) [21], not considering the dependencies between variables. The second is represented by the Bayesian network [22][23], considering the specific dependence relationship to some extent. The third is represented by the estimation of multivariate normal density algorithms (EMNA) [24], which involves the relationship between multiple variables fully.

Though GEDA acquires good results in many continuous problems, there are some limits, such as fast diversity loss, poor evolution direction, and small search scope. All of them affect the generation of solutions. The quality of the solution generated by GEDA depends on the probabilistic model and the way of updating the model. Probabilistic models are constructed based on the mean and covariance. When the mean or covariance are improved in different ways, many EDA variants are produced to solve different problems. Yang et al. presented an adaptive covariance scaling estimation of distribution algorithm (ACSEDA) calculating the covariance according to an enlarged number of promising solutions, and devised an adaptive promising individual selection strategy for the estimation of the mean vector and an adaptive covariance scaling strategy for the covariance estimation further [25], but the framework is complicated. In addition to adjusting covariance directly, regulating the eigenvalues of the covariance matrix is also an effective way. For example, Wagner et al. adjusted variance by replacing the maximum eigenvalue with the minimum one [26]. Moreover, some works of literature adopt the methods of changing the estimated covariance matrix. The covariance matrix adaptation evolution strategy (CMA-ES) [27] uses the rank-μ-update operation to increase the variance along the gradient direction, which improves the variance of the evolutionary direction effectively, but the complex framework leads to poor robustness and limited application. Liang et al. improved the performance of GEDA by exploiting the inferior solutions after a simple repair operation to adjust the covariance matrix, then, a better search direction and a more proper search scale are achieved [28], but the historical information is not made full use of.

GEDA utilizes the dominant information in the solutions generated by the current generation to estimate probabilistic models. The dominant solutions have a higher probability to be close to the global optimum [29], which accelerates the evolution speed, therefore, the dominant information is crucial. If the probabilistic model is updated only by the dominant solutions of the current generation, the excellent information obtained by the previous generations is lost. The historical learning mechanism takes advantage of the effective historical information to influence the subsequent evolutionary process. The Gaussian model is estimated to predict new solutions utilizing historical dominant data rather than limited information from the current generation. The historical information hidden in the archives is integrated into the EDA framework, increasing population diversity and improving the quality of solutions. The historical archive was first employed by Zhang et al. [30], and a new differential evolution (DE) algorithm, JADE, is proposed to improve optimization performance by implementing a new mutation strategy with an optional external archive utilizing historical data to provide information for evolution direction. Liang et al. preserved an archive and employed the solutions in the archive to assist in estimating the covariance matrix of the probability model to promote population diversity [16]. Inspired by this, the historical learning mechanism is introduced into the proposed KFHLEDA.

The elite strategies guide the population to approach the direction of the optimal solution quickly by adopting the elitist solutions [31], so the search efficiency is improved. The elite strategies have been employed in meta-heuristics algorithms. Yang et al. proposed a random neighbor elite guided differential evolution algorithm (RNEGDE), and a novel random neighbor elite guided mutation strategy was employed to mutate individuals to promising areas fast without serious loss of diversity [32]. A multi-population cooperative evolution mechanism guided the backtracking search optimization algorithm with hierarchical knowledge (HKBSA) and the elite strategy were proposed by Zhao et al. to improve the performance of BSA [33]. Li et al. suggested an opposition-based butterfly optimization algorithm with the adaptive elite mutation (OBOAEM) to solve the complex high-dimensional optimization problems [34]. This study draws on its good application in other meta-heuristic algorithms and applies it to the probabilistic model of GEDA.

The selection of population size is an important metric for many meta-heuristics algorithms [35], especially for EDA [17]. A large population size is beneficial to exploration, but the convergence speed is slow and computing resources are occupied. A small population size accelerates the convergence speed, but it is easy to fall into local optimum. Different population adjustment strategies are applied in meta-heuristic algorithms, and good results are obtained. Auger et al. introduced a restart-CMA-evolution strategy by increasing the population size, and the search characteristic became more global after each restart [36]. Tanabe et al. addressed L-SHADE with linear population size reduction (LPSR), continually decreasing the population size in accordance with a linear function [37].

Except for implementing different probabilistic model estimation methods and introducing learning mechanism and corresponding strategy to improve the performance of EDA, some other mechanisms are integrated into EDA by many works of literature. For example, Shi et al. proposed a multimodal EDA, equipped with a dynamic cluster sizing strategy and an alternative utilization of Gaussian and Cauchy distributions for multimodal problems [38]. Wang et al. proposed an adaptive estimation distribution distributed differential evolution (AEDDDE), where every individual formed its own niche to find the global optimum, and different parameter-free niches were coevolved with the master-slave mechanism to solve distributed multi-modal problems [39]. A new estimation of distribution algorithm based on normalized mutual information (NMIEDA) [40] provides a new updating mechanism, combined with stochastic sampling and the opposition-based learning scheme, accelerating the convergence speed of EDA. A novel multiple sub-models maintenance technique, named maintaining and processing sub-models (MAPS), aims to enhance the ability of estimation of distribution algorithms (EDAs) by the explicit detection of the promising areas on multimodal problems [41]. Li et al. proposed a novel estimation of distribution algorithm (EDA) with a hybrid-model EDA, adding adaptive hybrid-model learning (AHL) strategy, an orthogonal initialization (OI) strategy, and a surrogate-assisted multi-level evaluation (SME) method for the efficient hyperparameters optimization to deal with the challenge of large-scale search space efficiently [42].

Combining EDA with other meta-heuristic algorithms produces different EDA variations. These operations give full play to their own advantages. For example, EDA and DE [43], EDA and particle swarm optimization algorithms (PSO) [44] are combined to solve the continuous optimization problem. Cuevas et al. integrated the exploration ability of the invasive weed optimization (IWO) into the EDA probability model, and designed an adaptive mixed Gaussian-Cauchy distribution model [20]. Some frameworks are similar and simple, but only emphasize model structure.

EDA is equipped with a good self-learning ability, and acquires good results in continuous optimization problems. Meanwhile, EDA has also been used to solve a variety of practical problems. A Pareto-based estimation of distribution algorithm (PEDA) was presented by Shao et al. for solving the multi-objective distributed no-wait flow shop scheduling problem with sequence-dependent setup time [45]. A Pareto multi-objective optimization model based on the estimation of distribution algorithm (MOEDA) for the energy-efficient distributed blocking flow shops (EDBFSP) was proposed by Zhang et al. [46]. An innovative three-dimensional matrix-cube-based estimation of distribution algorithm (MCEDA) was proposed by Zhang et al. for the distributed assembly permutation flow-shop scheduling problem to minimize maximum completion time [47]. A hybrid multi-objective optimization algorithm of estimation of distribution algorithm and deep Q-network is proposed by Du to solve a flexible job shop scheduling problem with time-of-use electricity price constraints to optimize both maximum completion time and total electricity price [48].

To sum up, various works of literature solve the limitations of EDA in one or two different aspects by employing the corresponding learning mechanisms and adjustment strategies, or combining with other meta-heuristics. This paper adopts a different Kalman filtering and historical learning mechanism driven EDA to improve the population diversity, enlarge the search range, and adjust the search direction of EDA in the continuous optimization problems, simultaneously.

3. KFHLEDA

3.1 Kalman filtering

Kalman filtering [49] is an algorithm that utilizes the state equation to estimate the state of the system optimally through input and output observations. The state of the system, a set of minimal parameters containing all past inputs and disturbances, determines the entire behavior of the system together with future inputs and disturbances. Once the statistical properties of the disturbance and observation error are assumed properly, the estimated value of the real state is acquired by processing the observed variables.

The probability distribution adopted in this study is based on Gaussian type. According to the properties of Kalman filtering, the mean and covariance of Gaussian random variables are calculated by Kalman recursive formula, when the observation data and the state obey Gaussian distribution jointly. The updating process of conditional probability density is conducted by the minimum variance estimation. The state estimation is an important part of Kalman filtering. In general, quantitative inference of variables based on the observed data is regarded as an estimation problem. Especially, the state estimation of dynamic behavior is able to predict the current state.

It is very ideal to apply Kalman filtering in the continuously changing dynamic systems [50]. The reason is that it always points to the real state to make a reasonable prediction of the next direction for the system. The predicted value and observed one are calculated according to Eqs. (1)-(2), respectively.

$$\boldsymbol{X}(k) = \boldsymbol{A} * \boldsymbol{X}(k-1) + H * \boldsymbol{\theta}(k) + \boldsymbol{\varphi}(k) \tag{1}$$

$$\boldsymbol{Y}(k) = \boldsymbol{B} * \boldsymbol{X}(k-1) + \boldsymbol{\varepsilon}(k) \tag{2}$$

where $\boldsymbol{X}(k-1)$ represents the predicted value at the previous time, $\boldsymbol{X}(k)$ represents the predicted value at the current time, $\boldsymbol{Y}(k)$ represents the observed value at the current time, $\boldsymbol{\varepsilon}(k)$ represents the noise component during observation, $\boldsymbol{\varphi}(k)$ represents the prediction error matrix, $\boldsymbol{A}$ represents the state transition matrix from the previous state to the current state, $\boldsymbol{B}$ represents the coefficient matrix of input variables, and $\boldsymbol{H}$ represents the observation coefficient matrix.

In this study, Kalman filtering with the functions of the prediction, observation, and revision is introduced into GEDA. The process of Kalman filtering is shown in Algorithm 1.

Algorithm 1 Kalman filtering

Set the initial state variable and revision improvement matrix.
$nfes = 1$.
While $nfes \leq max\_nfes$ **do**
Perform prediction function:
    Obtain the predicted status value $\boldsymbol{X^{pre}}$ and revision improvement matrix $\boldsymbol{\varepsilon}$.
Perform observation function:
    Obtain the observed status value $\boldsymbol{X^{obs}}$.
Perform revision function:
    Calculate the gain vector $\boldsymbol{g}$.
    Update the status value $\boldsymbol{X^{rev}}$.
    Update revision improvement matrix $\boldsymbol{\varepsilon}$.
$nfes = nfes + 1$.
End while

3.2 GEDA with historical learning mechanism

The dominant solutions from Kalman filtering are stored into the history archive, and more hidden information in the archive is learnt. The information from the historical archives is employed to estimate the covariance matrix to revise the modeling data, which improves the model accuracy further. The dominant solutions acquired by different stages of Kalman filtering are stored in an archive. It is rich in edge information, which is beneficial for the evolution.

To emphasize this idea, on the one hand, the weighted fitness function of the dominant individuals in the archive is employed to modify the mean value to enhance the influence of the dominant information; on the other hand, the historical archive is used to update the mean and covariance matrix, as shown in Eqs. (3)-(4).

$$\mu(k) = \frac{1}{|\boldsymbol{X}|} * \sum_{i=1}^{|\boldsymbol{X}|} x_i(k) \tag{3}$$

$$C(k)=\frac{1}{|\boldsymbol{X}|}*\sum_{i=1}^{|\boldsymbol{X}|}(x_i(k)-\mu(k))\,(x_i(k)-\mu(k))^T \tag{4}$$

where $x_i(k)$ represents the ith solution in the history archive, $\boldsymbol{X}=\{x_1(k),x_2(k),\cdots,x_{|\boldsymbol{X}|}(k)\}$, and $|\boldsymbol{X}|$ represents the number of solutions in the archive.

In this study, matrices or sets are represented in bold. The operation process of GEDA is shown in Algorithm 2, where $\boldsymbol{POP}$ represents the whole population, $nfes$ represents the number of iterations, max_nfes represents the maximum number of iterations, τ is the selection rate, and NP is the population size.

Algorithm 2 GEDA

Input: τ, NP

Output: $\boldsymbol{NEW\_POP}$

```
Initialize all parameter values.
Generate the initialization population POP(int).
While nfes≤max_nfes do
    Evaluate the fitness function.
    Select the dominant population POP(sup) from the whole POP by truncation selection with the selection rate
    Estimate the mean μ and covariance matrix C.
    Sample the Gaussian probability model to generate the new dominant population placed into the archive.
    Merge dominant populations in the archive to form a new population NEW_POP.
    nfes = nfes + 1.
End while
```

3.3 Population adaptive adjustment strategy

EDA is a population-based algorithm, and the size of the population has a significant impact on its performance. In the early stage of evolution, a larger search space avoids getting stuck in a local optimum, hence, the exploration ability is vital. In the later stage, the exploitation ability is emphasized. A large population size is good for the population diversity and the exploration ability, while a small one contributes to the exploitation. The decreasing of the population size accelerates the convergence speed and achieves the optimal solution appropriately. Different from the conventional practice of a fixed population size, a population adaptive adjustment strategy makes a better balance between the exploration and exploitation better. It takes advantage of the specific characteristics of the population in different stages to adjust the population size with the iteration process adaptively, which avoids the waste of computing resources. The adjustment rules are as shown in Eq. (5).

$$NP(k+1)=round(NP_{max}-\frac{nfes}{max\_nfes}*(NP_{max}-NP_{min})) \tag{5}$$

where NP_{max}, and NP_{min} represent the maximum and the minimum number of the population, respectively; $nfes$, and max_nfes represent the fitness evaluation times of the current generation and the total maximum evaluation times, respectively; $round$ is a function.

The population size, which is maximized initially, decreases linearly until reaching the minimum during the iteration process. The maximum and minimum of the population size are vital. An oversize value leads to slow convergence, wasting computing resources, while a too small value weakens the exploration ability. Combined with the characteristics of GEDA, the covariance matrix contains $0.5*(n^2+n)$ estimated parameters, here, NP_{min} is set as $0.5*(n^2+n)$. NP_{max} is obtained by parameter analysis in section 4.2.

3.4 The proposed KFHLEDA

Kalman filtering is employed in various communication systems. In this study, the idea of system is combined with intelligent optimization algorithm and specific problem characteristics. It is a new attempt to transfer it into the optimization problem.

The pseudocode of KFHLEDA is shown in Algorithm 3.

Algorithm 3 KFHLEDA

Input: $\tau, NP_{max}, max\_nfes, h$

Output: The best individual

Initialize the population randomly.
$nfes = 1$;
Select the most excellent $\tau * NP$ individuals as the modeling population $\boldsymbol{POP^{mod}(1)}$.
Find the optimal individual.
Establish the model according to Eqs. (3)-(4).
Sample and obtain $(1-\tau) * NP_{max}$ observed individuals $\boldsymbol{POP^{obs}(1)}$.
Merge $\boldsymbol{POP^{mod}(1)}$ and $\boldsymbol{POP^{obs}(1)}$ to form a new population $\boldsymbol{NEW\_POP_{(1)}}$.
Calculate the revised improvement matrix $\boldsymbol{\pi(1)}$ according to Eq. (6).
$nfes = 2$;
While $nfes \leq max\_nfes$ **do**
Perform the prediction operation of Kalman filtering according to Eq. (7) to obtain the predicted population $\boldsymbol{POP^{pre}}$.
Find the best individual x_{pbest} in the archive.
If the conditions for the first revision are satisfied **do**
Perform the first revision.
End
Perform the second revision.
Select the dominant individuals of $\boldsymbol{POP^{pre}}$, and $\boldsymbol{POP^{rev}}(nfes-1)$ in the archive h as $\boldsymbol{POP^{mod}}$.
Establish the model according to Eqs. (11)-(12).
Calculate the population size of every generation according to Eq. (5).
Sample and obtain the observed population $\boldsymbol{POP^{obs}}$.
Merge the selected dominant individuals of $\boldsymbol{POP^{mod}}$,and $\boldsymbol{POP^{obs}}$.
Generate a new population $\boldsymbol{NEW\_POP}$.
Calculate the revision gain coefficient $\boldsymbol{g}$ according to Eq. (9).
Obtain a revised population $\boldsymbol{POP^{rev}}$ according to Eq. (10).
Calculate the revision improvement matrix $\boldsymbol{\pi}$ according to Eq. (8).
$nfes = nfes + 1$.
End while

where τ represents the truncated selection rate, that is, the proportion of selected dominant individuals, NP_{max} represents the population size of initialization, max_nfes represents the maximum number of iterations, h represents the size of the history archive, $\boldsymbol{POP^{pre}}$, $\boldsymbol{POP^{mod}}$, $\boldsymbol{POP^{obs}}$, *and* $\boldsymbol{POP^{rev}}$ represent the predicted , the modeling, the observed and the revised population, respectively. $\boldsymbol{NEW\_POP}$ represents a new population after merging, $\boldsymbol{\pi}$ represents the revision improvement matrix, and $\boldsymbol{g}$ represents the revision improvement matrix.

The probabilistic model of traditional GEDA does not make full use of the information from historical individuals. The dominant solutions from the prediction, observation, and revision stages of Kalman filtering are placed into the historical archive, and more potential regions in the search space are discovered fast. The enhanced information hidden in the archive is fed back to the probabilistic model through the history learning mechanism. The continuous optimization problem is simulated as a system. Correspondingly, the process of solving the problem is equivalent to that of reaching the optimal state of the system by revising the observed variables gradually. Inspired by the effective effect of the elite strategy on solving the optimization problems [33], the elite strategy is integrated into the revision improvement matrix to accelerate the search speed during the whole process of enhanced Kalman filtering.

Kalman filtering utilizes the state information at the previous generation to predict the current state and acquires the observed value. After revising the predicted and the observed value, the revised value is obtained, then it affects the solution of the next generation. Appropriate refinement operations lead the solutions more likely to move to the search center. The aim of filtering is to conduct the direction of evolution and improve the modeling accuracy.

3.4.1 Initialization

Based on the idea of Kalman filtering, the prior information is required. The updating of the first generation needs to be generated separately.

The dominant population $\boldsymbol{POP^{mod}}$ is obtained by the truncation selection method, and the optimal individual x_{pbest} is found. After modeling and sampling, $\boldsymbol{POP^{obs}}$ are obtained, ultimately, a new population $\boldsymbol{NEW\_POP}$ at the first generation is constructed with $\boldsymbol{NEW\_POP_{(1)}} = (\boldsymbol{POP^{mod}}, \boldsymbol{POP^{obs}})$. To improve the model accuracy further, the revision improvement matrix at the first generation is calculated according to Eq. (6), and the prediction accuracy in each generation is added to the revision process of the filtering stage.

$$\boldsymbol{\pi(1)} = \begin{bmatrix} \pi_1(1) \\ \vdots \\ \pi_m(1) \end{bmatrix} = \begin{bmatrix} x_{pbest}(1) - x_1^{obs}(1) \\ \vdots \\ x_{pbest}(1) - x_m^{obs}(1) \end{bmatrix} \tag{6}$$

where $\boldsymbol{x_i^{obs}(1)} = \left(x_1^{obs}(1), \cdots, x_m^{obs}(1)\right), i = 1, \cdots m$ represents the selected dominant individuals of the observed population at the first generation.

3.4.2 The prediction and observation stages

From the second iteration onwards, the dominant individuals in the archive are used to predict the optimal solution. The predicted population, denoted as $\boldsymbol{POP^{pre}}(k)$, is generated by predicting the relevant information of the new population of the last generation, and the optimal individual x_{pbest} is achieved from the historical archive.

The prediction operation is performed according to Eq. (7).

$$x_i^{pre}(k) = C * x_{pbest}(k-1) + D * \pi_i(k-1) \tag{7}$$

where $x_i^{pre}(k)$ is the predicted individual of the kth generation, C represents an identity matrix, $x_{pbest}(k-1)$ is the best individual of previous generation, D is a diagonal matrix with the diagonal elements generated randomly, and $\pi_i(k-1)$ is the revision improvement of the ith individual in $(k-1)th$ generation. The revision improvement matrix $\boldsymbol{\pi}(k)$ is calculated by Eq. (8).

$$\boldsymbol{\pi}(k) = \begin{bmatrix} \pi_1(k) \\ \vdots \\ \pi_m(k) \end{bmatrix} = \begin{bmatrix} x_{pbest}(k) - x_1^{obs}(k) \\ \vdots \\ x_{pbest}(k) - x_m^{obs}(k) \end{bmatrix} \tag{8}$$

The number m of the selected dominant individuals is calculated by $m = \tau * NP(k)$, the value of τ is obtained based on the parameter analysis in section 4.2, and $NP(k)$ is the population size of the kth generation, which is calculated by the population adaptive adjustment strategy in Section 3.3. The dominant individuals both in $\boldsymbol{POP^{pre}}(k)$ and $\boldsymbol{POP^{rev}}(k-1)$ are used for modeling, and the observed solutions for observation are generated.

3.4.3 The revision procedure

Different from the conventional Kalman filtering process, the revision procedure is divided into two stages, namely, the first revision and the second revision, aiming at the specific problem of the optimization information.

(1) the first revision

During the evolution, if the condition of the first revision is satisfied, the first revision operation is performed.

Condition 1: The covariance matrix is not positive definite.

Condition 2: The difference between the maximum and minimum values of individuals in each variable is less than 10^{-4}.

Once at least one of the two conditions is met, the percentage of each type of population is constantly revised. The specific revision method is described as follows. With the selection rate τ, the dominant individuals in the archive are selected for modeling, which affects the evolution of the next generation. When at least one of these two conditions occurs, the selection rate of the predicted population is increased by 10%, and that of the other population is decreased by 10%. Similarly, if no better solution exists for two consecutive generations, the selection rate is changed again in the same way. Once these cases occur, the percentage of the predicted population is increased by 10% each time, and the other is decreased by 10%. If the percentage exceeds 100%, 100% is subtracted and the remaining percentage is kept; if it's less than 0, 100% is added as well. For example, when the proportion of the predicted population comes up to 95%, the other is 5% exactly. When the condition of the first revision operation is satisfied, the selection rate of the predicted population after revision is calculated as (95%+10%)-100%=5%; for the other population, the selection rate is (5% -10%) + 100% = 95%.

(2) The second revision

Dominant individuals of $\boldsymbol{POP}^{\boldsymbol{pre}}(k)$, and $\boldsymbol{POP}^{\boldsymbol{rev}}(k-1)$ are selected from the historical archive. After the operation of modeling and sampling is implemented, $\boldsymbol{POP}^{obs}(k)$ is acquired. $\boldsymbol{POP}^{mod}(k)$, and $\boldsymbol{POP}^{obs}(k)$ are combined to generate a new population $\boldsymbol{NEW\_POP}$, and the optimal individual is achieved.

The search process of EDA is not linear in general, and the transformation is necessary. It is carried out by revising the gain coefficient and the improvement matrix of the system. The gain coefficient $g_i(k)$ is shown in Eq. (9).

$$g_i(k)=\frac{\sum_{i=1}^{m}\pi_i(k-1)}{\sum_{i=1}^{m}\pi_i\,(k-1)*\sum_{i=1}^{m}{}^{\left(f\left(x_{pbest}(k-1)\right)-f(x_i^{obs}(k-1))\right)}\!\Big/_{m}} \tag{9}$$

where $f\left(x_{pbest}(k-1)\right)$, $f(x_i^{obs}(k-1)))$ are the fitness values of the optimal individual and the ith individual at the $(k-1)th$ generation, respectively.

The gain coefficient introduces the information of the revision improvement matrix. The fitness difference between the best and the observed individual at the previous generation as the weighted coefficient in the revision population, which is generated according to Eq. (10).

$$x_i^{res}(k)=x_i^{pre}(k)+g_i(k)*\left(f\left(x_{pbest}^{pre}(k)\right)-f\left(x_i^{pre}(k)\right)\right)*\pi_i(k),\qquad i=1,\cdots,m \tag{10}$$

The revision operations are executed in the direction of the fitness improvement. The revision solutions are also recycled in the next generation of modeling, which improves both the diversity of the population and the accuracy of modeling. The composition of the solutions after Kalman filtering is shown in Fig.1. The red dot represents the optimal solution, the black dots are the current solutions, the blue triangles are the predicted solutions, and the green squares are the revision solutions.

The elite solution accelerates the iteration speed of the population and reduces the consumption caused by randomness. [32][33] have verified that the elite strategy guiding the direction of the population evolution is a very effective measure. The proposed algorithm utilizes the specific knowledge information carried by the dominant and elite individuals, which optimizes the distribution of the individuals.

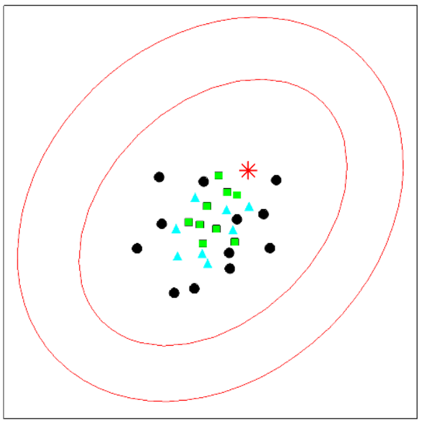


Fig.1 The solution composition after Kalman filtering operation

According to the definition of mean in Eq. (3), the improved mean value after revision is expressed as follows.

$$\tilde{\mu}(k) = \frac{1}{|X|} * \sum_{i=1}^{|X|} (x_i(k) + g * (f\left(x_{pbest}(k) - f\left(x_i(k)\right)\right) * \pi_i(k)) \tag{11}$$

The improved mean explores more potential solution spaces. On this basis, the improved covariance matrix is as shown in Eq. (12).

$$\tilde{C}(k) = \frac{1}{|X|} * \sum_{i=1}^{|X|} (x_i(k) - \tilde{\mu}(k))\,(x_i(k) - \tilde{\mu}(k))^T \tag{12}$$

The revision improvement of the mean is shown as follows.

$$\Delta(k) = \tilde{\mu}(k) - \mu(k) \tag{13}$$

After the operations of enhanced Kalman filtering, the improved mean moves to a more potential direction close to the search center. Better solutions are achieved quickly along the direction of the fitness improvement. In addition, the revision stage is guided by the optimal solution, which speeds up the search efficiency of EDA and reduces the time consumption caused by the random search.

The solution distribution and evolution direction of original GEDA are shown in Fig. 2. KFHLEDA expands the search space and improves the search direction, as shown in Fig. 3. The black dots represent the current solutions, the red dot represents the optimal solution, the green squares are the revision solutions, FID represents the fitness improvement direction, and ED represents the evolution direction, which is the direction of the principal axis of the probability density ellipse. ED is perpendicular to FID in Fig. 2. After the filtering, ED is parallel to FID.

In the process of Kalman filtering, firstly, the prediction, observation, and revision operations guided by the elite strategy are embedded into the EDA framework, increasing the search range and adjusting the search direction effectively. The detailed theoretical analysis is shown in Section 3.5. Secondly, a historical archive is added to GEDA with Kalman filtering to make full use of the dominant information of various types of individuals. Thirdly, in allusion to the specific problem of the search process, the revision operation of Kalman filtering is divided into two stages, namely, the first and the second revision, which are conducive to solving the problem of evolutionary stagnation during the search process.

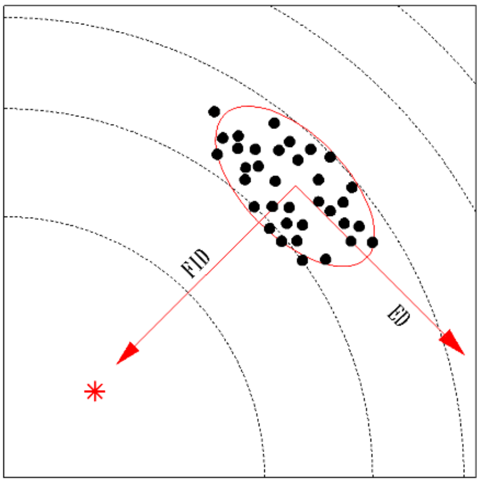


Fig.2 The solutions distribution and evolution direction of original GEDA

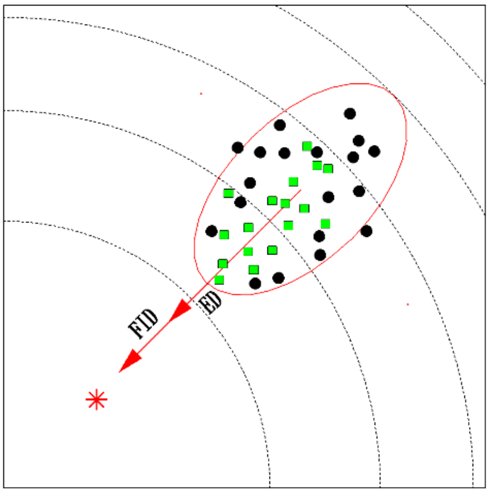


Fig.3 The solutions distribution and evolution direction after revision

3.5 Theoretical analysis

The covariance matrix determines the search range and direction [25]. As can be seen from Fig. 2, the individuals generated by the traditional GEDA evolve in the worst direction, that is, perpendicular to the direction of fitness improvement. After the enhanced Kalman filtering operation, the more potential revision solutions are achieved. According to the improved mean and covariance matrix, the principal axe of the probability density ellipsoid of the revised covariance matrix tilts in the direction of $\Delta(k) = \tilde{\mu}(k) - \mu(k)$, which is the direction of fitness improvement.

$$\begin{aligned}\tilde{C}(k) &= \frac{1}{|X|} * \sum_{i=1}^{|X|} \left(x_i(k) - \tilde{\mu}(k)\right)\left(x_i(k) - \tilde{\mu}(k)\right)^T \\ &= \frac{1}{|X|} * \sum_{i=1}^{|X|} \left(x_i(k) - \mu(k) - \Delta(k)\right)\left(x_i(k) - \mu(k) - \Delta(k)\right)^T \\ &= \frac{1}{|X|} * \sum_{i=1}^{|X|} \left(x_i(k) - \mu(k)\right)\left(x_i(k) - \mu(k)\right)^T - \frac{1}{|X|} * \sum_{i=1}^{|X|} \left(x_i(k) - \mu(k)\right)\left(\Delta(k)\right)^T \\ &- \frac{1}{|X|} * \sum_{i=1}^{|X|} \Delta(k)\left(x_i(k) - \mu(k)\right)^T + \Delta(k) * \left(\Delta(k)\right)^T = C(k) + \Delta(k) * \left(\Delta(k)\right)^T\end{aligned} \tag{14}$$

From Eq. (14), it is proved that the covariance matrix $\tilde{C}(k)$ improved by enhanced Kalman filtering operation is the rank-one correction of the original covariance matrix $C(k)$.

Through the analysis of the following two mathematical theories, the estimation method of the enhanced Kalman filtering extends the search range of EDA and adjusts the search in the appropriate direction adaptively.

Lemma 1: This operation is built on a set of excellent samples. For a given set, if $\tilde{\mu}(k) \neq \mu(k)$, the coverage of the probability density ellipsoid of the covariance matrix corrected by Kalman filtering is greater than or equal to that without Kalman filtering.

Proof : In the Gaussian distribution, the major axis of the probability density ellipsoid is the eigendirection of the covariance matrix, and the length of the half-axis is equal to the square root of the eigenvalues [51]. The longer the length of the major axis is, the larger the coverage area of the probability density ellipsoid becomes. The eigenvalues of the improved and original estimate of the covariance matrix are defined as $\tilde{\lambda}$ and λ, respectively, where $\lambda = (\lambda_1, \lambda_2, \cdots, \lambda_i), i = 1,2, \cdots, m$. From the definition of the covariance matrix, it is easy to prove that the covariance matrix is positive semi-definite. Therefore, the covariance matrix $C(k)$ is decomposed into $C(k) = \psi(k) * \Lambda(\mathrm{k}) * \psi(k)^{\mathrm{T}}$, where $\Lambda(\mathrm{k})$ is made up of the eigenvalues, and $\varphi(k)$ is an orthogonal matrix. By Eq. (13), $\Delta(k) = \tilde{\mu}(k) - \mu(k)$. When $\varphi(k) = \psi(k)^T \Delta(k)$, the following conclusion holds.

$$\begin{aligned} \psi(k)^T * \tilde{C}(k) * \psi(k) &= \psi(k)^T * (C(k) + \Delta(k) * \Delta(k)^T) * \psi(k) \\ &= \psi(k)^T * C(k) * \psi(k) + \left(\psi(k)^T * \Delta(k)\right) * \left(\psi(k)^T * \Delta(k)\right)^T \\ &= \Lambda(k) + \varphi(k) * \varphi(k)^T \end{aligned} \tag{15}$$

Based on Lemma 1.4 involved in [52], the eigenpolynomial of $\Lambda(k) + \varphi(k) * \varphi(k)^T$ is equal to that of $\tilde{C}(k)$.

$$\begin{aligned} f(\lambda) = \prod_{i=1}^{m} (\lambda - \lambda_1) \cdots (\lambda - \lambda_m) &= \det(\lambda * E - \Lambda(k) - \varphi(k) * \varphi(k)^T) \\ &= \det\left(\lambda * E - \Lambda(k)\right) * \det\left(E - \left(\lambda * E - \Lambda(k)\right)^{-1} * \varphi(k) * \varphi(k)^{\mathrm{T}}\right) \\ &= \prod_{i=1}^{m} (\lambda - \lambda_i) - \sum_{j=1}^{m} \left(\varphi_j^2 \prod_{i=1, i \neq j}^{m} (\lambda - \lambda_i)\right) \end{aligned} \tag{16}$$

E represents an identity matrix. Let $\lambda = 0$, Eq. (16) is written as:

$$\prod_{i=1}^{m} \tilde{\lambda_i} = \sum_{i=1}^{m} \lambda_i + \sum_{j=1}^{m} (\varphi_j^2 \prod_{i=1, i \neq j}^{m} \lambda_i) \geq \prod_{i=1}^{m} \lambda_i \tag{17}$$

Lemma 1 is proved.

The analysis of Lemma 1 shows that compared with the traditional GEDA, GEDA with the Kalman filtering mechanism expands the coverage of the probability density ellipsoid, hence, it expands the search range of GEDA.

Lemma 2: For a selected set of excellent sample combinations, if $\tilde{\mu}(k) \neq \mu(k)$, the orientation angle between $\Delta(k)$ and the major axis of the probability density ellipsoid revised with Kalman filtering is less than or equal to that without Kalman filtering.

Proof: Define the corresponding eigenvector of the maximum eigenvalue $\tilde{\lambda}_m$ of $\Lambda(k) + \varphi(k) * \varphi(k)^T$ as $\tilde{\phi}(k)$, and that of the maximum eigenvalue λ_m of $\Lambda(k)$ as $\phi(k)$. Distinctly, in accordance with the Eq. (15) of lemma 1, $\Lambda(k) + \varphi(k) * \varphi(k)^T$ and $\Lambda(k)$ are orthogonal transformations of the covariance matrix with revision and the one without revision, respectively.

To prove lemma 2, the conclusion of $\angle(\varphi(k), \tilde{\phi}(k)) \leq \angle(\varphi(k), \phi(k))$ needs to be drawn.

It is certain that there are acute and obtuse angles between two vectors, expressed in acute angle here, therefore, $\varphi(k)^T * \tilde{\phi}(k) \geq 0$, $\varphi(k)^T * \phi(k) \geq 0$.

According to the definition of eigenvalues, Eq. (18) is valid.

$$(\Lambda(k) + \varphi(k) * \varphi(k)^T) * \tilde{\phi}(k) = \lambda_m \tilde{\phi}(k) \tag{18}$$

After transformation, Eq. (19) is valid.

$$\tilde{\phi}(k) = (\tilde{\lambda}_m * E - \Lambda(k))^{-1} * \varphi(k) * \varphi(k)^T * \tilde{\phi}(k) \tag{19}$$

Let λ in Eq. (16) be replaced by $\tilde{\lambda}_m$, it is written as:

$$f(\tilde{\lambda}_m) = \left(1 - \sum_{i=1}^{m} \frac{\varphi_i^2}{\tilde{\lambda}_m - \lambda_i}\right) * \prod_{i=1}^{m} (\tilde{\lambda}_m - \lambda_i) = 0 \tag{20}$$

Next, the following two cases are discussed.

Case 1: Because $\angle\left(\varphi(k), \tilde{\phi}(k)\right) \leq 90^{\circ}$, when $\angle(\varphi(k), \phi(k)) = 90^{\circ}$, it is quite clear that $\angle\left(\varphi(k), \tilde{\phi}(k)\right) \leq \angle(\varphi(k), \phi(k))$. Lemma 2 is proved.

The following focuses on the second case.

Case 2: When $\angle(\varphi(k), \phi(k)) \neq 90^{\circ}$, that is, $\varphi(k)^{\mathrm{T}} * \phi(k) \neq 0$. According to the proof of Eq. (14) above, $\Lambda(k) + \varphi(k) * \varphi(k)^{T}$ is the rank-one correction of $\Lambda(k)$, and the product of the two vectors is not zero, so neither vector is equal to 0, then $\tilde{\lambda}_m > \lambda_m$.

Eq. (20) is transformed into Eq. (21).

$$\sum_{i=1}^{m} \frac{\varphi_i^2}{\tilde{\lambda}_m - \lambda_i} = 1 \tag{21}$$

Then, the following equation holds.

$$1 \geq \frac{\varphi_m^2}{\tilde{\lambda}_m - \lambda_m} > 0 \tag{22}$$

$$\begin{aligned}
\cos\angle\left(\varphi(k), \tilde{\phi}(k)\right) &= \frac{\varphi(k)^{\mathrm{T}} * \tilde{\phi}(k)}{\|\varphi(k)\| * \|\tilde{\phi}(k)\|} = \frac{\varphi(k)^{T} * \tilde{\phi}(k)}{\|\varphi(k)\| * \left\|\left(\tilde{\lambda}_m * E - \Lambda(k)\right)^{-1} * \varphi(k)\right\| * \varphi(k)^{\mathrm{T}} * \tilde{\phi}(k)} \\
&= \frac{1}{\|\varphi(k)\| * \sqrt{\sum_{i=1}^{m} \frac{\varphi_i^2}{(\tilde{\lambda}_m - \lambda_i)^2}}} \geq \frac{|\varphi_m(k)|}{\|\varphi(k)\| * \sqrt{\left(\tilde{\lambda}_m - \lambda_m\right) * \sum_{i=1}^{m} \frac{\varphi_i^2}{(\tilde{\lambda}_m - \lambda_i)^2}}} \\
&\geq \frac{|\varphi_m(k)|}{\|\varphi(k)\| * \sqrt{\left(\tilde{\lambda}_m - \lambda_m\right) * \sum_{i=1}^{m} \frac{\varphi_i^2}{\left(\tilde{\lambda}_m - \lambda_m\right) * \left(\tilde{\lambda}_m - \lambda_i\right)}}} \\
&= \frac{|\varphi_m(k)|}{\|\varphi(k)\| * \sqrt{\sum_{i=1}^{m} \frac{\varphi_i^2}{\left(\tilde{\lambda}_m - \lambda_i\right)}}} = \frac{|\varphi_m(k)|}{\|\varphi(k)\|} = \frac{\varphi(k)^{T} * \phi(k)}{\|\varphi(k)\| * \|\phi(k)\|} \\
&= cos\angle(\varphi(k), \phi(k))
\end{aligned} \tag{23}$$

$$\angle(\varphi(k), \tilde{\phi}(k)) \leq \angle(\varphi(k), \phi(k)) \tag{24}$$

The proof of lemma 2 illustrates that the direction of fitness improvement is exactly that of evolution after enhanced Kalman filtering.

4. Experimental results and analysis

4.1 Test Suites

KFHLEDA is tested in IEEE CEC 2017 test suite to prove the effectiveness of the proposed algorithm [53]. The algorithm is run on 30 benchmark test functions and the performance is verified. The 30 functions in 2017 test suite include three unimodal functions $f_1 - f_3$, seven simple multimodal functions $f_4 - f_{10}$, ten hybrid functions $f_{11} - f_{20}$, and ten composition functions $f_{21} - f_{30}$. It is difficult to find the optimal solution because of the complex structural characteristics and multiple local optima of most functions among them.

KFHLEDA and the adopted comparison algorithm are evaluated by the experiments on 10, 30, 50 and 100 dimensions respectively, denoted as 10D, 30D, 50D, and 100D. Each function is run for 51 times independently, and $\max\_nfes = 10000 * D$, which is taken as the experimental termination condition for each run. The function error value (the difference between the best solution and the known global optimal of the function in each run) is applied as evaluation criterion. When this value is less than 10^{-8}, it is set to 0. All the conditions and experimental environments are the same for all algorithms in this study. All experiments are performed by MATLAB on the

Tencent Cloud CVM with a 2.50GHz Intel(R) Xeon(R) Gold 6133 CPU, 8 GB of RAM.

4.2 Parameters analysis

The parameter setting has an important influence on the performance of the algorithm [12]. In the proposed KFHLEDA, three parameters are involved, which are the initial size of the population NP_{max}, the selection rate τ, and the size of the historical archive H.

It is necessary to choose a suitable initial population size. If the value is too enormous, the algorithm converges too slowly and wastes computing resources. On the contrary, the population diversity is reduced and the exploration ability is damaged further. The selection rate τ affects the selection of the dominant solutions and determines the quality of each generation of solutions directly. If it is too little, the influence of the dominant population on subsequent iterations is weakened, and it is easy to fall into the local optimum, affecting the modeling accuracy. If it is too colossal, the quality and efficiency of the selected solutions are reduced.

The size of the historical archive H determines how much historical information is retained, that is to say, the degree of learning the dominant solution makes the solution acquisition depend on both the current state and the helpful history information. If the value of H is too large, the search space is huge and contain the superabundant inferior information, which is disadvantageous to evolution and retards the convergence speed. Similarly, if it is too small, the dominant information in the archive is hard to be learned completely. Therefore, the choice of H has a significant impact on the search efficiency.

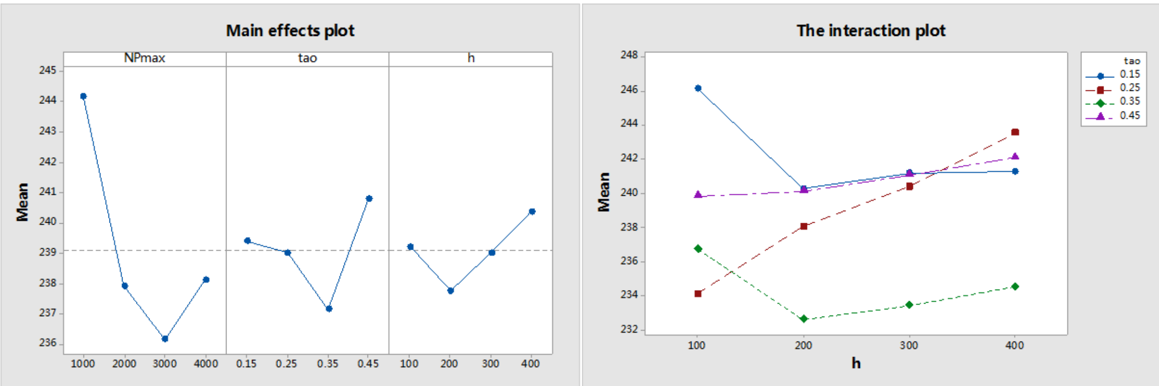


Fig. 4 Main effects plot and interaction plot of parameters

Table 1
ANOVA results for parameter settings of KFHLEDA

Source	Sum of squares	Degrees of freedom	Mean Square	F-ratio	p-value
NP_{max}	582.95	3	194.317	10.38	0.0001
τ	107.71	3	35.902	1.92	0.1506
h	54.42	3	18.14	0.97	0.4219
$NP_{max} * \tau$	339.62	9	37.736	2.01	0.0771
$NP_{max} * h$	308.31	9	34.25	1.87	0.1203
$\tau * h$	711.11	9	79.013	4.22	0.0017
$Error$	505.66	27	18.728		
$Total$	3054.18	63			

Different combinations of the three values result in the various performance. In order to achieve the best combination of these parameters to improve the performance of the algorithm, the differences are analyzed through the results in the Design of the Experiments (DOE) and Multifactor Analysis of Variance (ANOVA) [12]. The value range of the three parameters is set as $NP_{max} = \{1000, 2000, 3000, 4000\}, \tau = \{0.15, 0.25, 0.35, 0.45\}, H = \{100, 200, 300, 400\}$.

Therefore, the total number of the parameter combinations is $4*4*4=64$, and each parameter combination is run on 30 functions until the termination condition is met.

From Table 1, the p-value of NP_{max} is less than 0.05, and the corresponding F-ratio is also the largest, indicating that NP_{max} has an important influence within the 95% confidence interval. τ and h greater than 0.05 indicate an interaction between them. Refer to the main effects plot and interaction plot shown in Fig. 4. When $NP_{max}=3000$, the optimal result is achieved. When $NP_{max}=1000$, the worst is obtained. The smaller NP_{max} is, the more disadvantageous the exploration ability is, resulting in the poorer performance. A large NP_{max} wastes the computing resources. When $\tau=0.35$, the best results are achieved. $NP_{max}=3000, \tau=0.35, H=200$ are the optimal combination. The analysis above is consistent with the results shown in Fig.4.

4.3 Comparison with the state-of-the-art algorithms

In this study, some classical and the state-of-the-art algorithms are compared and tested with KFHLEDA. APBIL [54] and CMA-ES [27] are two representative different types of model-based evolution strategies. LSHADE [37], jSO [55] and PID-DE [56] are the classical representatives of differential evolution. EDA2 [16], IWOEDA [20], ACSEDA [25], AEDDDE [39] are EDA variants up to date. The population size NP_{max} and selection rate τ are set to the same values as that in this paper to ensure the experimental fairness, while other corresponding parameters in comparison algorithms are set to be consistent with those in the original literature.

The overall performance of the algorithm is evaluated through the values in Table 2-5, listing the mean and variance of KFHLEDA and comparison algorithms on 10D, 30D, 50D, and 100D for 30 functions, respectively. The numbers in bold in each line indicate that the value on this function achieves the minimum one, which means that the overall performance is the best. The smaller these values are, the better the performance is. When the mean of two algorithms is the same, the variances are compared.

For the unimodal function f_1-f_3, KFHLEDA, jSO, IWOEDA and CMA-ES find the optimal solution on 10D, 30D, 50D and 100D.

For the simple multimodal functions f_4-f_{10}, KFHLEDA shows the obvious advantages. When the dimension is 10D, LSHADE, jSO, IWOEDA, CMA-ES, EDA2 and AEDDDE are slightly better than KFHLEDA on f_4; KFHLEDA achieves the best performance on f_6-f_{10} compared with other algorithms, some of which obtain the similar results, such as LSHADE, jSO, APBIL, ACSEDA, AEDDDE on f_7; AEDDDE is the best on f_5, followed by KFHLEDA. When the dimension is 30D, CMA-ES performs best on f_4, and KFHLEDA has the excellent results on f_6, f_7, f_9, f_{10}, among which, LSHADE, jSO, APBIL, IWOEDA, EDA2 and ACSEDA also reach the optimal solution on f_6, so does f_9. AEDDDE is superior to the comparison algorithms on f_5 and f_8. When the dimension is 50D, both IWOEDA and ACSEDA overmatch other algorithms on f_6, followed by KFHLEDA and APBIL. KFHLEDA are better than others on f_7, f_8, f_9, and LSHADE, jSO, IWOEDA, EDA2, ACSEDA, KFHLEDA also achieve the similar performance. When the dimension is 100D, CMA-ES is the best on f_4. KFHLEDA is superior to other algorithms on f_5, f_8, f_{10}, respectively.

For the hybrid functions $f_{11}-f_{20}$, when the dimension is 10D, KFHLEDA achieves the excellent results compared with other algorithms on $f_{11}, f_{14}, f_{15}, f_{16}$, and some algorithms also obtain the similar good results. For example, the performance of jSO, IWOEDA, ACSEDA and AEDDDE on f_{11} is equal to that of KFHLEDA. While jSO shows obvious advantages compared with other algorithms on f_{12} and f_{13}, LSHADE, which is the best on f_{17} and f_{19}, achieves the optimal solution on f_{20}. PID-DE performs best on f_{18}. When the dimension is 30D, jSO is superior to others on f_{11}; KFHLEDA overmatches $f_{13}, f_{15}, f_{16}, f_{17}, f_{18}, f_{19}$, among which LSHADE, jSO, EDA2, ACSEDA and AEDDDE also have similar good results on f_{18}; EDA2 is the best on f_{14} and f_{20} compared with other algorithms. When the dimension is 50D, KFHLEDA is superior to others on $f_{11}, f_{13}, f_{14}, f_{16}, f_{18}, f_{19}, f_{20}$, among which ACSEDA and EDA2 achieve the equivalent results on f_{18} and f_{20}, respectively; EDA2 performs best on f_{17}. When the dimension is 100D, KFHLEDA is obviously better than other algorithms on f_{11}, and also preferable on $f_{12}, f_{14}, f_{15}, f_{16}, f_{18}, f_{19}$; AEDDDE, ACSEDA and EDA2 also achieve better performance on f_{18}, f_{19}, f_{20}.

For the composition functions $f_{21} - f_{30}$, the results of most algorithms are similar. In comparison, KFHLEDA also performs well on the whole. When the dimension is 10D, KFHLEDA shows the best performance on f_{23}, f_{25}; PID-DE is superior to other algorithms on f_{22}, f_{28}; LSHADE is the best on f_{21}; Both APBIL and IWOEDA perform best on f_{27}. When the dimension is 30D, KFHLEDA achieves the excellent performance on f_{22}, and other algorithms except PID-DE obtain similar results. In addition, KFHLEDA also performs better than other comparison algorithms on $f_{21}, f_{23}, f_{24}, f_{26}$, and EDA2, jSO have the best results on f_{27}, f_{28}. CMA-ES is superior to others on f_{25}, f_{29}, f_{30}. When the dimension is 50D, KFHLEDA also shows advantages on $f_{22}, f_{23}, f_{24}, f_{26}, f_{29}$ compared with the comparison algorithm, and both AEDDDE and KFHLEDA achieve the best results on f_{28}. When the dimension is 100D, KFHLEDA overmatches on $f_{21}, f_{22}, f_{23}, f_{24}, f_{25}, f_{26}$.

The box plots show the stability of the algorithm. A function is selected as a representative from the four types of functions with f_3, f_8, f_{15}, f_{23} here. In Fig. 5-8, the horizontal axis represents all algorithms, and the vertical axis is the error value between the candidate solution and the optimal one. Compared with other algorithms, KFHLEDA has the best stability as a whole.

The convergence curves show the convergence degree of the algorithms. The CEC 2017 test suit specifies that a convergence curve is drawn with a specified 14 points. In Fig. 9-12, the horizontal axis is the number of function evaluations, and the vertical axis is the logarithm of the error value. The red line represents KFHLEDA with best convergence precision on most functions. As can be seen from Fig. 9-12, some algorithms converge faster, but with lower accuracy. The reason is that it always tries to obtain the optimal solution, however, some algorithms fail to find the global optimal solution due to the limitation of the algorithm performance design. From the experimental data, some algorithms have stalled prematurely in the later iteration process. The overall performance of KFHLEDA is outstanding in most cases. It performs well on most functions but poorly on others.

4.4 Friedman test

Friedman test is employed to evaluate the performance of all algorithms to obtain a statistical conclusion. Fig.13 visualizes the data in Table 6 to indicate the effects of all algorithms. In Fig.13, the horizontal axis represents all algorithms, and the vertical axis represents an average rank. The bar of KFHLEDA is highlighted.

The ranking results of different dimensions reveal KFHLEDA has the best ranking. The average rank of the competitive KFHLEDA is under the critical differences. There are significant differences between KFHLEDA and most algorithms on 10D, 30D and 50D except for LSHADE and jSO on 10D, EDA2 on 30D, AEDDDE on 50D in the 90% and 95% confidence intervals. There are significant differences between KFHLEDA and all other algorithms on 100D in the 90% and 95% confidence intervals.

Table 2

The comparison results of all algorithms (10D)

Fun	PID-DE		LSHADE		jSO		APBIL		IWOEDA		CMA-ES		EDA^2		ACSEDA		AEDDDE		KFHLEDA	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean		mean	std	mean	std	mean	std	mean	std
1	6.16E-06	7.00E-06	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	4.92E+02	1.43E+03	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00
2	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	2.31E+06	2.68E+06	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00
3	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	1.78E+02	1.57E+02	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00
4	9.93E-01	5.86E-01	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	6.93E+00	6.52E-01	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	1.17E-03	8.36E-03	**0.00E+00**	0.00E+00	1.44E-07	1.03E-06
5	2.83E+01	4.17E+00	2.38E+00	8.22E-01	2.57E+00	1.06E+00	4.58E+00	2.14E+00	4.89E+00	2.06E+00	5.88E+01	4.61E+01	1.85E+01	4.25E+00	6.05E-01	8.69E-01	**2.15E-01**	5.74E-01	3.12E-01	5.45E-01
6	4.48E-05	2.05E-05	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	3.55E-07	5.59E+01	7.46E+00	6.94E-06	1.98E-06	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00
7	4.04E+01	4.18E+00	1.31E+01	9.54E-01	1.29E+01	1.06E+00	1.36E+01	1.63E+00	1.12E+01	2.20E+00	9.09E+01	3.39E+01	3.07E+01	3.34E+00	1.13E+01	6.83E-01	1.07E+01	3.34E-01	**1.01E+01**	3.42E-01
8	2.87E+01	4.15E+00	2.45E+00	1.09E+00	2.38E+00	1.05E+00	4.65E+00	1.77E+00	2.92E+00	2.01E+00	3.05E+01	2.22E+00	1.99E+01	3.75E+00	4.29E-01	8.27E-01	5.85E-02	2.36E-01	**4.15E-02**	2.99E-01
9	**0.00E+00**	3.83E-13	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	5.43E-04	1.14E-03	**0.00E+00**	2.31E-07	8.04E+02	1.09E+02	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00
10	6.55E+02	3.09E+02	2.38E+01	3.71E+01	1.72E+02	1.24E+02	1.84E+02	1.84E+02	3.68E+00	2.86E+00	1.46E+03	3.10E+02	1.15E+01	2.51E+01	1.38E+01	3.32E+01	6.18E+00	2.34E+01	**2.90E+00**	4.70E+00
11	5.14E+00	1.31E+00	4.44E-01	6.89E-01	**0.00E+00**	5.68E-14	6.61E+01	6.16E+01	**0.00E+00**	2.87E-07	3.53E-01	4.84E-01	6.78E-01	9.98E-01	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00
12	7.38E+01	4.41E+01	4.46E+01	5.53E+01	**3.02E-01**	1.60E-01	4.72E+05	4.37E+05	1.18E+02	6.34E+01	1.88E+02	1.13E+02	5.03E+00	2.34E+01	4.31E+01	5.71E+01	4.32E+01	5.56E+01	4.14E+01	5.69E+01
13	1.05E+01	1.97E+00	3.58E+00	2.23E+00	**3.17E+00**	2.51E+00	5.96E+03	3.25E+03	6.69E+00	3.20E+00	4.36E+00	3.22E-01	7.65E+00	1.33E+00	5.45E+00	3.57E+00	4.65E+00	2.04E+00	4.77E+00	2.63E+00
14	1.18E+01	5.64E+00	2.94E-01	4.68E-01	3.61E-01	3.72E-01	6.44E+02	6.34E+02	1.49E+00	1.14E+00	1.32E+01	1.07E+01	2.32E+00	2.06E+00	1.69E+00	5.41E+00	4.12E-01	2.80E+00	**2.82E-01**	2.80E+00
15	3.32E-01	2.29E-01	4.76E-01	2.08E-01	2.69E-01	2.26E-01	1.51E+03	9.08E+02	1.48E+00	5.48E-01	4.40E-01	1.91E-01	4.65E-01	5.36E-02	5.43E-01	2.91E-01	4.67E-01	1.04E-01	**2.08E-01**	1.48E-01
16	2.08E+00	3.79E+00	3.36E-01	1.73E-01	8.85E-01	3.95E-01	2.92E+00	2.61E+00	1.79E+00	1.13E+00	3.83E+02	9.29E+01	8.21E-01	1.91E-01	8.12E-01	4.06E-01	8.25E-01	2.71E-01	**1.44E-01**	2.87E-01
17	2.04E+01	5.57E+00	**1.45E-01**	1.80E-01	8.66E+00	8.11E+00	1.66E+01	1.35E+01	2.77E+01	5.95E+00	1.26E+02	1.09E+02	2.72E+01	3.81E+00	1.76E+01	9.56E+00	2.16E+01	5.71E+00	1.24E+01	1.00E+01
18	**1.89E-01**	1.32E-01	1.94E-01	2.07E-01	2.24E-01	2.23E-01	5.31E+03	3.06E+03	2.10E+00	1.97E+00	1.66E+01	8.06E+00	7.87E-01	1.24E+00	8.92E-01	2.81E+00	8.81E-01	2.81E+00	4.77E-01	6.45E-02
19	3.95E-01	3.47E-01	**1.04E-02**	1.07E-02	1.28E-02	2.71E-02	3.53E+03	1.57E+03	2.46E+01	6.57E+00	1.85E+00	1.29E+00	2.10E-01	2.97E-01	2.50E-01	2.46E-01	1.82E-01	2.36E-01	2.03E-01	2.50E-01
20	2.43E+00	3.93E+00	**0.00E+00**	0.00E+00	1.70E+00	4.98E+00	8.78E+00	9.46E+00	2.19E+01	8.85E+00	4.37E+02	2.56E+01	1.72E+00	1.71E+00	5.84E+00	8.89E+00	4.55E+00	8.35E+00	1.54E+00	5.37E+00
21	1.92E+02	6.01E+01	**1.51E+02**	5.20E+01	1.57E+02	5.25E+01	1.93E+02	3.64E+01	1.93E+02	3.64E+01	2.03E+02	1.17E+00	2.16E+02	9.28E+00	1.87E+02	2.73E+01	1.76E+02	1.51E+02	1.55E+02	1.07E+01
22	**9.31E+01**	2.89E+01	1.00E+02	5.58E-02	1.00E+02	6.25E-02	1.00E+02	5.80E-01	1.00E+02	5.80E+01	1.00E+02	0.00E+00	1.00E+02	0.00E+00	1.00E+02	0.00E+00	1.25E+02	0.00E+00	1.00E+02	0.00E+00
23	3.18E+02	4.57E+01	3.03E+02	1.54E+00	3.02E+02	1.70E+00	3.08E+02	2.37E+00	3.08E+02	2.37E+01	3.06E+02	1.58E+02	3.00E+02	8.12E-01	3.01E+02	1.93E+00	3.04E+02	2.73E+00	**2.97E+02**	2.53E+01
24	3.35E+02	7.76E+01	3.02E+02	8.09E+01	2.56E+02	1.09E+02	3.35E+02	3.11E+00	3.35E+02	3.11E+01	**1.00E+02**	0.00E+00	3.16E+02	1.62E+01	3.28E+02	3.67E-01	3.06E+02	3.21E+00	2.28E+02	4.47E+00
25	4.06E+02	1.80E+01	4.16E+02	2.26E+01	4.08E+02	1.93E+01	4.47E+02	2.45E+00	4.47E+02	2.45E+00	4.42E+02	8.20E+00	4.35E+02	1.74E+01	4.40E+02	1.23E+01	4.15E+02	1.74E+01	**3.98E+02**	1.55E-02
26	3.00E+02	0.00E+00	3.00E+02	0.00E+00	3.00E+02	0.00E+00	3.66E+02	7.27E+01	3.03E+02	7.27E+01	**2.68E+02**	4.75E+01	3.00E+02	0.00E+00	3.00E+02	3.06E-13	3.33E+02	2.97E+02	3.00E+02	0.00E+00
27	3.89E+02	4.39E-01	3.89E+02	1.22E-01	3.89E+02	2.36E-01	**3.73E+02**	1.84E+00	**3.73E+02**	1.84E+00	4.71E+02	2.36E+01	3.90E+02	2.95E-01	4.03E+02	2.52E+00	4.07E+02	1.34E+00	3.81E+02	1.15E+00
28	**3.28E+02**	8.07E+01	3.46E+02	1.08E+02	3.32E+02	8.95E+01	4.79E+02	1.88E+01	4.79E+02	1.88E+02	4.80E+02	1.34E+01	4.54E+02	1.48E+02	5.44E+02	1.30E+02	3.85E+02	1.36E+02	3.30E+02	1.00E+02
29	2.53E+02	6.32E+00	2.34E+02	2.94E+00	2.37E+02	4.80E+00	2.49E+02	6.18E+00	2.49E+02	6.18E+00	**2.31E+02**	8.06E+00	2.43E+02	3.55E+00	2.39E+02	6.57E+00	2.40E+02	3.78E+00	2.32E+02	4.03E+00
30	4.41E+02	2.26E+02	3.25E+04	1.60E+05	5.31E+04	2.04E+05	2.02E+02	7.21E-01	1.36E+03	7.21E+01	**2.01E+02**	3.58E-01	9.65E+04	2.66E+05	2.12E+05	5.32E+05	4.18E+04	4.34E+04	2.13E+05	5.90E+05

Table 3

The comparison results of all algorithms (30D)

Fun	PID-DE		LSHADE		jSO		APBIL		IWOEDA		CMA-ES		EDA^2		ACSEDA		AEDDDE		KFHLEDA	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
1	1.07E-02	2.57E-02	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	3.19E+06	1.58E+07	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	7.05E-06	1.60E-06	**0.00E+00**	0.00E+00
2	2.76E+14	1.13E+15	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	1.59E+22	8.70E+22	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00
3	1.16E+04	4.56E+03	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	1.90E+04	5.76E+03	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00
4	8.02E+01	7.71E+00	7.86E+01	0.00E+00	7.86E+01	0.00E+00	1.32E+02	1.06E+01	9.69E+00	2.50E-01	**0.00E+00**	0.00E+00	5.86E+01	0.00E+00	1.04E+02	3.40E+01	7.95E+01	7.25E+00	6.51E+01	3.60E+00
5	1.89E+02	1.11E+01	6.61E+00	1.32E+00	1.64E+01	1.17E+01	2.58E+01	7.79E+00	3.02E+01	8.85E+00	5.55E+00	1.94E+00	6.18E+00	1.99E+00	2.24E+00	1.42E+00	**5.07E-01**	8.05E-01	1.64E+00	1.07E+00
6	4.18E-04	1.62E-04	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	2.30E+01	2.66E+01	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	5.26E-05	1.18E-05	**0.00E+00**	0.00E+00
7	2.22E+02	1.17E+01	3.73E+01	1.35E+00	7.09E+01	1.24E+01	4.53E+01	4.25E+00	3.93E+01	9.91E+00	3.58E+01	1.58E+00	1.17E+02	5.33E+01	3.36E+01	8.04E-01	9.96E+01	2.43E+01	**3.30E+01**	6.73E-01
8	1.91E+02	1.04E+01	6.64E+00	1.79E+00	1.51E+01	1.14E+01	2.35E+01	7.55E+00	4.97E+00	2.31E+00	5.04E+00	1.85E+00	6.50E+00	2.44E+00	2.13E+00	1.38E+00	**2.54E-01**	6.55E-01	1.52E+00	1.20E+00
9	**0.00E+00**	**0.00E+00**	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	1.36E-01	3.54E-01	**0.00E+00**	0.00E+00	4.20E+03	1.04E+03	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00
10	6.88E+03	4.65E+02	1.45E+03	1.92E+02	4.11E+03	3.22E+02	1.83E+03	4.76E+02	3.10E+02	1.68E+02	3.43E+03	5.79E+02	2.12E+02	9.85E+01	3.71E+02	2.73E+02	6.55E+02	4.41E+02	**1.29E+02**	1.28E+02
11	7.85E+01	2.49E+01	2.27E+01	2.69E+01	**5.17E+00**	1.06E+01	4.83E+02	2.13E+02	3.94E+01	1.18E+01	3.54E+01	3.25E+01	1.79E+01	2.65E+01	6.18E+01	1.11E+01	4.97E+01	2.25E+01	1.28E+01	2.38E+01
12	1.46E+04	8.50E+03	9.50E+02	3.63E+02	2.08E+02	1.47E+02	1.81E+06	1.07E+06	8.61E+02	2.35E+02	1.26E+03	3.41E+02	1.16E+02	9.56E+01	1.17E+02	9.82E+01	**8.82E+01**	6.72E+01	9.96E+01	8.37E+01
13	9.69E+01	1.03E+01	1.54E+01	5.82E+00	1.74E+01	4.13E+00	9.10E+03	5.07E+03	3.82E+01	1.00E+01	1.55E+01	3.95E+00	1.47E+01	3.61E+00	2.77E+01	1.65E+01	2.56E+01	5.79E+00	**9.21E+00**	7.53E+00
14	6.86E+01	5.15E+00	2.05E+01	4.85E+00	2.18E+01	1.18E+00	2.26E+05	1.32E+05	2.32E+01	1.14E+01	1.36E+02	3.36E+01	**8.71E+00**	1.01E+01	1.80E+01	6.60E+00	1.60E+01	8.38E+00	1.38E+01	9.45E+00
15	4.76E+01	6.44E+00	3.62E+00	2.18E+00	8.51E-01	6.17E-01	1.71E+03	1.53E+03	1.03E+01	2.39E+00	3.40E+01	4.51E+01	4.80E-01	3.72E-01	1.40E+00	1.08E+00	7.61E-01	4.37E-01	**3.52E-01**	2.32E-01
16	1.19E+03	2.41E+02	5.43E+01	5.14E+01	1.74E+02	9.23E+01	4.58E+02	2.15E+02	3.31E+01	2.19E+01	7.77E+01	1.02E+02	1.54E+01	2.34E+01	1.43E+01	2.65E+01	9.30E+00	5.10E+00	**7.96E+00**	1.74E+01
17	2.48E+02	1.66E+02	3.23E+01	5.93E+00	7.91E+01	1.34E+01	1.00E+02	5.99E+01	3.12E+01	3.60E+00	5.99E+01	4.26E+01	2.87E+01	2.57E+00	2.66E+01	4.19E+00	3.41E+01	7.02E+00	**2.29E+01**	2.72E+00
18	4.79E+01	4.72E+00	2.19E+01	1.62E+00	2.08E+01	3.46E-01	6.89E+05	7.16E+05	2.24E+01	9.28E+00	9.57E+01	4.23E+01	2.01E+01	3.13E+00	2.04E+01	2.86E+00	2.04E+01	2.32E+00	**1.95E+01**	4.80E+00
19	3.10E+01	2.69E+00	5.22E+00	1.33E+00	5.93E+00	2.27E+00	5.37E+03	4.46E+03	5.55E+00	2.54E+00	9.26E+01	2.61E+01	3.36E+00	6.21E-01	3.62E+00	7.05E-01	4.33E+00	5.48E-01	**2.39E+00**	3.98E-01
20	9.30E+01	9.14E+01	3.18E+01	4.94E+00	9.41E+01	2.04E+01	1.96E+02	7.16E+01	3.36E+01	7.04E+00	4.58E+02	2.87E+02	**6.51E+00**	9.13E+00	3.04E+01	2.86E+01	2.04E+01	4.27E+00	2.03E+01	4.90E+00
21	3.79E+02	9.77E+00	2.07E+02	1.65E+00	2.13E+02	6.88E+00	2.26E+02	6.61E+00	2.09E+02	8.99E+00	2.08E+02	3.44E+00	2.07E+02	3.26E+00	2.03E+02	2.11E+00	2.01E+02	8.90E-01	**1.94E+02**	3.23E+00
22	2.17E+03	3.25E+03	**1.00E+02**	0.00E+00	**1.00E+02**	0.00E+00	1.01E+02	1.56E+00	1.12E+02	1.61E+01	1.89E+02	1.54E+02	**1.00E+02**	0.00E+00	**1.00E+02**	0.00E+00	**1.00E+02**	1.60E-06	**1.00E+02**	0.00E+00
23	5.35E+02	8.48E+00	3.50E+02	2.71E+00	3.57E+02	9.81E+00	3.85E+02	9.68E+00	3.18E+02	1.30E+01	3.62E+02	6.50E+00	3.48E+02	7.85E+00	3.56E+02	9.23E+00	3.54E+02	6.14E+00	**3.36E+02**	6.81E+00
24	6.04E+02	8.24E+00	4.26E+02	1.56E+00	4.29E+02	2.81E+00	4.44E+02	9.12E+00	4.29E+02	9.74E+00	4.28E+02	2.66E+00	4.19E+02	5.55E+00	4.20E+02	4.30E+00	4.19E+02	1.14E+01	**4.05E+02**	9.64E+00
25	3.92E+02	2.42E-02	3.89E+02	2.66E-02	3.89E+02	1.19E-02	4.23E+02	1.56E+01	3.80E+02	1.81E+00	**3.78E+02**	1.88E-02	3.87E+02	6.22E-03	3.90E+02	6.14E-01	3.87E+02	2.73E-03	3.88E+02	1.46E-01
26	2.80E+03	1.35E+02	9.24E+02	3.80E+01	9.79E+02	4.92E+01	1.38E+03	3.56E+02	9.47E+02	1.20E+02	9.64E+02	2.27E+02	6.93E+02	1.99E+02	7.72E+02	1.21E+02	7.55E+02	9.51E+01	**5.96E+02**	1.42E+02
27	4.97E+02	9.67E+00	5.03E+02	5.91E+00	5.01E+02	6.18E+00	5.72E+02	9.14E+00	4.99E+02	6.73E+00	5.00E+02	2.45E-04	**4.96E+02**	1.49E+01	5.88E+02	8.18E+00	5.11E+02	4.41E+00	5.00E+02	9.06E+00
28	3.55E+02	5.15E+01	3.42E+02	5.52E+01	**3.07E+02**	2.72E+01	5.14E+02	1.61E+01	4.71E+02	3.28E+01	5.00E+02	4.91E-04	3.17E+02	4.04E+01	3.84E+02	4.47E+01	3.43E+02	5.35E+01	3.13E+02	3.46E+01
29	8.82E+02	2.09E+02	4.34E+02	5.37E+00	5.40E+02	3.22E+01	6.08E+02	1.32E+02	4.32E+02	6.79E+01	**3.59E+02**	4.21E+01	4.00E+02	9.82E+00	4.32E+02	3.41E+01	4.10E+02	1.09E+01	4.08E+02	7.32E+00
30	2.19E+03	1.12E+02	2.18E+03	4.36E+01	2.17E+03	1.01E+01	2.45E+04	1.34E+04	2.57E+03	1.32E+03	**7.54E+02**	1.54E+03	1.98E+03	2.43E+01	2.63E+03	1.81E+02	1.99E+03	5.17E+01	2.06E+03	2.54E+01

Table 4

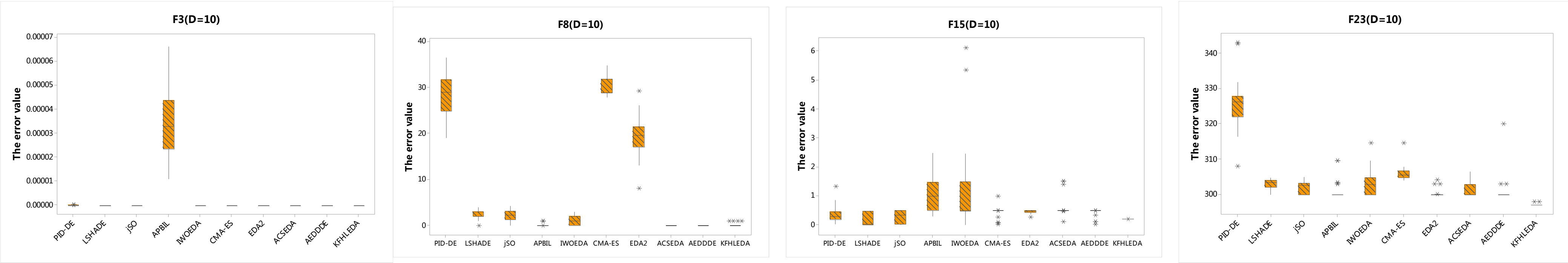
The comparison results of all algorithms (50D)

Fun	PID-DE		LSHADE		jSO		APBIL		IWOEDA		CMA-ES		EDA^2		ACSEDA		AEDDDE		KFHLEDA	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
1	5.85E+03	6.49E+03	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	1.49E+06	1.65E+06	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	1.63E-02	3.37E-03	**0.00E+00**	0.00E+00
2	9.33E+29	2.26E+29	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	1.00E+30	2.86E+14	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	4.47E+05	1.88E+06	1.55E+13	8.03E+13	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00
3	1.73E+05	2.57E+04	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	5.20E+04	1.12E+04	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	2.13E-07	0.00E+00	**0.00E+00**	0.00E+00
4	1.03E+02	5.71E+01	9.73E+01	4.46E+01	6.65E+01	5.06E+01	2.65E+02	3.19E+01	2.09E+02	5.25E+01	**0.00E+00**	0.00E+00	8.61E+01	5.09E+01	2.15E+02	3.24E+01	6.49E+00	4.27E+01	1.45E+01	6.57E+00
5	3.69E+02	1.56E+01	1.21E+01	2.49E+00	1.51E+01	5.77E+00	5.34E+01	1.47E+01	1.09E+01	1.75E+01	1.33E+01	3.32E+00	2.60E+01	5.62E+00	3.65E+00	2.41E+00	**6.94E-01**	1.04E+00	3.53E+00	1.49E+00
6	1.04E-04	4.25E-05	4.12E-05	2.12E-04	9.56E-06	1.07E-06	2.88E-07	5.24E-07	**0.00E+00**	0.00E+00	3.95E+00	1.53E+01	2.29E-03	9.43E-03	**0.00E+00**	0.00E+00	7.15E-04	1.17E-04	5.52E-06	1.73E-06
7	4.22E+02	1.24E+01	6.35E+01	1.63E+00	1.48E+02	4.27E+01	8.53E+01	7.30E+00	5.85E+01	1.80E+02	6.19E+01	2.34E+00	1.00E+02	1.19E+01	5.66E+01	1.38E+00	3.27E+02	1.51E+01	**5.57E+01**	1.24E+00
8	3.66E+02	1.30E+01	1.22E+01	2.31E+00	1.56E+01	7.02E+00	5.22E+01	1.09E+01	1.09E+01	1.75E+01	1.19E+01	2.31E+00	2.46E+01	7.76E+00	4.29E+00	3.25E+00	3.25E+00	1.89E+00	**2.71E+00**	1.61E+00
9	1.10E-01	4.14E-01	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	7.35E+00	5.98E+00	**0.00E+00**	0.00E+00	1.22E+04	8.68E+01	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	1.05E-07	0.00E+00	**0.00E+00**	0.00E+00
10	1.34E+04	4.11E+02	3.16E+03	2.70E+02	8.94E+03	5.70E+02	3.51E+03	7.09E+02	8.02E+02	1.84E+02	3.71E+03	9.70E+02	1.02E+03	3.64E+02	5.58E+02	3.44E+02	3.98E+03	8.74E+02	**2.06E+02**	1.71E+02
11	1.57E+02	1.39E+01	4.75E+01	8.56E+00	2.67E+01	3.66E+00	1.66E+03	9.88E+02	1.83E+02	1.33E+02	1.13E+02	3.31E+01	3.03E+01	4.20E+00	6.68E+01	3.12E+01	2.02E+01	1.20E+00	**1.91E+01**	9.21E-01
12	1.72E+05	1.18E+05	2.35E+03	4.87E+02	1.67E+03	4.89E+02	7.57E+06	4.22E+06	1.57E+03	4.87E+02	2.63E+03	4.04E+02	1.61E+03	4.41E+02	**4.54E+02**	1.67E+02	1.31E+03	4.73E+02	4.62E+02	1.94E+02
13	1.09E+03	1.98E+03	5.78E+01	3.57E+01	3.66E+01	1.83E+01	2.65E+03	1.90E+03	7.78E+02	5.70E+01	7.21E+01	4.93E+01	7.28E+01	3.49E+01	8.43E+01	4.58E+01	1.32E+02	4.25E+01	**1.17E+01**	1.44E+01
14	1.40E+02	8.05E+00	2.96E+01	3.07E+00	2.42E+01	1.75E+00	9.92E+05	6.34E+05	3.24E+01	1.62E+01	2.74E+02	9.02E+01	2.34E+01	1.85E+00	2.14E+01	1.50E+00	2.05E+01	2.90E+00	**1.89E+01**	5.85E+00
15	1.27E+02	1.04E+01	4.04E+01	9.29E+00	2.40E+01	2.42E+00	3.89E+03	2.43E+03	1.04E+02	9.88E+01	4.15E+02	1.17E+02	2.37E+01	3.18E+00	2.60E+01	4.15E+00	**1.78E+01**	1.15E+00	1.79E+01	1.45E+00
16	2.98E+03	2.54E+02	3.83E+02	1.51E+02	7.57E+02	1.41E+02	9.23E+02	3.08E+02	1.84E+02	1.00E+02	2.00E+02	9.96E+01	1.73E+02	8.60E+01	1.28E+02	1.89E+00	1.32E+02	1.01E+01	**1.27E+02**	1.81E+00
17	1.76E+03	2.41E+02	2.53E+02	7.73E+01	4.97E+02	1.04E+02	5.30E+02	2.02E+02	1.77E+02	7.59E+01	3.39E+02	1.35E+02	**6.98E+01**	6.59E+01	1.33E+02	8.71E+01	2.34E+02	1.03E+02	1.03E+02	6.77E+01
18	1.20E+03	1.08E+03	4.14E+01	1.39E+01	2.46E+01	2.05E+00	2.73E+06	1.42E+06	3.08E+01	9.15E+00	2.50E+02	9.44E+01	2.19E+01	1.11E+00	**2.10E+01**	1.84E-01	2.13E+01	1.03E-01	**2.10E+01**	2.83E-01
19	7.29E+01	4.43E+00	2.59E+01	6.53E+00	1.53E+01	3.51E+00	1.32E+04	3.68E+03	1.80E+01	3.33E+00	1.40E+02	5.99E+01	1.04E+01	2.38E+00	7.58E+00	2.08E+00	1.08E+01	1.69E+00	**6.67E+00**	1.18E+00
20	1.60E+03	2.54E+02	1.54E+02	5.31E+01	3.74E+02	7.09E+01	3.37E+02	2.42E+02	6.51E+01	4.02E+01	1.55E+03	2.23E+02	2.49E+01	3.35E+00	3.56E+01	1.73E+01	8.12E+01	2.95E+01	**2.47E+01**	3.66E+00
21	5.67E+02	1.42E+01	2.13E+02	2.41E+00	2.18E+02	4.63E+00	2.50E+02	1.29E+01	**2.03E+02**	2.76E+01	2.14E+02	4.09E+00	2.25E+02	5.92E+00	2.06E+02	2.91E+00	2.03E+02	1.78E+00	2.06E+02	1.86E+00
22	1.33E+04	1.93E+03	2.13E+03	1.80E+03	3.45E+03	4.30E+03	2.53E+03	2.57E+03	1.16E+02	1.10E+01	6.72E+02	3.44E+02	1.24E+02	1.71E+02	2.80E+02	3.15E+02	**1.00E+02**	1.69E-04	**1.00E+02**	0.00E+00
23	7.90E+02	1.33E+01	4.31E+02	4.26E+00	4.44E+02	8.20E+00	5.06E+02	2.11E+01	4.24E+02	1.89E+01	4.40E+02	1.12E+01	4.39E+02	1.49E+01	4.26E+02	1.20E+01	4.20E+02	7.91E+00	**3.98E+02**	6.63E+00
24	8.61E+02	1.33E+01	5.07E+02	2.75E+00	5.12E+02	4.67E+00	5.66E+02	1.83E+01	5.04E+02	5.80E+01	5.11E+02	4.11E+00	5.11E+02	1.52E+01	4.92E+02	7.79E+00	5.02E+02	5.32E+00	**4.55E+02**	3.75E+01
25	4.96E+02	2.95E+01	4.82E+02	6.12E+00	4.83E+02	1.49E+01	7.03E+02	4.02E+01	5.58E+02	2.11E+01	**4.31E+02**	1.01E-02	4.82E+02	4.51E+00	5.40E+02	2.40E+01	4.80E+02	6.89E-01	4.81E+02	3.68E-01
26	4.60E+03	1.80E+02	1.16E+03	5.05E+01	1.17E+03	6.46E+01	2.16E+03	2.62E+02	1.02E+03	1.55E+02	1.60E+03	1.96E+02	1.09E+03	2.02E+02	7.14E+02	1.31E+02	6.02E+02	1.74E+02	**4.55E+02**	9.57E+01
27	5.56E+02	6.66E+01	5.35E+02	2.06E+01	5.13E+02	1.07E+01	8.74E+02	4.37E+01	5.10E+02	7.55E+01	**5.00E+02**	6.00E-04	5.70E+02	8.48E+01	6.64E+02	1.29E+01	5.38E+02	1.09E+01	5.12E+02	7.76E+00
28	4.70E+02	1.96E+01	4.83E+02	2.47E+01	4.60E+02	8.77E+00	8.88E+02	6.68E+01	5.44E+02	3.80E+01	5.00E+02	5.42E-04	4.64E+02	1.52E+01	4.89E+02	2.55E+01	**4.59E+02**	6.03E-01	**4.59E+02**	5.95E-01
29	1.81E+03	2.08E+02	3.51E+02	1.15E+01	5.89E+02	6.98E+01	9.18E+02	1.91E+02	4.16E+02	2.53E+01	4.28E+02	1.21E+02	3.39E+02	3.07E+01	3.83E+02	3.98E+01	3.36E+02	3.92E+01	**3.14E+02**	1.99E+01
30	6.03E+05	3.51E+04	6.68E+05	9.53E+04	6.14E+05	3.94E+04	3.90E+06	8.43E+05	6.06E+05	8.62E+04	**4.78E+02**	4.68E+02	6.34E+05	4.80E+04	2.13E+06	2.19E+05	6.23E+05	5.07E+04	8.31E+05	3.08E+04

Table 5

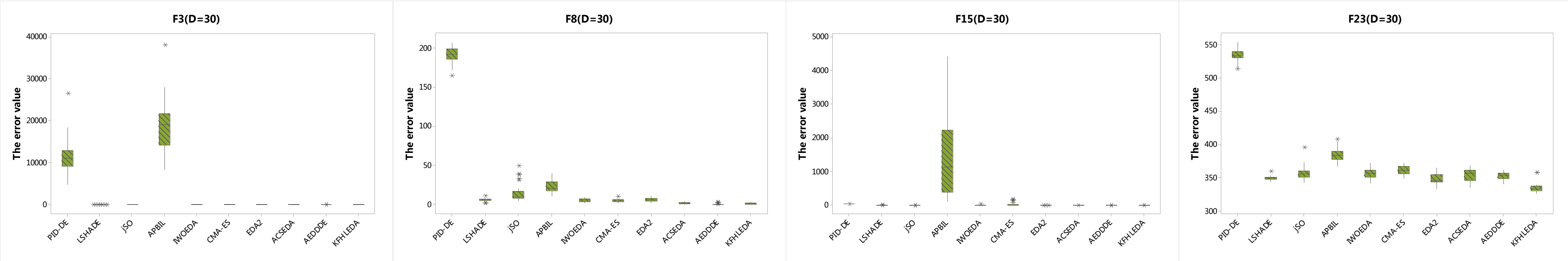
The comparison results of all algorithms (100D)

Fun	PID-DE		LSHADE		jSO		APBIL		IWOEDA		CMA-ES		EDA^2		ACSEDA		AEDDDE		KFHLEDA	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
1	6.63E+03	8.40E+03	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	8.69E+07	1.20E+08	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	2.90E+09	2.19E+09	6.35E-02	6.50E-03	6.26E+01	1.53E+01	**0.00E+00**	0.00E+00
2	1.00E+30	2.84E+14	2.29E+02	1.22E+03	2.05E+02	5.85E+02	1.00E+30	2.86E+14	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	1.00E+30	2.84E+14	1.00E+30	2.84E+14	1.82E+12	1.25E+13	2.01E+02	6.97E+01
3	5.54E+05	4.74E+04	8.49E-07	1.20E-06	**0.00E+00**	1.00E-07	2.30E+05	2.53E+04	**0.00E+00**	0.00E+00	**0.00E+00**	0.00E+00	2.84E+04	9.84E+03	1.58E-01	1.10E+00	3.24E-04	5.41E-05	**0.00E+00**	0.00E+00
4	2.32E+02	3.14E+01	1.96E+02	9.65E+00	1.94E+02	1.75E+01	6.49E+02	1.38E+02	4.13E+02	1.01E+02	**1.73E+01**	5.50E-01	4.04E+02	7.59E+01	4.45E+02	2.73E+01	2.21E+02	2.42E+00	1.89E+02	4.21E+00
5	8.42E+02	2.73E+01	3.92E+01	5.59E+00	3.22E+01	4.96E+00	1.57E+02	3.70E+01	1.89E+01	8.44E+00	3.45E+01	5.59E+00	3.04E+02	3.76E+01	2.85E+01	1.05E+01	6.20E+01	7.75E+01	**1.09E+01**	3.78E+00
6	6.51E-03	1.91E-02	7.57E-03	4.50E-03	1.58E-03	3.59E-04	6.45E-03	9.12E-03	**4.67E-05**	1.76E-04	2.46E+01	2.94E+01	9.89E+00	2.52E+00	1.01E-03	4.73E-05	7.32E-03	6.98E-04	3.27E-04	7.86E-04
7	9.60E+02	2.90E+01	3.41E+02	3.84E+00	3.50E+02	3.02E+01	2.03E+02	2.17E+01	**1.16E+02**	7.02E+01	1.33E+02	4.68E+00	6.37E+02	7.65E+01	4.83E+02	1.18E+02	8.07E+02	1.74E+01	6.11E+02	3.54E+01
8	8.41E+02	1.89E+01	3.68E+01	4.37E+00	3.03E+01	5.29E+00	1.71E+02	3.16E+01	3.08E+01	1.13E+01	3.41E+01	5.35E+00	3.13E+02	3.77E+01	2.19E+01	5.55E+01	4.50E+01	7.38E+01	**1.46E+01**	5.00E+00
9	1.18E+01	1.73E+01	4.29E-01	4.20E-01	2.02E-02	3.81E-02	3.29E+02	1.02E+02	4.56E-06	1.61E-05	2.21E+04	9.08E+01	1.30E+03	8.08E+02	**4.55E-07**	0.00E+00	3.63E-04	5.35E-05	1.27E-04	2.35E-05
10	3.02E+04	5.57E+02	1.04E+04	5.44E+02	2.30E+04	7.71E+02	1.04E+04	1.41E+03	6.38E+03	9.99E+02	9.47E+03	1.52E+03	8.18E+03	1.03E+03	5.88E+03	4.02E+03	1.94E+04	1.02E+03	**5.31E+03**	1.82E+03
11	7.00E+02	8.63E+01	4.35E+02	1.12E+02	1.08E+02	3.41E+01	2.66E+04	4.92E+03	6.40E+02	3.24E+01	1.08E+03	3.79E+02	5.92E+02	1.23E+02	2.39E+02	1.06E+02	1.29E+02	5.61E+01	**3.64E+01**	2.52E+01
12	9.15E+05	3.90E+05	2.18E+04	7.39E+03	1.83E+04	8.96E+03	1.26E+08	2.92E+07	4.55E+03	8.86E+02	5.36E+03	6.75E+02	9.54E+05	8.50E+05	2.66E+03	3.68E+02	7.60E+03	4.60E+02	**2.34E+03**	5.12E+02
13	6.26E+03	6.45E+03	5.18E+02	2.46E+02	**1.42E+02**	4.16E+01	1.83E+05	9.73E+05	1.65E+03	2.57E+02	4.01E+03	8.88E+02	4.57E+03	1.22E+03	3.09E+02	8.55E+01	1.56E+02	3.83E+01	2.32E+02	3.11E+01
14	1.53E+04	2.14E+04	2.58E+02	3.16E+01	6.52E+01	1.02E+01	6.25E+06	2.76E+06	5.91E+01	2.38E+01	3.72E+02	1.35E+02	8.94E+01	2.04E+01	2.87E+01	2.53E+01	7.29E+01	5.77E+01	**2.35E+01**	1.68E+00
15	5.92E+03	7.50E+03	2.48E+02	4.62E+01	1.64E+02	4.20E+01	1.31E+03	4.46E+02	2.31E+02	2.39E+01	4.13E+02	8.31E+01	4.47E+02	7.62E+01	9.76E+01	3.14E+01	1.71E+02	4.64E+01	**7.31E+01**	2.62E+01
16	7.92E+03	3.23E+02	1.68E+03	2.61E+02	2.84E+03	4.78E+02	3.16E+03	7.02E+02	2.74E+02	2.31E+02	4.45E+02	2.93E+02	1.01E+03	2.97E+02	2.94E+02	1.83E+02	1.08E+03	4.80E+02	**1.40E+02**	1.21E+02
17	5.03E+03	2.31E+02	1.10E+03	2.22E+02	1.86E+03	2.57E+02	1.97E+03	4.45E+02	**2.49E+02**	4.36E-01	1.26E+03	3.60E+02	4.64E+02	1.67E+02	5.49E+02	2.74E+02	1.19E+03	3.68E+02	3.61E+02	7.67E+01
18	2.33E+05	1.21E+05	2.42E+02	4.79E+01	1.87E+02	3.71E+01	4.97E+06	1.93E+06	5.33E+01	3.50E+01	2.25E+02	4.59E+01	2.85E+02	5.65E+01	2.38E+01	2.05E+00	3.52E+01	1.39E+01	**2.32E+01**	1.13E+00
19	8.63E+03	1.02E+04	1.69E+02	2.29E+01	1.06E+02	2.04E+01	7.52E+04	2.91E+05	7.42E+01	3.16E+01	3.10E+02	7.75E+01	1.58E+02	3.46E+01	3.91E+01	1.54E+01	7.40E+01	2.57E+01	**3.36E+01**	3.41E+00
20	4.96E+03	1.86E+02	1.54E+03	1.93E+02	2.16E+03	2.07E+02	1.64E+03	4.69E+02	**2.13E+02**	2.14E+02	3.82E+03	3.29E+02	3.74E+02	8.02E+01	5.99E+02	2.96E+02	1.11E+03	2.46E+02	4.15E+02	1.01E+02
21	1.07E+03	2.29E+01	2.57E+02	7.21E+00	2.58E+02	7.17E+00	3.82E+02	3.08E+01	2.69E+02	2.29E+01	2.58E+02	4.85E+00	5.28E+02	3.79E+01	4.22E+02	2.15E+02	2.96E+02	6.95E+01	**2.32E+02**	4.39E+00
22	3.08E+04	5.70E+02	1.13E+04	5.92E+02	2.32E+04	9.16E+02	1.16E+04	2.97E+03	2.90E+03	1.58E+03	1.73E+03	8.00E+02	6.22E+03	4.54E+03	2.57E+03	2.29E+03	1.15E+04	7.84E+03	**1.25E+03**	1.64E+03
23	1.33E+03	4.78E+01	5.69E+02	9.89E+00	5.77E+02	9.65E+00	8.13E+02	5.79E+01	7.31E+02	2.37E+01	5.67E+02	1.23E+01	9.57E+02	5.47E+01	9.29E+02	1.48E+02	5.42E+02	1.15E+01	**5.37E+02**	1.10E+01
24	1.70E+03	4.01E+01	9.09E+02	7.56E+00	9.13E+02	1.07E+01	1.10E+03	5.28E+01	1.11E+03	1.98E+02	8.96E+02	5.50E+00	1.50E+03	1.05E+02	9.62E+02	2.03E+02	8.93E+02	8.92E+00	**3.50E+02**	1.28E+02
25	7.59E+02	5.07E+01	7.46E+02	3.27E+01	7.35E+02	3.69E+01	1.46E+03	1.34E+02	9.74E+02	7.22E+01	7.28E+02	3.50E+01	1.09E+03	1.14E+02	8.27E+02	7.14E+01	6.79E+02	3.28E+01	**6.34E+02**	1.43E+01
26	1.13E+04	3.42E+02	3.30E+03	9.88E+01	3.30E+03	1.08E+02	6.35E+03	7.89E+02	3.03E+03	1.15E+02	3.28E+03	1.57E+02	6.08E+03	1.81E+03	2.56E+03	1.93E+02	2.66E+03	1.38E+02	**1.64E+03**	2.77E+02
27	6.28E+02	2.53E+01	6.30E+02	2.07E+01	5.90E+02	1.44E+01	1.07E+03	6.38E+01	7.47E+02	8.77E+01	**5.00E+02**	4.78E-04	7.15E+02	4.98E+01	8.50E+02	4.06E+01	6.75E+02	1.46E+01	5.39E+02	6.84E+00
28	5.72E+02	3.74E+01	5.23E+02	2.50E+01	5.21E+02	3.11E+01	2.17E+03	3.22E+02	8.84E+02	2.35E+02	**5.00E+02**	4.94E-04	7.14E+02	5.91E+01	7.47E+02	4.47E+01	5.49E+02	2.47E+01	5.16E+02	2.61E+01
29	5.54E+03	3.54E+02	1.24E+03	1.89E+02	2.20E+03	3.76E+02	3.06E+03	5.99E+02	**6.77E+02**	1.54E+02	1.71E+03	3.83E+02	1.68E+03	2.92E+02	1.13E+03	1.11E+02	1.27E+03	1.91E+02	9.99E+02	9.79E+01
30	4.76E+03	3.08E+03	2.39E+03	1.35E+02	2.34E+03	1.56E+02	2.09E+06	3.94E+06	2.12E+04	3.97E+03	**1.35E+03**	3.28E+02	6.85E+03	1.46E+03	5.73E+03	7.92E+02	2.39E+03	2.72E+02	2.31E+03	4.55E+01



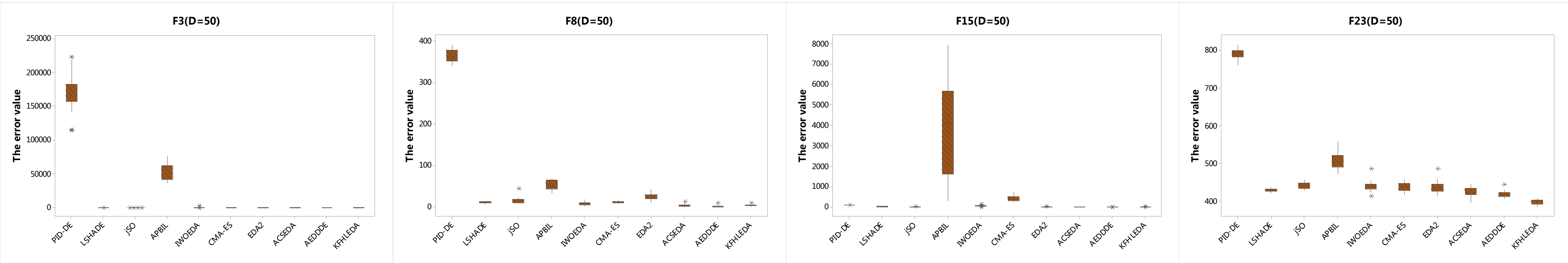
(a). f_3 (b). f_8 (c). f_{15} (d). f_{23}

Fig.5 Box plots of four typical functions (10D)



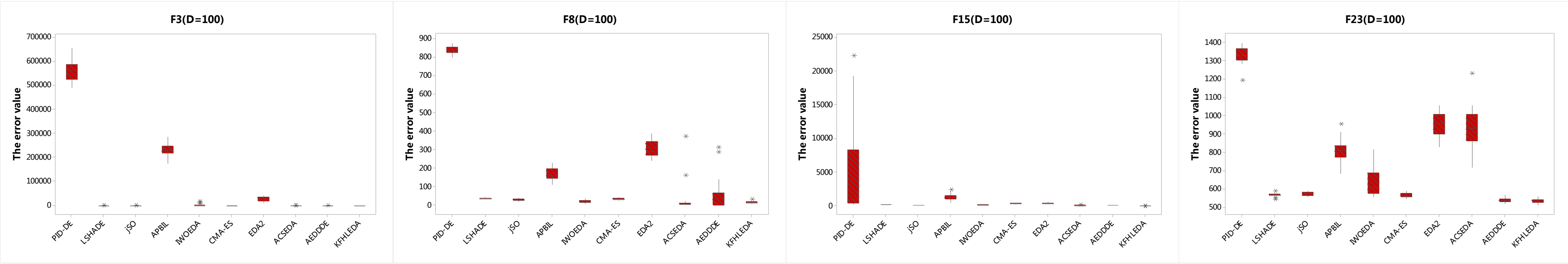
(a). f_3 (b). f_8 (c). f_{15} (d). f_{23}

Fig.6 Box plots of four typical functions (30D)



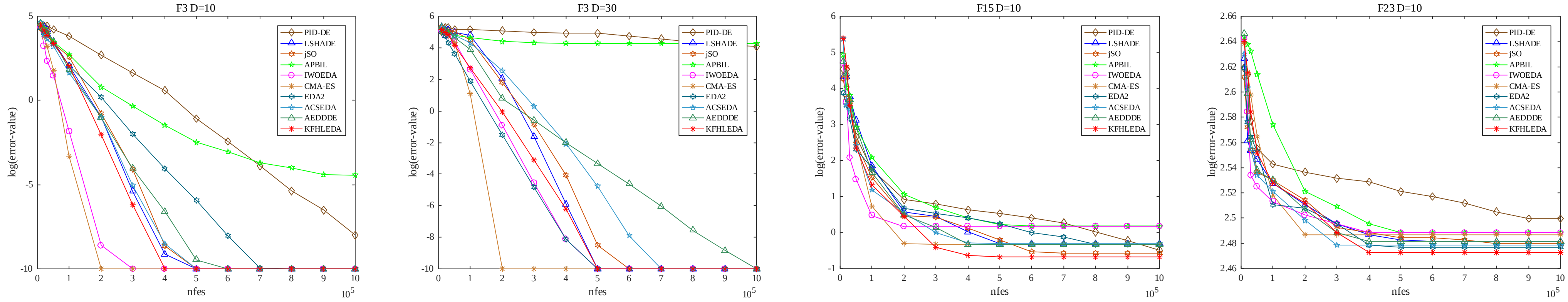
(a). f_3 (b). f_8 (c). f_{15} (d). f_{23}

Fig.7 Box plots of four typical functions (50D)



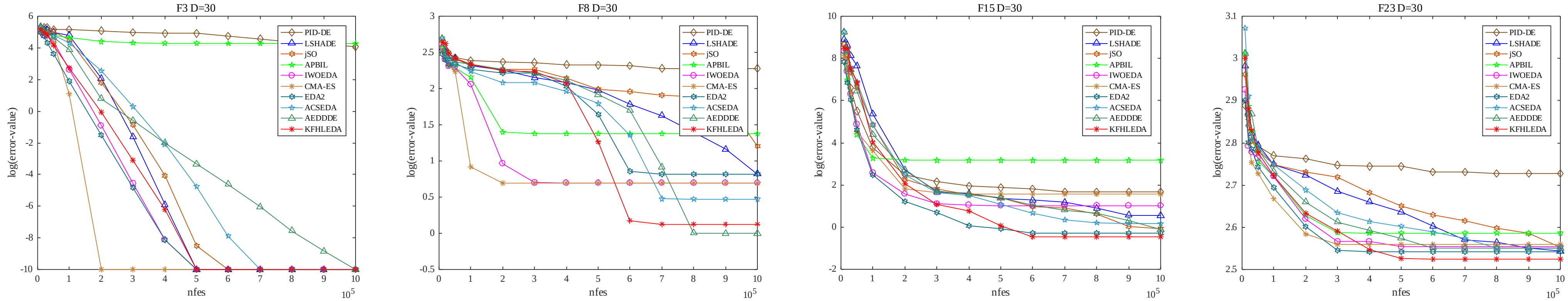
(a). f_3 **(b).** f_8 **(c).** f_{15} **(d).** f_{23}

Fig.8 Box plots of four typical functions (100D)



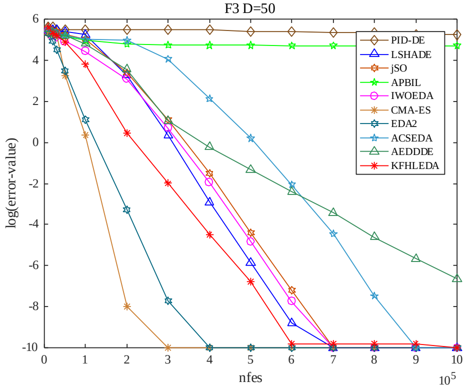
(a). f_3 **(b).** f_8 **(c).** f_{15} **(d).** f_{23}

Fig.9 Convergence curves of four typical functions (10D)

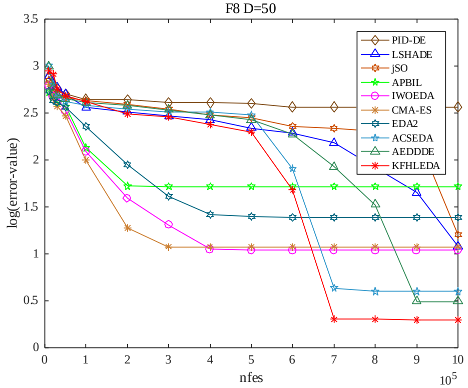


(a). f_3 **(b).** f_8 **(c).** f_{15} **(d).** f_{23}

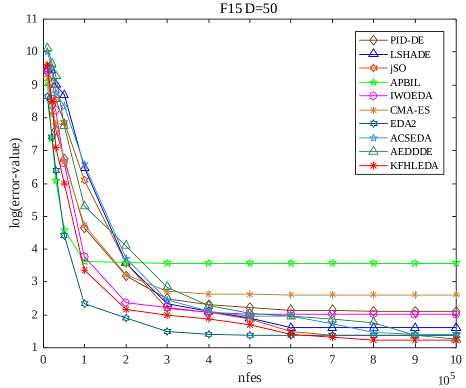
Fig.10 Convergence curves of four typical functions (30D)



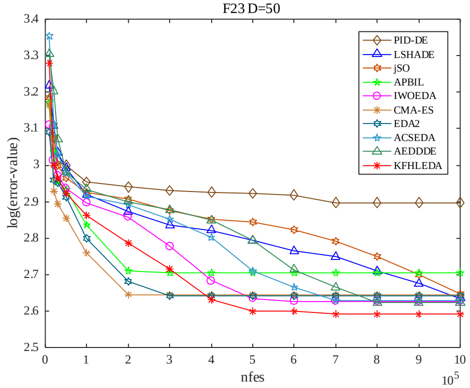
(a). f_3



(b). f_8

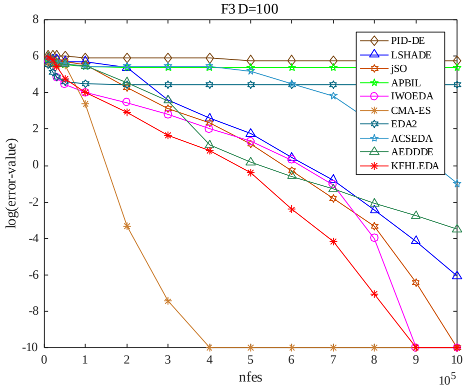


(c). f_{15}

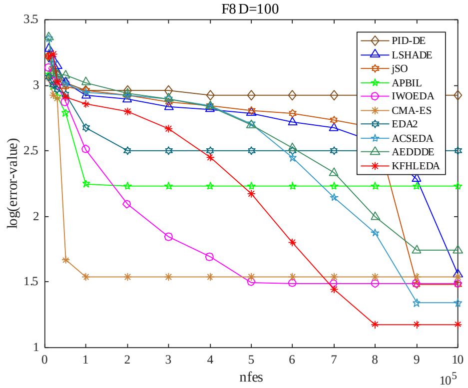


(d). f_{23}

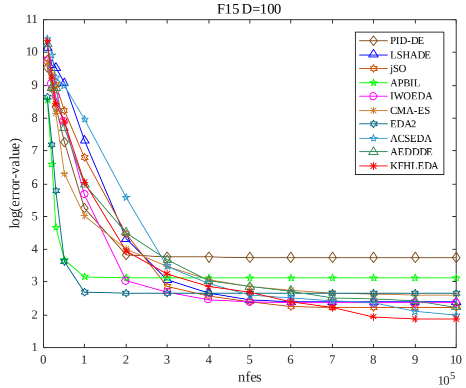
Fig.11 Convergence curves of four typical functions (50D)



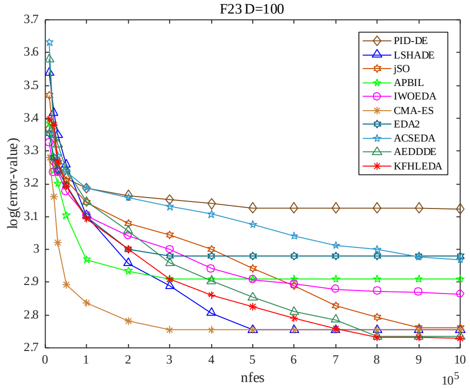
(a). f_3



(b). f_8

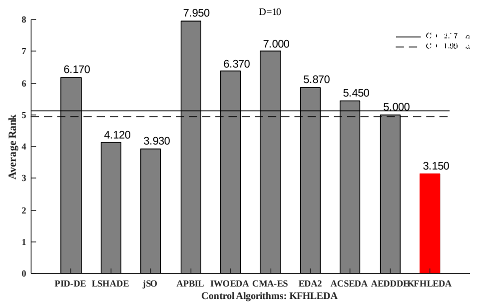


(c). f_{15}

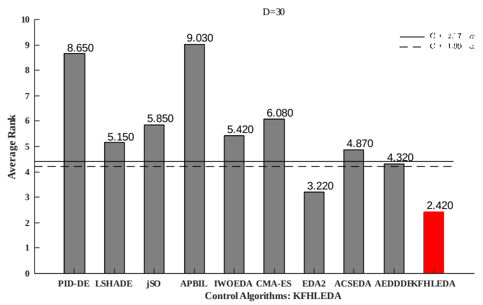


(d). f_{23}

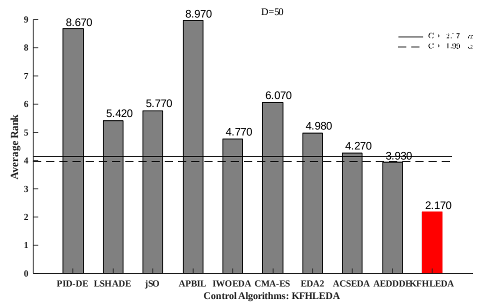
Fig.12 Convergence curves of four typical functions (100D)



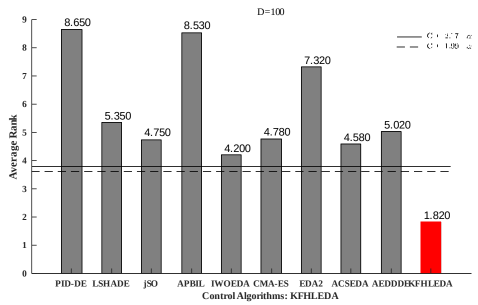
(a).D=10



(b).D=30



(c). D=50



(d). D=100

Fig. 13. Rankings obtained through Friedman test

Table 6

Friedman-Test results of ten algorithms.

Algorithms	Mean rank			
	10*D*	30*D*	50*D*	100*D*
PID-DE	6.17	8.65	8.67	8.65
LSHADE	4.12	5.15	5.42	5.35
jSO	3.93	5.85	5.77	4.75
APBIL	7.95	9.03	8.97	8.53
IWOEDA	6.37	5.42	4.77	4.20
CMAES	7.00	6.08	6.07	4.78
EDA2	5.87	3.22	4.98	7.32
ACSEDA	5.45	4.87	4.27	4.58
AEDDDE	5.00	4.32	3.93	5.02
KFHLEDA	3.15	2.42	2.17	1.82
Crit. Diff $\alpha = 0.05$	1.99	1.99	1.99	1.99
Crit. Diff $\alpha = 0.10$	1.79	1.79	1.79	1.79

4.5 Wilcoxon symbolic rank test

Table 7

Rankings obtained through the Wilcoxon test.

D	**algorithms**	**vs**	**+**	**-**	**=**	**R+**	**R-**	**Z**	**p-value**	**0.05**	**0.1**
		PID-DE	22	4	4	2.99E+02	5.20E+01	-3.14E+00	2.00E-03	**yes**	**yes**
		LSHADE	15	8	7	1.92E+02	8.40E+01	-1.64E+00	1.01E-01	no	no
		jSO	15	7	8	1.77E+02	7.60E+01	-1.64E+00	1.01E-01	no	no
		APBIL	26	2	2	3.71E+02	3.50E+01	-3.83E+00	0.00E+00	**yes**	**yes**
10	KFHLEDA	IWOEDA	20	3	7	2.40E+02	3.60E+01	-3.10E+00	2.00E-03	**yes**	**yes**
		CMA-ES	20	6	4	2.85E+02	6.60E+01	-2.78E+00	5.00E-03	**yes**	**yes**
		EDA2	21	3	6	2.56E+02	4.40E+01	-3.03E+00	2.00E-03	**yes**	**yes**
		ACSEDA	21	1	8	2.31E+02	2.20E+01	-3.39E+00	1.00E-03	**yes**	**yes**
		AEDDDE	19	5	6	2.63E+02	3.70E+01	-3.23E+00	1.00E-03	**yes**	**yes**
		PID-DE	28	1	1	4.32E+02	3.00E+00	-4.64E+00	0.00E+00	**yes**	**yes**
		LSHADE	24	0	6	3.00E+02	0.00E+00	-4.29E+00	0.00E+00	**yes**	**yes**
		jSO	22	2	6	2.87E+02	1.30E+01	-3.91E+00	0.00E+00	**yes**	**yes**
		APBIL	29	0	1	4.35E+02	0.00E+00	-4.70E+00	0.00E+00	**yes**	**yes**
30	KFHLEDA	IWOEDA	21	4	5	2.85E+02	4.00E+01	-3.30E+00	1.00E-03	**yes**	**yes**
		CMA-ES	22	4	4	2.95E+02	5.60E+01	-3.04E+00	2.00E-03	**yes**	**yes**
		EDA2	17	7	6	2.15E+02	8.55E+01	-1.84E+00	6.50E-02	no	**yes**
		ACSEDA	24	0	6	3.00E+02	0.00E+00	-4.29E+00	0.00E+00	**yes**	**yes**
		AEDDDE	21	5	4	2.90E+02	6.10E+01	-2.91E+00	4.00E-03	**yes**	**yes**
		PID-DE	29	1	0	4.36E+02	2.90E+01	-4.19E+00	0.00E+00	**yes**	**yes**
		LSHADE	25	1	4	3.25E+02	2.60E+01	-3.80E+00	0.00E+00	**yes**	**yes**
		jSO	25	1	4	3.25E+02	2.60E+01	-3.80E+00	0.00E+00	**yes**	**yes**
		APBIL	29	1	0	4.64E+02	1.00E+00	-4.76E+00	0.00E+00	**yes**	**yes**
50	KFHLEDA	IWOEDA	22	4	4	3.18E+02	3.30E+01	-3.62E+00	0.00E+00	**yes**	**yes**
		CMA-ES	23	4	3	3.28E+02	5.00E+01	-3.34E+00	1.00E-03	**yes**	**yes**
		EDA2	25	2	3	3.37E+02	4.10E+01	-3.56E+00	0.00E+00	**yes**	**yes**
		ACSEDA	23	2	5	3.16E+02	9.00E+00	-4.13E+00	0.00E+00	**yes**	**yes**
		AEDDDE	21	6	3	3.00E+02	7.80E+01	-2.67E+00	8.00E-03	**yes**	**yes**
		PID-DE	30	0	0	4.65E+02	0.00E+00	-4.78 E+00	0.00E+00	**yes**	**yes**
		LSHADE	28	1	1	4.16E+02	1.90E+01	-4.29 E+00	0.00E+00	**yes**	**yes**
		jSO	26	2	2	3.72E+02	3.40E+01	-3.85 E+00	0.00E+00	**yes**	**yes**
		APBIL	29	1	0	4.58E+02	7.00E+00	-4.64 E+00	0.00E+00	**yes**	**yes**
100	KFHLEDA	IWOEDA	21	7	2	3.30E+02	7.60E+01	-2.89 E+00	4.00E-03	**yes**	**yes**
		CMA-ES	22	6	2	3.42E+02	6.40E+01	-3.17 E+00	2.00E-03	**yes**	**yes**
		EDA2	29	1	0	4.62E+02	3.00E+00	-4.72 E+00	0.00E+00	**yes**	**yes**
		ACSEDA	28	2	0	4.52E+02	1.30E+01	-4.52 E+00	0.00E+00	**yes**	**yes**
		AEDDDE	29	1	0	4.50E+02	1.50E+01	-4.47 E+00	0.00E+00	**yes**	**yes**

The value of significance level α is set to 0.05 and 0.1. In Table 7, the symbols "“+”, “-” and “=” represent that the performance of KFHLEDA is better than, worse than, or similar to that of the corresponding algorithm, respectively. R+ and R- represent that the sum of rank of KFHLEDA are superior to, and inferior to the comparison algorithm. If p-value is less than α, it indicates that there is a significant difference, indicated by “yes”. From the results of Table 7, except that there are no significant differences between KFHLEDA and LSHADE, jSO on 10D with $\alpha = 0.05\ and\ 0.1$, and also EDA2 on 30D with $\alpha = 0.05$, there are significant differences under all other cases on 10D, 30D, 50D, and 100D with $\alpha = 0.05\ and\ 0.1$.

From the results of two non-parametric tests, it can be seen that KFHLEDA has the best performance on 10D, 30D, 50D, and 100D compared with other comparison algorithms, which also verifies that the enhanced Kalman filtering mechanism, the history learning mechanism and the population adaptive adjustment strategy adopted in this study are conducive to the performance of the algorithm effectively. Kalman filtering feeds the enhanced information back to EDA model through the history learning mechanism. The modified operation of the algorithm improves the quality of solutions, balances the exploration and exploitation capabilities availably.

The experimental results show that KFHLEDA is valid. According to No Free Lunch Theory [57], there is no single algorithm with optimal performance on all problems. Compared to these classical and the state-of-the-art algorithms, KFHLEDA is competitive clearly due to the improvement of the problem feature design based, including the enhanced Kalman filtering operation, history learning mechanism and population adaptive adjustment strategy. The following experiments verify this point by effective components availability analysis.

4.6 Component availability analysis

The proposed KFHLEDA covers three important strategies: the Kalman filtering, the history learning mechanism, and the population adaptive adjustment strategy. This section evaluates the effectiveness of the three strategies.

Three variants of KFHLEDA are adopted and tested against it on CEC 2017 test suit to demonstrate the effectiveness of each strategy. The three variants are denoted as KFHLEDA1, KFHLEDA2 and KFHLEDA3, which represent KFHLEDA without the enhanced Kalman filtering, KFHLEDA without the history learning mechanism and KFHLEDA without the population adaptive adjustment strategy, respectively. The three variants are detailed below.

(1) KFHLEDA1: This variant eliminates the strategy of enhanced Kalman filtering and retains other operations of KFHLEDA. By comparing with it, the effect of the enhanced Kalman filtering is displayed. The mean and variance of the probabilistic model are updated according to Eq. (5)-(6), instead of Eq. (14)-(15) based on the filtering operation, that is, the mean and variance are guided by the dominant information in the historical archive.

(2) KFHLEDA2: This variant omits the history learning mechanism and reserves the other operations of KFHLEDA, which means that the modification of the individuals by the enhanced Kalman filtering operation is no longer based on the historical archive, but the information obtained in the last operation.

(3) KFHLEDA3: This variant removes the population adaptive adjustment strategy, while keeping other operations of KFHLEDA. Therefore, the population size is no longer updated adaptively with iteration, but with a fixed value, which is set to 2000.

From Table 8, the performance of KFHLEDA is optimal as a whole compared with the three variants. KFHLEDA1 obtains the worse results than KFHLEDA2 and KFHLEDA3, all of which are worse than KFHLEDA. KFHLEDA1 is the worst due to lack of the filtering mechanism, and KFHLEDA3 wins the best results with Kalman filtering and historical learning mechanism among three variants. These results show the effectiveness and significance of the selected strategies.

Table 8

The comparison results of 30 functions

Fun	KFHLEDA		KFHLEDA1		KFHLEDA2		KFHLEDA3	
	mean	std	mean	std	mean	std	mean	std
1	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
2	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
3	0.00E+00	0.00E+00	1.14E-04	4.73E-05	0.00E+00	0.00E+00	0.00E+00	0.00E+00
4	1.44E-07	1.03E-06	9.69E+00	1.14E-02	8.47E-07	1.77E-07	7.35E-07	5.15E-06
5	3.12E-01	5.45E-01	1.72E+01	3.58E+00	8.62E+00	1.34E+00	1.99E+00	1.12E+00
6	0.00E+00	0.00E+00	3.82E+00	1.77E-02	0.00E+00	0.00E+00	0.00E+00	0.00E+00
7	1.01E+01	3.42E-01	8.25E+01	1.75E+01	5.62E+01	3.07E+00	1.99E+01	1.04E+00
8	4.15E-02	2.99E-01	1.75E+01	2.66E+00	4.19E+00	5.20E-01	9.95E-01	3.56E-01
9	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00	0.00E+00
10	2.90E+00	4.70E+00	5.12E+02	2.30E+02	5.72E+01	3.36E+01	2.51E+01	1.25E+01
11	0.00E+00	0.00E+00	6.20E+01	1.77E-01	6.71E-01	1.12E-01	0.00E+00	0.00E+00
12	4.14E+01	5.69E+01	2.14E+03	4.25E+02	7.64E+02	1.13E+02	2.31E+02	6.28E+01
13	4.77E+00	2.63E+00	5.23E+02	5.23E+01	5.73E+01	5.40E+00	1.21E+01	4.15E+00
14	2.82E-01	2.80E+00	8.22E+01	1.42E+01	4.17E+01	2.67E+00	2.00E+00	1.03E+00
15	2.08E-01	1.48E-01	1.92E+01	3.02E+00	6.14E+00	3.09E-01	9.89E-01	1.76E-01
16	1.44E-01	2.87E-01	5.56E+01	4.18E+00	8.15E+00	5.02E-01	7.14E-01	5.00E-01
17	1.24E+01	1.00E+01	7.76E+01	2.54E+01	4.16E+01	2.13E+01	3.11E+01	2.03E+01
18	4.77E-01	6.45E-02	2.31E+01	4.39E+00	3.04E+00	4.87E-01	8.20E-01	4.54E-01
19	2.03E-01	2.50E-01	3.65E+02	2.19E+01	8.48E+00	4.67E-01	1.21E+00	3.56E-01
20	1.54E+00	5.37E+00	2.22E+01	1.01E+01	8.97E+00	2.57E+00	4.13E+00	1.61E+00
21	1.55E+02	1.07E+01	3.84E+02	1.01E+02	2.86E+02	4.40E+01	2.38E+02	2.51E+01
22	1.00E+02	0.00E+00	2.74E+02	3.72E+01	2.08E+02	1.84E+01	1.47E+02	1.11E+01
23	2.97E+02	2.53E+01	4.22E+02	2.88E+02	3.43E+02	1.78E+02	3.15E+02	1.60E+02
24	2.28E+02	4.47E+00	5.11E+02	3.79E+01	3.41E+02	4.59E+00	2.81E+02	5.75E+00
25	3.98E+02	1.55E-02	6.18E+02	5.09E+00	5.03E+02	3.37E+00	4.41E+02	2.41E+00
26	3.00E+02	0.00E+00	3.68E+02	1.26E+01	3.34E+02	2.57E+00	3.12E+02	1.21E+00
27	3.81E+02	1.15E+00	4.89E+02	2.93E+01	4.13E+02	4.10E+00	3.92E+02	3.34E+00
28	3.30E+02	1.00E+02	6.72E+02	2.84E+02	5.43E+02	2.30E+02	4.34E+02	2.17E+02
29	2.32E+02	4.03E+00	2.86E+02	4.41E+01	2.57E+02	1.49E+01	2.42E+02	1.27E+01
30	2.13E+05	5.90E+05	2.83E+05	4.97E+05	2.57E+05	4.06E+05	2.38E+05	3.95E+05

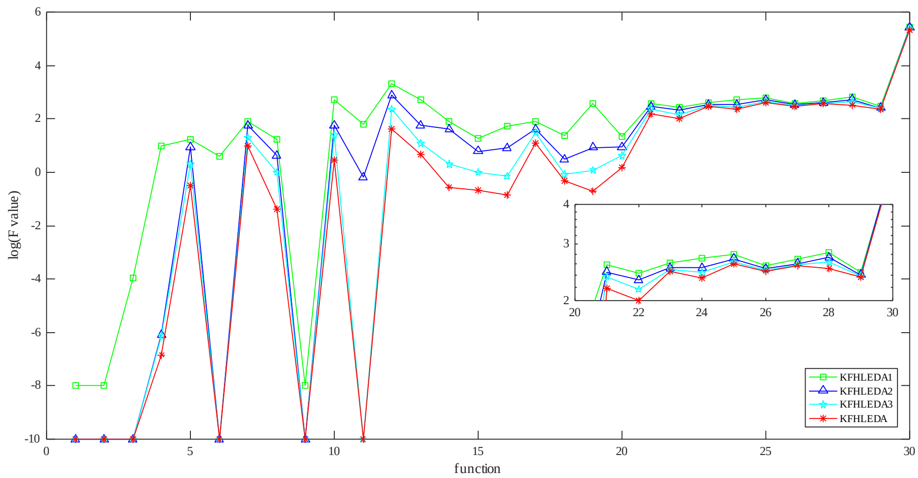


Fig.14 The comparison results of KFHLEDA and the three variants on 30 functions

The population adaptive adjustment strategy makes the proposed algorithm beneficial to exploration in the early stage to avoid premature convergence, and also profitable to exploitation in the late stage, avoiding the invalid search and upgrading the evolution speed. The historical learning mechanism helps the population explore the hidden information in the archive during the evolutionary process, and make it evolve towards the potential region. In the stages of the prediction, observation and two revision operation during the enhanced Kalman filtering, the elite strategy is joined to guide the population to jump out of the local optimum.

4.7 Analysis of fitness landscape

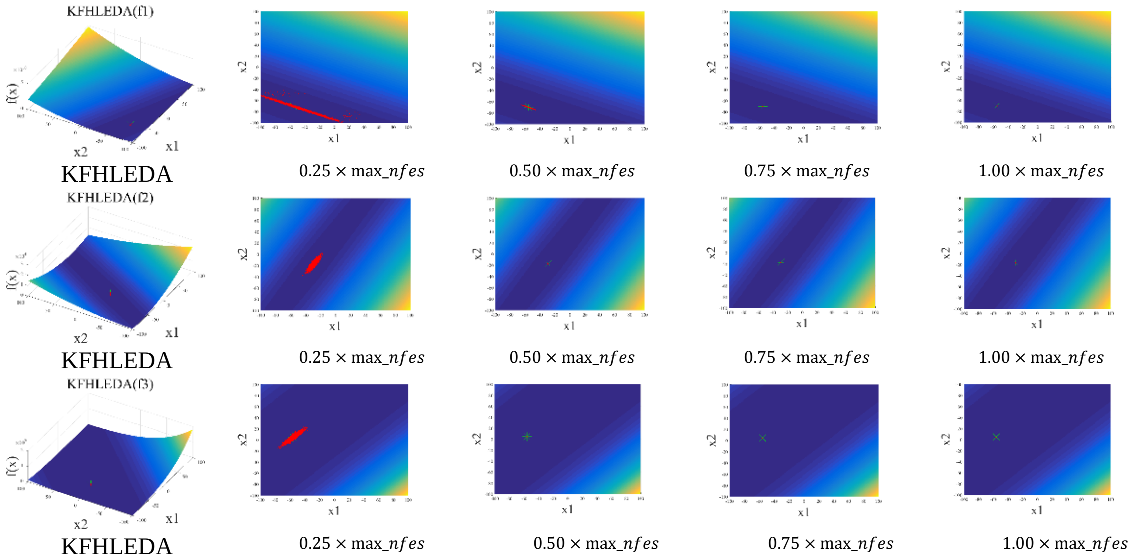
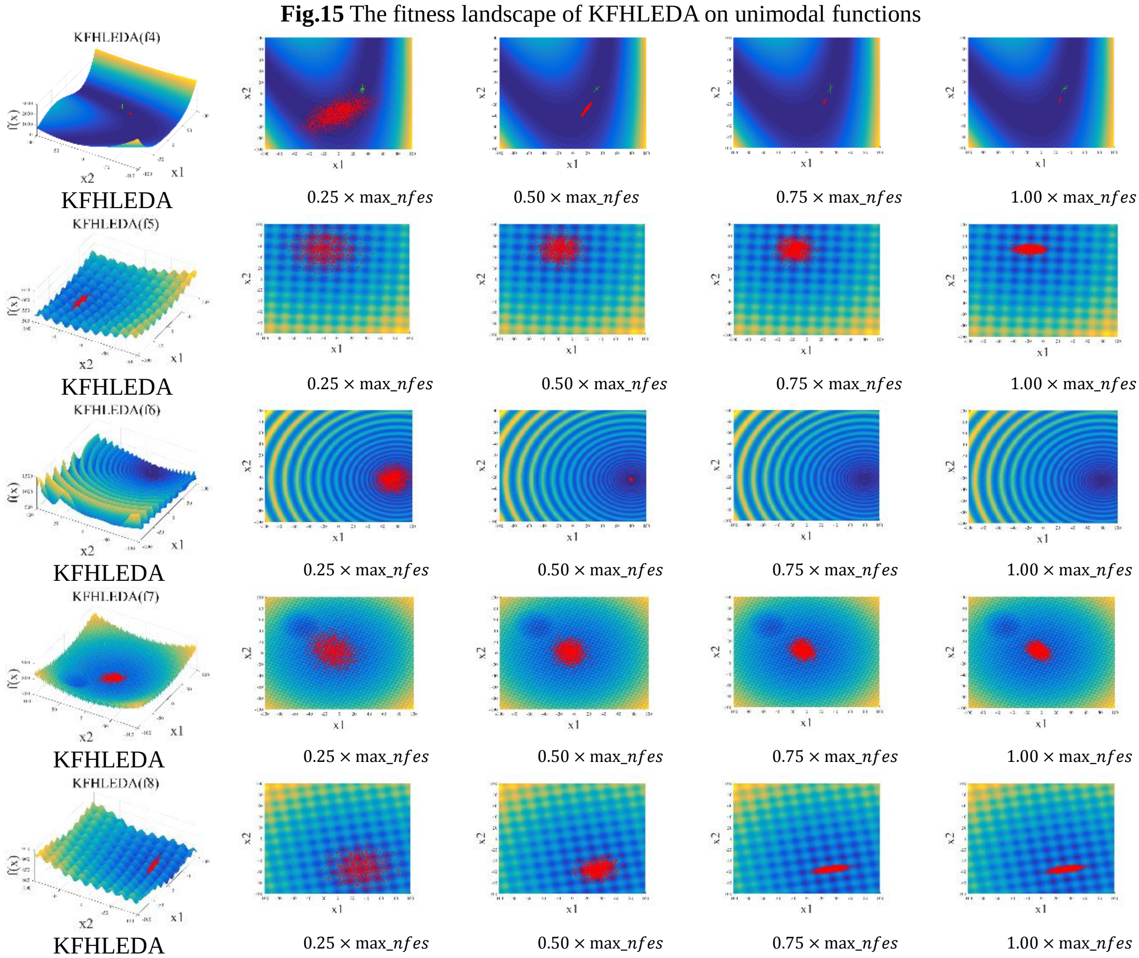


Fig.15 The fitness landscape of KFHLEDA on unimodal functions



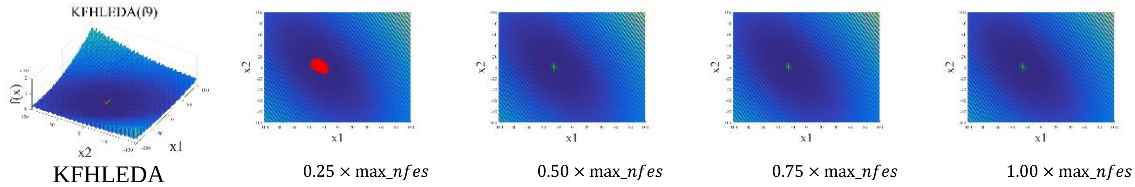


Fig.16 The fitness landscape of KFHLEDA on simple multimodal functions

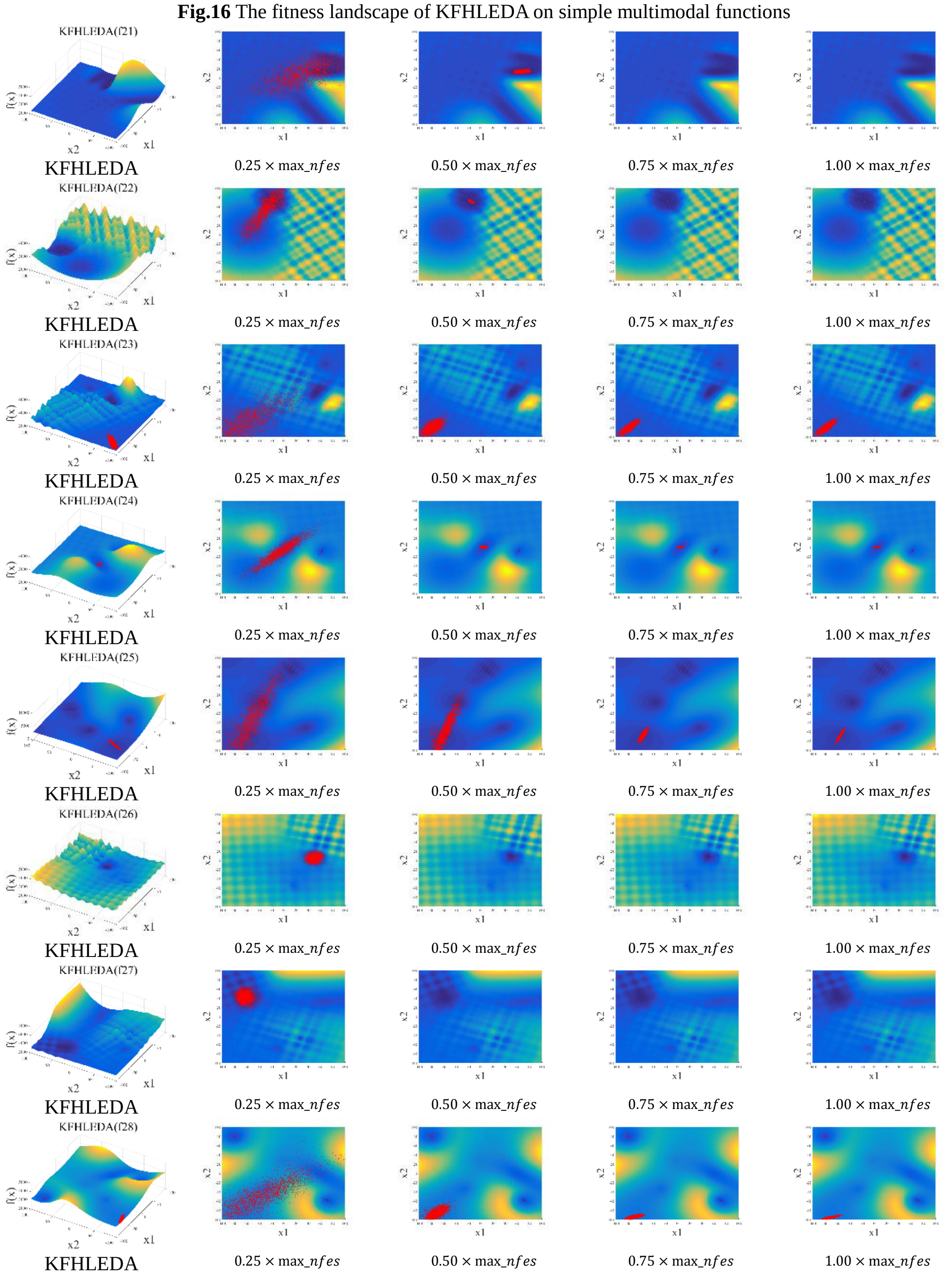


Fig.17 The fitness landscape of KFHLEDA on composition functions

The fitness landscape reproduces the evolution of the population in the search space, and responds to the specific phenomena with the progress of iteration intuitively [58]. It is very effective to analyze the optimization effect of the algorithms on the functions in the test suit, especially for the complex, diverse, multi-modal, funnel-shaped features. It visualizes the distribution of solutions. The red dots in Fig. 15-17 represent the individuals involved in the evolution. In the unimodal functions $f_1 - f_3$, the optimal position is found at express speed. $f_4 - f_{10}$ are the multimodel functions with the rugged landscape and many local optimal solutions. The global optimal position is difficult to be discovered. The composition functions with composition basins are of different shapes, so they are the most difficult to achieve the optimal solution.

APBIL owns the weak local search ability, and performs poorly on f_1 of non-seperable, smooth but narrow ridge. As can be seen from Fig. 15, the proposed KFHLEDA evolves rapidly on f_1, and find the optimal solution, attribute to the rich population diversity and excellent local search ability. CMA-ES performs poorly on f_6 of asymmetrical, and huge number of local optima, nevertheless, KFHLEDA finds the optimal solution quickly as shown in Fig. 16. For the composition functions, such as $f_{21}, f_{22}, f_{24}, f_{26}$ with many peaks and valleys in Fig. 17, KFHLEDA overcomes these obstacles better due to the good exploration and exploitation capability.

6. Conclusion and future work

An enhanced Kalman filtering and historical learning mechanism driven EDA (KFHLEDA) with the fitness landscape for continuous optimization problem is proposed. KFHLEDA enhances the conventional Kalman filtering operation. According to the specific problem characteristics in the optimization process, the prediction, observation, the first revision and the second revision operations are embedded into the EDA framework, and the filtering mechanism is based on the elite strategy to modify the system gain and improve the filtering effect. Theoretical analysis and experiments show that the modified probabilistic model with enhanced Kalman filtering takes the revised mean as a new search center, improves the direction of population evolution, and strengthens the quality of solutions, which overcomes the shortcomings of the traditional EDA. These operations improve the search efficiency effectively. In addition, the dominant information of the iterative process is put into the historical archive to guide the population evolution so as to avoid the premature convergence, instead of just utilizing the information of the current solution. Learning the information carried by dominant individuals also modifies the mean and covariance matrix, and reduces the invalid evolutionary iterations. Moreover, the population adaptive adjustment strategy lessons the population size to balance the exploration and exploitation further during the evolution process, and ameliorates the overall efficiency of the algorithm. The non-parametric tests display that there are significant differences between KFHLEDA and the comparison algorithms. Furthermore, the fitness landscape reemerges the evolution process and the distribution of the whole population completely, which has guiding significance for the design of the algorithm in this study. The search behavior is visualized to show the effectiveness of the designed algorithm. The experimental results show that KFHLEDA has superior robustness and stability on 30 functions in CEC 2017 test suit with 10D, 30D, 50D, and 100D.

Given the proposed algorithm achieving desired results, the following work will connect KFHLEDA with other evolutionary algorithms and learning mechanisms for solving the constrained multi-objective continuous optimization problems. Besides, it will also be applied to solve practical problems combined with hyper-heuristic algorithms, such as distributed no-wait flow shop scheduling problem with energy consumption constraints.

CRediT authorship contribution statement

Ningning Zhu: Investigation, Software, Original draft, Experiments of the algorithms. **Fuqing**

Zhao: Funding acquisition, Investigation, Supervision. **Ling Wang:** Methodology, Resources, Conceptualization. **Chenxin Dong:** Formal analysis, Review & editing, Visualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

This work was financially supported by the National Natural Science Foundation of China under grant 62063021. It was also supported by the Key talent project of Gansu Province (ZZ2021G50700016), the Key Research Programs of Science and Technology Commission Foundation of Gansu Province (21YF5WA086), and Lanzhou Science Bureau Project (2018-rc-98), respectively.

References

[1] F. Zhao, H. Bao, L. Wang, J. Cao, J. Tang, Jonrinaldi, A multipopulation cooperative coevolutionary whale optimization algorithm with a two-stage orthogonal learning mechanism, Knowledge-Based Syst. 246 (2022) 108664.

[2] J. Liang, X. Ban, K. Yu, K. Qiao, B. Qu, Constrained multiobjective differential evolution algorithm with infeasible-proportion control mechanism, Knowledge-Based Syst. 250 (2022) 109105.

[3] D. Bhati, P. Singh, Branch and bound computational method for multi-objective linear fractional optimization problem, Neural Comput. Appl. 28 (2017) 3341-3351.

[4] A.C. Luna, Student, IEEE, N.L. Diaz, Student, Mixed-Integer-Linear-Programming-Based Energy Management System for Hybrid PV-Wind-Battery Microgrids: Modeling, Design, and Experimental Verification, IEEE Trans. Power Electron. 32 (2017) 2769–2783.

[5] P.D. Wang, Y.J. Ma, M.H. Wang, A dynamic multi-objective optimization evolutionary algorithm based on particle swarm prediction strategy and prediction adjustment strategy, SWARM Evol. Comput. 75 (2022) 101164.

[6] F. Zhao, B. Zhu, L. Wang, T. Xu, N. Zhu, J. Jonrinaldi, An offline learning co-evolutionary algorithm with problem-specific knowledge, Swarm Evol. Comput. 75 (2022) 101148.

[7] F. Zhao, Z. Xu, L. Wang, N. Zhu, T. Xu, Jonrinaldi, A Population-Based Iterated Greedy Algorithm for Distributed Assembly No-Wait Flow-Shop Scheduling Problem, IEEE Trans. Ind. Informatics. (2022).

[8] F. Zhao, X. He, L. Wang, A Two-Stage Cooperative Evolutionary Algorithm With Problem-Specific Knowledge for Energy-Efficient Scheduling of No-Wait Flow-Shop Problem, IEEE Trans. Cybern. 51 (2020) 5291–5303.

[9] D.Y. Li, W.A. Guo, A. Lerch, Y.M. Li, L. Wang, Q.D. Wu, An adaptive particle swarm optimizer with decoupled exploration and exploitation for large scale optimization, SWARM Evol. Comput. 60 (2021) 1100789.

[10] H. Mühlenbein, J. Bendisch, H.M. Voigt, From recombination of genes to the estimation of distributions II. Continuous parameters, Springer, Berlin, Heidelb. (1996).

[11] Y. Yu, S.C. Gao, M.C. Zhou, Y.R. Wang, Z.Y. Lei, T.F. Zhang, J.H. Wang, Scale-free network-based differential evolution to solve function optimization and parameter estimation of photovoltaic models, SWARM Evol. Comput. 74 (2022) 101142.

[12] Z. Shao, D. Pi, W. Shao, P. Yuan, An efficient discrete invasive weed optimization for blocking

flow-shop scheduling problem, Eng. Appl. Artif. Intell. 78 (2019) 124–141.
[13] F. Zhao, X. Hu, L. Wang, J. Zhao, J. Tang, Jonrinaldi, A reinforcement learning brain storm optimization algorithm (BSO) with learning mechanism, Knowledge-Based Syst. 235 (2022) 107645.
[14] J. Zhang, L. Xing, An improved genetic algorithm for the integrated satellite imaging and data transmission scheduling problem, Comput. Oper. Res. 139 (2022) 105626.
[15] Q.L. Dang, W.F. Gao, M.G. Gong, An efficient mixture sampling model for gaussian estimation of distribution algorithm, Inf. Sci. (Ny). 608 (2022) 1157–1182.
[16] Y. Liang, Z. Ren, X. Yao, Z. Feng, A. Chen, W. Guo, Enhancing Gaussian Estimation of Distribution Algorithm by Exploiting Evolution Direction with Archive, IEEE Trans. Cybern. 50 (2020) 140–152.
[17] Z.G. Ren, Y.S. Liang, L. Wang, A.M. Zhang, B. Pang, B.Y. Li, Anisotropic adaptive variance scaling for Gaussian estimation of distribution algorithm, KNOWLEDGE-BASED Syst. 146 (2018) 142–151.
[18] Q. Yang, W.N. Chen, Y. Li, C.L.P. Chen, X.M. Xu, J. Zhang, Multimodal estimation of distribution algorithms, IEEE Trans. Cybern. 47 (2017) 636–650.
[19] A.M. Zhou, J.Y. Sun, Q.F. Zhang, An Estimation of Distribution Algorithm With Cheap and Expensive Local Search Methods, IEEE Trans. Evol. Comput. 19 (2015) 807–822.
[20] E. Cuevas, A. Rodriguez, A. Valdivia, D. Zaldivar, M. Perez, A hybrid evolutionary approach based on the invasive weed optimization and estimation distribution algorithms, SOFT Comput. 23 (2019) 13627–13668.
[21] B. Doerr, M.S. Krejca, The Univariate Marginal Distribution Algorithm Copes Well with Deception and Epistasis, Evol. Comput. 29 (2021) 543–563.
[22] A. Khajenezhad, M.A. Bashiri, H. Beigy, A distributed density estimation algorithm and its application to naive Bayes classification, Appl. Soft Comput. 98 (2021) 106837.
[23] M.S.R. Martins, M. El Yafrani, M. Delgado, R. Lüders, R. Santana, H. V. Siqueira, H.G. Akcay, B. Ahiod, Analysis of Bayesian Network Learning Techniques for a Hybrid Multi-objective Bayesian Estimation of Distribution Algorithm: a case study on MNK Landscape, J. Heuristics. 27 (2021) 549–573.
[24] C. Witt, Theory of estimation-of-distribution algorithms, GECCO 2020 Companion - Proc. 2020 Genet. Evol. Comput. Conf. Companion. (2020) 1254–1282.
[25] Q. Yang, Y. Li, X.D. Gao, Y.Y. Ma, Z.Y. Lu, S.W. Jeon, J. Zhang, An Adaptive Covariance Scaling Estimation of Distribution Algorithm, Mathematics. 9 (2021).
[26] M. Wagner, A. Auger, M. Schoenauer, E.A. New, R. Estimation, EEDA : A New Robust Estimation of Distribution Algorithms, (2006).
[27] N.B.T.G. and E.C.C. Hansen, A global surrogate assisted CMA-ES, GECCO 2019 Companion - Proc. 2019 Genet. Evol. Comput. Conf. Companion. (2019) 664–672.
[28] Y. Liang, Z. Ren, W. Lin, P. Bei, M.M.B.T.-E.C. Hossain, Inferior solutions in Gaussian EDA: Useless or useful?, IEEE Congress on Evolutionary Computationin. (2017) 301-307.
[29] Z.P. Liang, J.Y. Zeng, L. Liu, Z.X. Zhu, A many-objective optimization algorithm with mutation strategy based on variable classification and elite individual, SWARM Evol. Comput. 60 (2021) 100769.
[30] J. Zhang, S. Member, IEEE, Fellow, IEEE, JADE: Adaptive Differential Evolution With Optional External Archive, IEEE Trans. Evol. Comput. 13 (2009) 945–958.

[31] C. Huemmer, C. Hofmann, R. Maas, W. Kellermann, Estimating Parameters of Nonlinear Systems Using the Elitist Particle Filter Based on Evolutionary Strategies, IEEE/ACM Trans. Audio Speech Lang. Process. 26 (2018) 595–608.

[32] Q. Yang, J.-Q. Yan, X.-D. Gao, D.-D. Xu, Z.-Y. Lu, J. Zhang, Random neighbor elite guided differential evolution for global numerical optimization, Inf. Sci. (Ny). 607 (2022) 1408–1438.

[33] F. Zhao, J. Zhao, L. Wang, J. Cao, J. Tang, A hierarchical knowledge guided backtracking search algorithm with self-learning strategy, Eng. Appl. Artif. Intell. 102 (2021) 104268.

[34] Y. Li, X. Yu, J. Liu, An opposition-based butterfly optimization algorithm with adaptive elite mutation in solving complex high-dimensional optimization problems, Math. Comput. Simul. 24 (2022) 498–528.

[35] W.C. Yi, Y. Chen, Z. Pei, J.S. Lu, Adaptive differential evolution with ensembling operators for continuous optimization problems, SWARM Evol. Comput. 69 (2022) 100994.

[36] A. Auger, N.B.T.-I. Hansen, A restart CMA evolution strategy with increasing population size, IEEE Congress on Evolutionary Computation. (2005) 1769–1776.

[37] R. Tanabe, A.S.B.T.-E.C. Fukunaga, Improving the search performance of SHADE using linear population size reduction, IEEE Congress on Evolutionary Computation. (2014) 1658-1665.

[38] W. Shi, W.N. Chen, T.L. Gu, H. Jin, J. Zhang, Handling Uncertainty in Financial Decision Making: A Clustering Estimation of Distribution Algorithm With Simplified Simulation, IEEE Trans. Emerg. Top. Comput. Intell. 5 (2021) 42–56.

[39] H. Zhao, Z.-H. Zhan, J. Liu, Outlier Aware Differential Evolution for Multimodal Optimization Problems, SSRN Electron. J. (2022) 1–12.

[40] Z. Lin, Q. Su, G. Xie, NMIEDA: Estimation of distribution algorithm based on normalized mutual information, Concurr. Comput. Pract. Exp. 33 (2021) 1–16.

[41] P. Yang, K. Tang, X.F. Lu, Improving Estimation of Distribution Algorithm on Multimodal Problems by Detecting Promising Areas, IEEE Trans. Cybern. 45 (2015) 1438–1449.

[42] J.Y. Li, Z.H. Zhan, J. Xu, S. Kwong, J. Zhang, Surrogate-Assisted Hybrid-Model Estimation of Distribution Algorithm for Mixed-Variable Hyperparameters Optimization in Convolutional Neural Networks, IEEE Trans. NEURAL NETWORKS Learn. Syst. (2021) 3106399.

[43] Z.W. Liao, W.Y. Gong, Z.H. Cai, L. Wang, Y. Wang, IEEE, Random Walk Mutation-based DE with EDA for Nonlinear Equations Systems, 2019 IEEE Congr. Evol. Comput. (2019) 3118–3125.

[44] J. Li, J.Q. Zhang, C.J. Jiang, M.C. Zhou, Composite Particle Swarm Optimizer With Historical Memory for Function Optimization, IEEE Trans. Cybern. 45 (2015) 2350–2363.

[45] W.S. Shao, D.C. Pi, Z.S. Shao, A Pareto-Based Estimation of Distribution Algorithm for Solving Multiobjective Distributed No-Wait Flow-Shop Scheduling Problem With Sequence-Dependent Setup Time, IEEE Trans. Autom. Sci. Eng. 16 (2019) 1344–1360.

[46] X.H. Zhang, X.H. Liu, A. Cichon, G. Krolczyk, Z.X. Li, Scheduling of energy-efficient distributed blocking flowshop using pareto-based estimation of distribution algorithm, Expert Syst. Appl. 200 (2022) 116910.

[47] Z.Q. Zhang, B. Qian, R. Hu, H.P. Jin, L. Wang, A matrix-cube-based estimation of distribution algorithm for the distributed assembly permutation flow-shop scheduling problem, SWARM Evol. Comput. 60 (2021) 100785.

[48] Y. Du, J.Q. Li, X.L. Chen, P.Y. Duan, Q.K. Pan, Knowledge-Based Reinforcement Learning and Estimation of Distribution Algorithm for Flexible Job Shop Scheduling Problem, IEEE Trans. Emerg. Top. Comput. Intell. (2022) 3145706.

[49] B. Chen, L. Dang, Y. Gu, N. Zheng, J.C. Principe, Minimum Error Entropy Kalman Filter, IEEE Trans. Syst. Man, Cybern. Syst. 51 (2021) 5819–5829.

[50] N. Saleem, J. Gao, M.I. Khattak, H.T. Rauf, S. Kadry, M. Shafi, DeepResGRU: Residual gated recurrent neural network-augmented Kalman filtering for speech enhancement and recognition, Knowledge-Based Syst. 238 (2022) 107914.

[51] H. Zhao, Areas(Volumes) of n-dimension Ellipsoid by Quadratic Curve(Surface) Enclosed, Math. Pract. Theory. 43 (2013) 279–282.

[52] Q.X. Yin, The Inverse Problem of Rank -1 Modification of Real Symmetric Matrices, J. Nantong Inst. Technol. Sci. (2003).

[53] N.H. Awad, M.Z. Ali, P.N. Suganthan, J.J. Liang, B.Y. Qu, Problem Definitions and Evaluation Criteria for the CEC2017 Special Session and Competition on Single Objective Real-Parameter Numertical Optimization, 2016.

[54] K.A.B.T.-2019 I.S.S. on C.I. (SSCI) Folly, A Short Survey on Population-Based Incremental Learning Algorithm, IEEE SYMPOSIUM SERIES ON COMPUTATIONAL INTELLIGENCE. (2019) 339-344.

[55] J. Brest, M.S. Maucec, B.B.T.-2017 I.C. on E.C. (CEC) Boskovic, Single objective real-parameter optimization: Algorithm jSO, in: 2017.

[56] R. Jiang, R. Shankaran, S. Wang, T. Chao, A proportional, integral and derivative differential evolution algorithm for global optimization, Expert Syst. Appl. 206 (2022) 117669.

[57] D.H. Wolpert, W.G. Macready, No free lunch theorems for optimization, IEEE. (1997).

[58] M.H. Tayarani-N, A. Prugel-Bennett, Anatomy of the fitness landscape for dense graph-colouring problem, SWARM Evol. Comput. 22 (2015) 47–65.

Declaration of interests

☒ The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

☐The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

Author Agreement

All authors of the paper have seen and approved the final version of the manuscript being submitted. We warrant that the article is the authors' original work, hasn't received prior publication and isn't under consideration for publication elsewhere.