Cooperative Framework Fireworks Algorithm with Covariance Mutation

Chao Yu¹, Ling Chen Kelley²,³, and Ying Tan¹,⁎

¹ The Key Laboratory of Machine Perception and Intelligence (Ministry of Education), Department of Machine Intelligence, School of Electronics Engineering and Computer Science, Peking University, Beijing, China, 100871
² Department of History, School of Humanities, Tsinghua University, Beijing, China, 100084
³ Department of World Languages and Cultures, College of Liberal Arts and Sciences, Auburn University at Montgomery, Montgomery, AL, United States 36117

Email: chaoyu@pku.edu.cn, lingchenkelley@gmail.com, ytan@pku.edu.cn

Abstract—Since fireworks algorithm (FW A) debuted in 2010, a dozen proposals of improvement for FW A had been published in an effort to enhance, refine and optimize accuracy while minimizing calculation speed and volume. In this paper, we introduce the covariance mutation operator into cooperative framework fireworks algorithm (CoFFWA) to create cooperative framework fireworks algorithm with covariance mutation (CoFFWA-CM) in order to solve the Congress on Evolutionary Computation (CEC 2016) competition functions on single objective optimization. The experimental results are calculated on all 10-, 30-, 50- and 100-dimensional functions.

I. INTRODUCTION

Since its initial introduction by Tan and Zhu in 2010, FW A has witnessed several dozen designs that sought to improve, refine and optimize its output [1], [2]. FW A can be enhanced in three ways, as improvements in the explosion operator, the mutation operator and or the selection strategies.

The explosion operator was improved in 2013 by Zheng et al. that addressed to the weaknesses of conventional FW A and proposed the enhance FW A (EFWA) [3]. Liu et al. constructed a transfer function to generate sparks and proposed a constructed FW A (CFWA) [4]. In 2014, Zheng et al. improved the explosion amplitude of each firework and put forward dynamic search FW A (dynFW A) [5]. In 2015, Zheng et al. periodically reduced the number of explosion dimensions and put forward an exponentially decreased dimension FW A (eddynFW A) [6].

Enhancements of mutation operator include Yu et al., whom introduced the differential mutation operator into the EFW A and perpetuated a fireworks algorithm with differential mutation (FWA-DM) [7], [8], Li et al. calculated the difference values between the better and worse sparks and proposed orienting mutation based FW A (FWA-OM) [9]. The most effective improvement was the covariance mutation, which was proposed by Yu and Tan in 2015, called FW A with covariance mutation (FWACM) [10], [11].

The selection strategy was upgraded in 2012 by Pei et al., which utilized the approximation approaches to produce new sparks. If the fitness value of the new spark was better than the current selected spark, the new spark would be selected and passed down to the next generation. FW A was accelerated by this approximation method and the algorithm was named as accelerate FW A (AcFW A) [12]. In 2015, Zheng et al. invented an independent selection mechanism in the selection strategy. In addition, a crowd and bound back mechanism was introduced. When the selected sparks were too close, all but one would be generated randomly. The algorithm enhanced the cooperation between the sparks and called as the cooperative framework FW A (CoFFWA) [13].

Aside from the algorithmic improvements, FW A had been applied to many practical fields. Bacanin and Tuba applied FW A to solve the constrained portfolio optimization problems [14]. Bouarara et al. utilized FW A to solve the modern web information retrieval with visual results mining [15]. Rahmani et al. studied the privacy preservation problems in big data by FW A [16]. There were many other applications, as evidenced in article [17]–[19].

This study used CoFFWA with covariance mutation (CoFFWA-CM) to participate in the competition, as explained in section II [20]. By introducing the covariance mutation into the process of CoFFWA, the performance of the CoFFWA was further enhanced. The new mutation operator in CoFFWA-CM differed from the covariance mutation in FWACM while the new mutation operator is analyzed in section III. The experimental results are displayed in section IV, followed by discussion and conclusion.

II. COOPERATIVE FRAMEWORK FIREWORKS ALGORITHM WITH COVARIANCE MUTATION

CoFFWA-CM was proposed based on the previous work of CoFFWA and covariance mutation. Published in Oct, 2015, CoFFWA enhanced the diversity of the population by incorporating the independent selection strategy and bound back mechanism, thus outperforming both EFWA and dynFWA. In the same year, FWACM was put forward and defeated both EFWA and dynFWA. However, FWACM was based on the work of dynFWA, whereas dynFWA was not as good as CoFFWA. Therefore, covariance mutation was introduced to the most advanced FW A, which was CoFFWA, to proof its effectiveness.

CoFFWA-CM generated two kinds of sparks: explosion sparks from the explosion operator and mutation sparks from the covariance mutation. Similar to the FW A variants, the
sparks were then systematically evaluated and selected for the next generation. This iteration continued until the termination conditions were met, usually when the maximum number of function evaluations and or the accuracy requirements had been reached. CoFFWA-CM proposed to further improve accuracy with fewer generations. The covariance mutation was a mutation with Gaussian distribution and utilized the better sparks in each generation, instead of using the single best spark. The aim of covariance mutation was to use more effective information and thus increase the information utilization ratio.

CoFFWA-CM consisted of explosion operator, mutation operator, mapping rules and selection strategy. The details of them were as follows.

### A. Explosion Operator

The explosion operator emulated a firework explosion, where sparks were generated encompassing a central firework. Before the explosion occurred, the number of sparks and the amplitude of the explosion were determined, and were generated in the same fashion as dynFWA, FW ACM and CoFFWA. The explosion amplitude and the number of explosion of each firework all factored into the optimization results.

![Real fireworks explosion is illustrated in (a), and the explosion operation in FWA is shown in (b).](image)

Before a firework exploded, the number of the sparks and the amplitude of the explosion were calculated in advance. In CoFFWA-CM, the number of sparks was inherited from the conventional FWA.

Let $S_i$ denotes the number of sparks for the $i^{th}$ firework.

$$S_i = \hat{S} \ast \frac{f(X_w) - f(X_i) + \varepsilon}{\sum_{(i=1)}^{N} (f(X_w) - f(X_i)) + \varepsilon}$$

where the parameter $\hat{S}$ determines the sum of sparks in each generation. The Function $f(x)$ represents the fitness value of the input $x$. Since $X_w$ is the individual with the worst fitness value, function $f(X_w)$ gives the fitness value of that individual $X_w$. The parameter $N$ is the number of fireworks in a generation, whereas $\varepsilon$ is used to prevent the denominator from becoming zero. There are lower and upper boundaries for $S_i$, which are set empirically.

The sign $A_{CF}$ stands for the explosion amplitude of the core firework. When saying the ‘core firework’, it refers to the firework with the best fitness value in the current iteration. The calculation of the $A_{CF}$ is as in Eq. 2.

$$A_{CF}(g) = \begin{cases} A_{CF}(1), & \text{if } g = 1 \\ A_{CF}(g - 1) \ast C_a, & \text{if } f(X_{CF}(g)) < f(X_{CF}(g - 1)) \\ A_{CF}(g - 1) \ast C_r, & \text{otherwise} \end{cases}$$

where the parameter $g$ is the number of generations, $C_a$ and $C_r$ stands for the amplification and reduction factors of the explosion amplitude, and $X_{CF}(g)$ is the core firework in the $g^{th}$ generation.

Different from the core firework, the other fireworks have the traditional way to calculate the explosion amplitude, which is inherited from the conventional FWA. The amplitude of explosion of the $i^{th}$ firework is denoted as $A_i$.

$$A_i = \hat{A} \ast \frac{f(X_i) - f(X_{CF}) + \varepsilon}{\sum_{i=1}^{N} (f(X_i) - f(X_{CF})) + \varepsilon}$$

where the parameter $\hat{A}$ is a constant used to control the amplitude of explosions, $X_i$ is the $i^{th}$ individual. The parameter $N$ and $\varepsilon$ are the same as mentioned in Eq. 1.

Under the effect of the explosion operator, a firework produces a certain number of sparks within a preset amplitude, both of which are vital to the experimental results. However, compared to the number of sparks, the experimental results are much worse if the explosion amplitudes are set improperly. For example, if the number of sparks enlarges 10 times, the experimental results are The explosion amplitudes are always worth for studying.

### B. Mutation Operator

In CoFFWA-CM, the covariance mutation utilizes information from the sparks generated by the core firework, instead of solely focusing on the single best spark. Unlike previous mutation operators, covariance mutation uses both the sparks produced by the fireworks with best fitness values from the current generation and from the single most recent generation to calculate an optimal solution. By using sparks from the previous generation, the covariant mutation selects better sparks with proper fitness values than previous algorithms.

The covariance mutation selects the sparks with better fitness values from the sparks produced by a firework, calculates the mean value of the selected sparks and the covariance matrix of all the sparks. With the mean value and covariance matrix, covariance mutation estimates the local distribution of a function and produces sparks according to normal distribution, aiming to find potential sparks with better fitness values. The covariance mutation contains three steps.

Firstly, the covariance mutation selects the sparks produced by a firework and calculates the mean value. Let the sign $\lambda$ represents the sum of sparks and $\mu$ stands for the number of the selected sparks. The mean value of the selected sparks is represented as $m$.

$$m = \frac{\mu}{\lambda} \ast \sum_{i=1}^{\lambda} x_i$$

Fig. 1. Real fireworks explosion is illustrated in (a), and the explosion operation in FWA is shown in (b).
where $x_i$ is the $i^{th}$ selected spark. Note that $m$ is the mean value of the $\mu$ individuals, rather than the mean value of all individuals.

Secondly, the mutation operator calculates the covariance matrix $C$ of all the $\lambda$ sparks. The $i^{th}$ row and $j^{th}$ column of the matrix $C$ is represented as $C_{ij}$ in Eq. 5.

$$C_{ij} = \text{cov}(d_i, d_j)(i, j = 1, ..., D),$$  \hspace{1cm} (5)

where constant $D$ is the dimension of the benchmark function and sample $d_i$ is the sparks in their $i^{th}$ dimension. The $\text{cov}(d_i, d_j)$ stands for the covariance of the sparks in $i^{th}$ and $j^{th}$ dimensions and it is calculated in Eq. 6.

$$\text{cov}(d_i, d_j) = \frac{\sum_{k=1}^{n} (a_k - \bar{A})(b_k - \bar{B})}{\mu}.$$  \hspace{1cm} (6)

where $a_k$ and $b_k$ are the $k^{th}$ spark in its $i^{th}$ and $j^{th}$ dimensions, $\bar{A}$ and $\bar{B}$ are the mean value of all the $\lambda$ sparks in dimension $i$ and $j$. Different from calculating the covariance, the denominator here is not $\mu - 1$ as usual.

Thirdly, the mutation sparks are generated according to normal distribution, using the $m$ and $C$ obtained before. Figure 2 shows the contour of the mutation sparks in ellipse.

It can be seen from Fig. 2 that the possibility of finding useful sparks increased, as the mutation sparks produced by covariance mutation can now better navigate towards the optimal direction.

The algorithm of the covariance mutation is given in Alg. 1. The two groups of sparks are produced by the current and the most recent generation best sparks. The sign $N(m, C)$ denotes a normal distribution with mean $m$ and covariance $C$.

**Algorithm 1 Covariance Mutation**

1. collect two groups of sparks
2. calculate the mean value $m$ of the better sparks
3. obtain the covariance matrix $C$
4. generate mutation sparks with $N(m, C)$

**C. Mapping Rules**

Mapping rules are proposed to handle sparks that fell outside of the boundaries. When a firework is close to the boundary, the generated sparks can easily stray outside of the boundary when distributed by a relatively large explosion amplitude. In CoFFW A-CM, the mapping rules are the same as in conventional FWA, where the outlying sparks are generated randomly back into the feasible space.

$$X = X_i + \text{rand}(0, 1) \ast (X_u - X_l),$$  \hspace{1cm} (7)

where $X$ is the location of the spark $X$ in the feasible space, $X_u$ and $X_l$ denote the upper and lower boundaries. Function $\text{rand}(0, 1)$ generates a random number within the area from 0 to 1 with uniform distribution.

**D. Selection Strategies**

The two strategies used in CoFFW A-CM are independent selection strategy and the crowd bounce off strategy. The independent selection strategy selects $N$ individuals independently of the next generation, where $N$ fireworks produce $N$ groups of fireworks, followed by the best spark in each group being selected. The crowd and bounce off strategy filters the selected sparks. If two selected sparks are too close, then one of them is bounced off randomly to another location. The effectiveness of these two strategies are carefully researched in the reference [13].

To algorithm of CoFFW A-CM is shown in Alg. 2.

**Algorithm 2 The algorithm of CoFFW A-CM**

1. generate $N$ fireworks with uniform distribution randomly.
2. evaluate the fitness values of the $N$ fireworks
3. while terminate condition is not met do
4. calculate the number of sparks $S_i$
5. calculate the amplitude of explosion $A_{GF}$ and $A_i$
6. generate $S_i$ explosion sparks within the amplitudes
7. calculate $m$ and covariance matrix $C$
8. generate Gaussian sparks according with $N(m, C)$
9. evaluate all the fitness values of explosion and Gaussian sparks
10. keep $N$ individuals for the next generation using selection strategies
11. end while
12. return the best individual and its fitness value

**III. ANALYSIS OF THE MUTATION IN COFFWA-CM**

Covariance mutation is an implementation of CoFFWA. In CoFFWA, there is no mutation operator, as such, it is easy to add the mutation operator to the CoFFWA without changing the original structure. The new algorithm firstly utilized the explosion operator, followed by the mutation operator and the selection strategies. If any sparks are produced out of the boundaries, they will be discarded and a new spark will be generated randomly within the feasible space.

A two-dimensional function $f(x) = x_1 + x_2$ is taken as an example to visualize the covariance mutation in CoFFWA.

In Fig. 3, the processes of generating mutation sparks by covariance mutation is shown one by one. The contour lines
is drawn in Fig. 3(a), where the bottom left indicates better fitness values. Figure 3(b) illustrates the selected \( \lambda \) sparks. The better sparks are shown in Fig. 3(c) marked as black ‘x’. The mutation sparks are represented in Fig. 3(d) marked with red ‘+’.

Fig. 3. The process of generating mutation sparks by covariance mutation.

It can be seen from Fig. 3 that the sparks marked with red ‘+’ distributed in a direction from the bottom left to the upper right, which is close to the direction of the gradient. Besides, the relatively better sparks in Fig. 3d are all produced by covariance mutation, which are marked with red ‘+’. Therefore, the covariance mutation operator is effective at finding local optimal values with the information of the sparks. The characteristic of covariance mutation can be extended to higher dimensions and more complicated functions.

To validate the performance of covariance mutation, the rotated high conditioned elliptic function was utilized. This function was the number 1 function in CEC 2016 single objective optimization function, which was a unimodal function. The comparison experiments show the CoFFWA with and without covariance mutation. All of the four dimensions were considered, as 10, 30 and 50 dimensions. Each algorithm ran for 51 times and the average mean errors were recorded.

<table>
<thead>
<tr>
<th>TABLE I. THE COMPARISON OF COFFWA WITH COFFWA-CM ON FUNCTION 1 IN 10, 30 AND 50 DIMENSIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dimensions</td>
</tr>
<tr>
<td>10</td>
</tr>
<tr>
<td>30</td>
</tr>
<tr>
<td>50</td>
</tr>
</tbody>
</table>

The experimental results in Table I indicated that when the dimension increased to 50, CoFFWA-CM still had the better performance than the original CoFFWA. Note that the parameters chosen for CoFFWA were its default parameters, whereas the parameters of CoFFWA-CM were the same as indicated in the experiments section. The better performance of CoFFWA-CM proof its effectiveness on 10, 30 and 50 dimensional of a unimodal function.

IV. EXPERIMENTS

This part first introduce the benchmark functions, followed by the parameter settings and experimental results. Note that a value of 100 was added to function 1, 200 to function 2, ..., and a value of 3000 to function 30, all of which were not subtracted. In this way, the experimental results were displayed as they were calculated from their corresponding benchmark functions.

A. Benchmark Functions

CoFFWA-CM was used to find the global optimum values of 30 benchmark functions from the CEC’16 competition. The details of the functions could be found in [21]. The names and numbers of the functions were given in Table II. The functions included unimodal, multi-modal, hybrid and composition functions, which were divided by horizontal lines in Table II.

<table>
<thead>
<tr>
<th>TABLE II. BENCHMARK FUNCTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>No.</td>
</tr>
<tr>
<td>1.</td>
</tr>
<tr>
<td>2.</td>
</tr>
<tr>
<td>3.</td>
</tr>
<tr>
<td>4.</td>
</tr>
<tr>
<td>5.</td>
</tr>
<tr>
<td>6.</td>
</tr>
<tr>
<td>7.</td>
</tr>
<tr>
<td>8.</td>
</tr>
<tr>
<td>9.</td>
</tr>
<tr>
<td>10.</td>
</tr>
<tr>
<td>11.</td>
</tr>
<tr>
<td>12.</td>
</tr>
<tr>
<td>13.</td>
</tr>
<tr>
<td>14.</td>
</tr>
<tr>
<td>15.</td>
</tr>
<tr>
<td>16.</td>
</tr>
<tr>
<td>17.</td>
</tr>
<tr>
<td>18.</td>
</tr>
<tr>
<td>19.</td>
</tr>
<tr>
<td>20.</td>
</tr>
<tr>
<td>21.</td>
</tr>
<tr>
<td>22.</td>
</tr>
<tr>
<td>23.</td>
</tr>
<tr>
<td>24.</td>
</tr>
<tr>
<td>25.</td>
</tr>
<tr>
<td>26.</td>
</tr>
<tr>
<td>27.</td>
</tr>
<tr>
<td>28.</td>
</tr>
<tr>
<td>29.</td>
</tr>
<tr>
<td>30.</td>
</tr>
</tbody>
</table>

B. Parameters Settings

The parameters were set identically for all the dimensions. Some of the parameters’ settings were shown in Table III. It was clearly that the parameters were simple and easy to set. The other parameters are demonstrated in [13].

<table>
<thead>
<tr>
<th>TABLE III. PART OF THE PARAMETERS’ SETTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parameters</td>
</tr>
<tr>
<td>( N )</td>
</tr>
<tr>
<td>( S )</td>
</tr>
<tr>
<td>( A )</td>
</tr>
<tr>
<td>( \mu/\lambda )</td>
</tr>
</tbody>
</table>

C. Experimental Results

The experimental platform is Matlab 2015a and the program is run on a Windows 10 operating system. The experimental results on 10, 30, 50 and 100-dimension functions of CEC 2016 single objective optimization competition are given in Table IV, V, VI and VII, respectively.

The computational complexity of CoFFWA-CM is given in Table VIII.
V. Discussion

The covariance mutation is effective in estimating local distribution in all functions. In CoFFWA-CM, each group of sparks is produced around a firework, thus the local distribution is used. There are other ways to estimate the local distribution of the functions, such as sampling and covariance mutation is used. There are other ways to estimate the local distribution of the functions, but now the geologically closest and the fitness values of these sparks. Early in 2012, Pei et al. proposed an approximate method to estimate the local distribution in all functions. In CoFFWA-CM, each group of sparks is produced around a firework, thus the local distribution is used. There are other ways to estimate the local distribution of the functions, such as sampling and covariance mutation is used. There are other ways to estimate the local distribution of the functions, but now the geologically closest and the fitness values of these sparks. Early in 2012, Pei et al. proposed an approximate method to estimate the local distribution in all functions. In CoFFWA-CM, each group of sparks is produced around a firework, thus the local distribution is used. There are other ways to estimate the local distribution of the functions, such as sampling and covariance mutation is used. There are other ways to estimate the local distribution of the functions, but now the geologically closest and the fitness values of these sparks. Early in 2012, Pei et al. proposed an approximate method to estimate the local distribution in all functions. In CoFFWA-CM, each group of sparks is produced around a firework, thus the local distribution is used. There are other ways to estimate the local distribution of the functions, such as sampling and covariance mutation is used. There are other ways to estimate the local distribution of the functions, but now the geologically closest and the fitness values of these sparks. Early in 2012, Pei et al. proposed an approximate method to estimate the local distribution in all functions. In CoFFWA-CM, each group of sparks is produced around a firework, thus the local distribution is used. There are other ways to estimate the local distribution of the functions, such as sampling and covariance mutation is used. There are other ways to estimate the local distribution of the functions, but now the geologically closest and the fitness values of these sparks. Early in 2012, Pei et al. proposed an approximate method to estimate the local distribution in all functions. In CoFFWA-CM, each group of sparks is produced around a firework, thus the local distribution is used. There are other ways to estimate the local distribution of the functions, such as sampling and covariance mutation is used. There are other ways to estimate the local distribution of the functions, but now the geologically closest and the fitness values of these sparks. Early in 2012, Pei et al. proposed an approximate method to estimate the local distribution in all functions. In CoFFWA-CM, each group of sparks is produced around a firework, thus the local distribution is used. There are other ways to estimate the local distribution of the functions, such as sampling and covariance mutation is used. There are other ways to estimate the local distribution of the functions, but now the geologically closest and the fitness values of these sparks. Early in 2012, Pei et al. proposed an approximate method to estimate the local distribution in all functions. In CoFFWA-CM, each group of sparks is produced around a firework, thus the local distribution is used. There are other ways to estimate the local distribution of the functions, such as sampling and
based on the probabilistic distribution model, all of which are worth researching in the future.

As there is no comparison algorithm, the performance can be analyzed if the global best value has been found. From the experimental results of 10-dimensional functions, the optimal values were found on function 7, 12, 13, 14, 15, 16 and 19. However, when the values were subtracted and the optimal values became zero, the algorithm found no value that equals zero. The experimental results on the 3rd function in its 30-, 50- and 100-dimension were great, as the algorithm found all the values as zero. Therefore, even the global optimal values were not found on most functions, it didn’t necessary mean the algorithm performed poorly. The only reasonable explanation was the functions that were complicated and the global optimal values of the functions were hard to pinpoint.

Another way to measure the experimental results was to judge the standard deviation. From the experimental results of 10-dimensional functions, 9 of the 30 standard deviation were lower than 1. This meant the results were steady with a relatively small variation after 51 cycles. The experimental results on the 30-dimensional functions were different. The algorithm found the global optimal value on function 3 and the local optimal value on function 5, both of the standard deviations were zero. It can be seen from to the experimental results on 50 dimensional functions that the algorithm found the global best value on function 3, whereas the algorithm found the local optimal on function 5 in 100-dimension functions. Both of the standard deviations were zero.

Although the overall experimental results were acceptable, there is still much space for improvement of the algorithms to ascertain better results. CoFFWA-CM uses sparks with better fitness values from previous generations with the current generation. However, it would be an interesting study to weigh how many previous generations could be used to refine the most recent spark before the refinement process becomes moot. Nevertheless, we expect to see progressive research to improve CoFFWA-CM in the near future.

VI. CONCLUSION

This paper utilized the most advanced FWA variant to participate the CEC 2016 single objective optimization problems. The CoFFWA-CM utilized the covariance mutation and increased the local search ability. By successfully integrating the covariant mutation into CoFFWA, CoFFWA-CM further enhanced the accuracy of CoFFWA. CoFFWA-CM outperformed its predecessors on all accounts and has achieved better accuracy while minimizing computation time.

ACKNOWLEDGMENT

This work was supported by the Natural Science Foundation of China (NSFC) under grant no. 61375119 and the Beijing Natural Science Foundation under grant no. 4162029, and partially supported by National Key Basic Research Development Plan (973 Plan) Project of China under grant no. 2015CB352302.

REFERENCES

Fig. 4. Convergence curves for CoFFWA-CM on function 1 to 15 in Dimension 10, 30, 50 and 100.
Fig. 5. Convergence curves for CoFFWA-CM on function 16 to 20 in Dimension 10, 30, 50 and 100.