Frontiers of Information Technology & Electronic Engineering
www.jzus.zju.edu.cn; engineering.cae.cn; www.springerlink.com
ISSN 2095-9184 (print); ISSN 2095-9230 (online)
E-mail: jzus@zju.edu.cn

Unusual phenomenon of optimizing the Griewank function with the increase of dimension*

Yan HUANG[1,2,3], Jian-ping LI‡[2], Peng WANG‡[4]

[1]*School of Computer Science and Technology, Huaiyin Normal University, Huai'an 223000, China*

[2]*School of Computer Science and Engineering, University of Electronic Science and Technology of China, Chengdu 611731, China*

[3]*Jiangsu Key Laboratory of Media Design and Software Technology, Jiangnan University, Wuxi 214122, China*

[4]*School of Computer Science and Technology, Southwest Minzu University, Chengdu 610225, China*

E-mail: hep128@qq.com; jpli2222@uestc.edu.cn; qhoalab@163.com

Received Mar. 19, 2019; Revision accepted Aug. 21, 2019; Crosschecked Oct. 10, 2019

Abstract: The Griewank function is a typical multimodal benchmark function, composed of a quadratic convex function and an oscillatory nonconvex function. The comparative importance of Griewank's two major parts alters in different dimensions. Different from most test functions, an unusual phenomenon appears when optimizing the Griewank function. The Griewank function first becomes more difficult and then becomes easier to optimize with the increase of dimension. In this study, from the methodology perspective, this phenomenon is explained by structural, mathematical, and quantum analyses. Furthermore, frequency transformation and amplitude transformation are implemented on the Griewank function to make a generalization. The multi-scale quantum harmonic oscillator algorithm (MQHOA) with quantum tunnel effect is used to verify its characteristics. Experimental results indicate that the Griewank function's two-scale structure is the main reason for this phenomenon. The quantum tunneling mechanism mentioned in this paper is an effective method which can be generalized to analyze the generation and variation of solutions for numerous swarm optimization algorithms.

Key words: Griewank; Two-scale structure; Multi-scale quantum harmonic oscillator algorithm; Quantum tunnel effect

https://doi.org/10.1631/FITEE.1900155 **CLC number:** TP3-05

1 Introduction

The Griewank function was first used as a test function to verify the generalized descent method for global unconstrained minimization by Andreas GRIEWANK from Humboldt University, Germany, in 1981 (Griewank, 1981). It is now a typical multimodal test function in Congress on Evolutionary Computation (CEC) (Liang et al., 2013), and has been widely used to test state-of-the-art global optimization algorithms and their variants, such as particle swarm optimization (PSO) (Shi and Eberhart, 1999), simulated annealing (SA) (Wang and Chen, 1996), differential evolution (DE) (Qin et al., 2009), ant colony optimization (ACO) (Karaboga and Basturk, 2007), hunting search (Oftadeh et al., 2010), brain storm optimization (BSO) (Zhou et al., 2012), the fireworks algorithm (FWA) (Tan and Zhu, 2010), and the polar bear optimization (PBO) algorithm (Połap and Woźniak, 2017).

‡ Corresponding authors

* Project supported by the Natural Science Foundation of Huai'an, China (No. HAB201828), the Fundamental Research Funds for the Central Universities of China (No. 2019NYB22), and the Open Foundation of Jiangsu Key Laboratory of Media Design and Software Technology, China (Nos. 19ST0204 and 18ST0203)

ORCID: Yan HUANG, http://orcid.org/0000-0003-3896-5636

The global minimum of the Griewank function is unique, located in the origin. The Griewank function has a large number of local minima, exponentially increasing with the dimension (Cho et al., 2008). The complex structure of the Griewank function makes it difficult to analytically solve the derivative of the Griewank function to find the position of minima (Cho et al., 2008).

It is generally believed that with the increase of dimension, the global minimum becomes more and more difficult to detect. However, this is not the case. An unusual phenomenon can be found by optimizing the Griewank function with L-BFGS (Locatelli, 2003). With the increase of dimension, the expected number of local searches to first hit the global minimum decreases dramatically. This phenomenon is explained by mathematical analysis on the gradient with the increase of dimension (Locatelli, 2003).

As the Griewank function has been widely used as a test function by swarm optimization algorithms, a similar phenomenon can be found by analyzing the experimental data in this research area. In Zhang et al. (2003), the Griewank function was used to test the PSO algorithm with a random inertia weight, opposite to other test functions, in which the mean fitness value decreases with the increase of dimension for every population size. Karaboga and Basturk (2007) proposed the artificial bee colony algorithm (ABC). The Griewank function was used to compare the performance of ACO, PSO, and the particle swarm inspired evolutionary algorithm (PS-EA). Mean and standard deviations of the function values obtained by these algorithms showed a similar phenomenon that it is easier to optimize a low-dimensional Griewank function than a high-dimensional Griewank function. In Akbari et al. (2010), the Griewank function was used to compare the performance of the bee swarm optimization algorithm with that of ABC and the bee and foraging algorithm (BFA). When the dimension increases from 10 to 30, the average fitness decreases from 10^{-4} to 10^{-10}. The same phenomenon can be found in Chen and Zhao (2009), Akay and Karaboga (2012), Gao et al. (2012), and Rao et al. (2012). In these works, the dimensions of the Griewank function were all higher than 10 and the phenomenon of the Griewank function whose dimension lower than 10 was neglected.

Can we draw the conclusion that the higher the Griewank function's dimension is, the easier it can be optimized by swarm optimization algorithms? Can the high success rate (SR) of optimizing a high-dimensional Griewank function be used to verify the performance of optimization algorithms?

In this study, we optimize the Griewank function by swarm intelligence algorithms and find an unusual phenomenon that the Griewank function first becomes more difficult and then becomes easier to optimize with the increase of dimension (abbreviated as "the phenomenon").

We conduct a deep analysis on the Griewank function's mathematical structure, and find out that the Griewank function is composed of two major parts. One is a quadratic convex function with a unique global minimum, and the other is an oscillatory nonconvex function with numerous local minima. These two parts represent the two scales of the Griewank function. With the increase of dimension, the comparative importance of these two parts alters. For a low-dimensional Griewank function, the oscillatory nonconvex function has a greater impact on the optimization results, while for a high-dimensional Griewank function, the quadratic convex function has a dominant influence on the optimization results. The two-scale structure is the core difficulty in the optimization process. Structure, mathematical, and quantum analyses are made to reveal the reason behind this phenomenon.

In addition, the generalization of the Griewank function is proposed to make an in-depth analysis on this phenomenon. The frequency transformed parameter is used to adjust the distribution density of the local optimal solution. The amplitude transformed parameter is used to adjust the amplitude of the quadratic convex function. The multi-scale quantum harmonic oscillator optimization algorithm (MQHOA) with quantum tunnel effect is used to optimize the generalized Griewank function.

The contributions of this study are listed below:

1. The phenomenon of optimizing the Griewank function by swarm intelligence algorithms is found, which is different from Locatelli (2003) and was not mentioned in the literature on swarm intelligence algorithms.

2. Three analytical methods are proposed to interpret this phenomenon, including structure analysis, mathematical analysis, and quantum analysis. It is worth mentioning that the quantum tunneling

mechanism is effective which can be generalized to analyze the generation and variation of solutions in optimization algorithms.

3. Frequency transformation and amplitude transformation are implemented on the Griewank function to make a generalization. The phenomenon is further verified and explained by experimental analysis on the generalized Griewank function.

Fig. 1 shows the roadmap of this study.

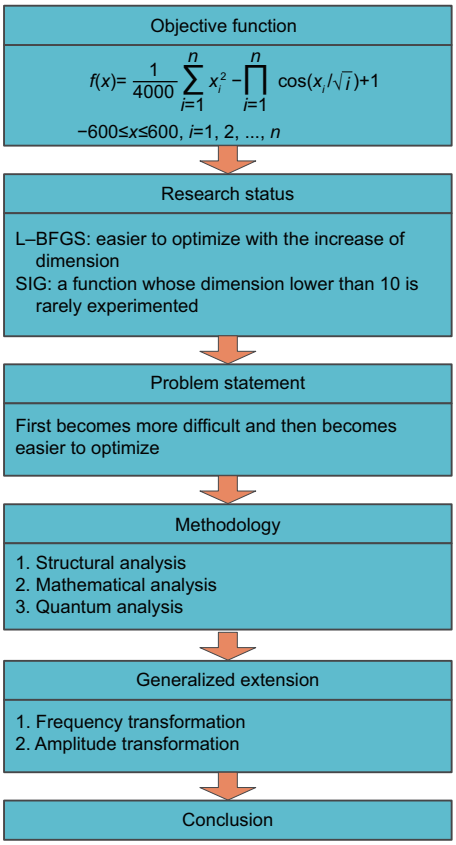

Fig. 1 Roadmap of this study

2 Statement of the problem

2.1 Definition of the Griewank function

The Griewank function is defined as follows:

$$f(x)=\frac{1}{4000}\sum_{i=1}^{n}x_i^2-\prod_{i=1}^{n}\cos\left(x_i/\sqrt{i}\right)+1, \quad (1)$$

where $-600\le x_i\le 600$, $i=1,2,\ldots,n$.

The global minimum of the Griewank function is unique, located in the origin. From Eq. (1), we can find that the Griewank function is composed mainly of two parts:

$$f(x)=g(x)-h(x)+1. \quad (2)$$

The first part is $g(x)$, which is a quadratic convex function as follows:

$$g(x)=\frac{1}{4000}\sum_{i=1}^{n}x_i^2. \quad (3)$$

The unique global minimum of $g(x)$ is located in the origin, which is the same as that of the Griewank function.

The second part is $h(x)$, which is an oscillatory nonconvex function as follows:

$$h(x)=\prod_{i=1}^{n}\cos\left(x_i/\sqrt{i}\right). \quad (4)$$

The graphs of two-dimensional (2D) $f(x)$, $g(x)$, and $h(x)$ over the box $[-20,20]\times[-20,20]$ are shown in Figs. 2a–2c. The Griewank function is the superposition of $g(x)$ and $h(x)$. $h(x)$ has a large number of equivalent global minima. The oscillations introduced by $h(x)$ give rise to the many local minima of the Griewank function.

2.2 Phenomenon of optimizing the Griewank function with the increase of dimension

To better illustrate this phenomenon, we optimize the Griewank function with SPSO2011 (Zambrano-Bigiarini et al., 2013) in this subsection. The search space of the Griewank function is [−100, 100]. The function dimension gradually increases from 2 to 100. The scope of the inertia weight is set within [0.4, 0.9]. The program is coded in Matlab R2013a and executed on a Microsoft Surface Pro4 (Intel Core$^{\text{TM}}$ i5-7300U CPU @2.60 GHz, 6 GB RAM, and operation system Windows 10). Experimental results are obtained based on 51 independent trials. For each run, the difference between the current function value $f_{\min}(x)$ and theoretical minimum function value $f'_{\min}(x)$ is calculated. If $|f_{\min}(x)-f'_{\min}(x)|<10^{-3}$, this run is considered to be successful.

The following evaluation criteria are calculated: best fitness value (Best), mean fitness value (Mean), standard deviation of fitness values (Std), success rate (SR), lower bound of 95% confidence limit (L95), and upper bound of 95% confidence limit (U95). Experimental results in different dimensions are listed in Table 1.

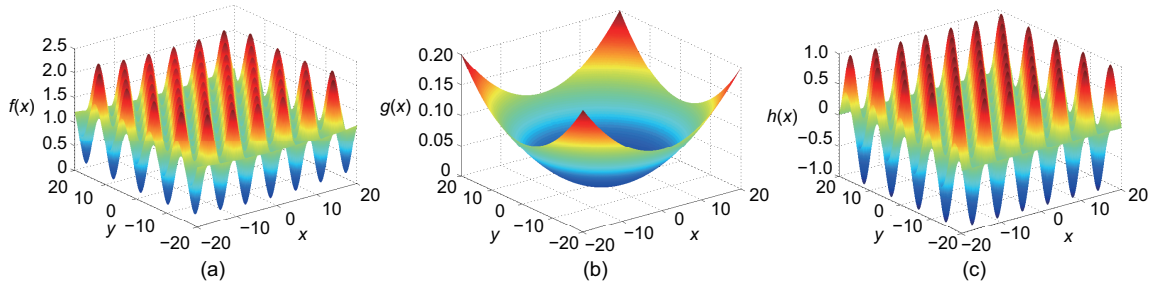

Fig. 2 Three-dimensional images of the Griewank function and its components: (a) Griewank ($f(x)$); (b) $g(x)$; (c) $h(x)$

Table 1 Best fitness value, mean fitness value, standard deviation of fitness values, and the success rate of SPSO2011 on the Griewank function with function dimension gradually increasing from 2 to 100

Dimension	Best	Mean	Std	SR (%)	L95	U95
D=2	9.19×10^{-10}	4.95×10^{-7}	3.09×10^{-7}	100.00	4.13×10^{-7}	5.78×10^{-7}
D=3	4.32×10^{-7}	2.74×10^{-3}	3.35×10^{-3}	52.94	1.82×10^{-3}	3.66×10^{-3}
D=4	3.35×10^{-4}	1.04×10^{-2}	6.15×10^{-3}	5.88	8.71×10^{-3}	1.21×10^{-2}
D=5	5.02×10^{-5}	2.21×10^{-2}	1.31×10^{-2}	3.92	1.85×10^{-2}	2.57×10^{-2}
D=6	4.70×10^{-4}	2.90×10^{-2}	1.16×10^{-2}	1.96	2.58×10^{-2}	3.22×10^{-2}
D=7	9.92×10^{-7}	3.02×10^{-2}	1.44×10^{-2}	1.96	2.63×10^{-2}	3.42×10^{-2}
D=8	7.84×10^{-7}	3.53×10^{-2}	4.34×10^{-2}	7.84	2.34×10^{-2}	4.72×10^{-2}
D=9	7.30×10^{-7}	2.26×10^{-2}	1.35×10^{-2}	1.96	1.89×10^{-2}	2.63×10^{-2}
D=10	6.77×10^{-7}	1.71×10^{-2}	1.17×10^{-2}	13.73	1.39×10^{-2}	2.03×10^{-2}
D=11	6.23×10^{-7}	1.27×10^{-2}	1.10×10^{-2}	25.49	9.68×10^{-3}	1.57×10^{-2}
D=12	6.23×10^{-7}	8.93×10^{-3}	1.15×10^{-2}	50.98	5.77×10^{-3}	1.21×10^{-2}
D=13	5.34×10^{-7}	6.86×10^{-3}	8.05×10^{-3}	47.06	4.65×10^{-3}	9.07×10^{-3}
D=14	6.70×10^{-7}	6.38×10^{-3}	7.70×10^{-3}	49.02	4.27×10^{-3}	8.50×10^{-3}
D=15	4.60×10^{-7}	7.63×10^{-3}	9.43×10^{-3}	45.10	5.04×10^{-3}	1.02×10^{-2}
D=20	7.13×10^{-7}	7.05×10^{-3}	9.72×10^{-3}	49.02	4.83×10^{-3}	9.72×10^{-3}
D=30	8.74×10^{-7}	5.99×10^{-3}	8.05×10^{-3}	52.94	3.78×10^{-3}	8.20×10^{-3}
D=40	8.90×10^{-7}	5.94×10^{-3}	7.68×10^{-3}	56.86	3.83×10^{-3}	8.05×10^{-3}
D=50	9.30×10^{-7}	5.99×10^{-3}	6.95×10^{-3}	49.02	4.08×10^{-3}	7.90×10^{-3}
D=60	9.55×10^{-7}	2.22×10^{-3}	4.27×10^{-3}	76.47	1.05×10^{-3}	3.39×10^{-3}
D=80	9.67×10^{-7}	3.48×10^{-3}	6.03×10^{-3}	70.59	1.83×10^{-3}	5.14×10^{-3}
D=100	9.59×10^{-7}	4.73×10^{-3}	6.85×10^{-3}	58.82	2.85×10^{-3}	6.61×10^{-3}

Best: best fitness value; Mean: mean fitness value; Std: standard deviation of fitness values; SR: success rate; L95: lower bound of 95% confidence limit; U95: upper bound of 95% confidence limit

It can be seen from the experimental data that when $D=2$, SR=100%. SR drops to the minimum when $D=6$ and rises to a relatively stable value when $D=12$. With the increase of dimension, the SR of SPSO2011 first falls and then rises. When dimension is lower than six, the SR of SPSO2011 falls with the increase of dimension, exhibiting the opposite phenomenon (Zhang et al., 2003; Karaboga and Basturk, 2007; Chen and Zhao, 2009; Akbari et al., 2010; Akay and Karaboga, 2012; Gao et al., 2012; Rao et al., 2012). This indicates that the structure of the 2D Griewank function is simple and is easy to optimize. SR decreases when the dimension increases and reaches the minimum when the dimension is around six. When the dimension continues to rise, the number of local minima increases, but the impact of the oscillation part of the Griewank function decreases. This makes it easier to find the global optimum of the high-dimensional Griewank function.

Furthermore, we optimize Ackley and Levy with SPSO2011.

The definition of the Ackley function is as follows:

$$f(x)=-20\exp\Bigg(-0.2\sqrt{\frac{1}{n}\sum_{i=1}^{n}x_i^2}\Bigg)$$

$$-\exp\left(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)\right)+20+\mathrm{e}, \tag{5}$$

where $-600\le x_i\le 600$, $i=1,2,\ldots,n$.

The definition of the Levy function is as follows:

$$\begin{aligned} f(x)=&\sin^2(\pi w_1)+(w_n-1)^2\left[1+\sin^2(2\pi w_n)\right]\\ &+\sum_{i=1}^{n-1}(w_i-1)^2\left[1+10\sin^2(\pi w_i+1)\right],\\ w_i=&1+\frac{x_i-1}{4}, \end{aligned} \tag{6}$$

where $-600\le x_i\le 600$, $i=1,2,\ldots,n$.

The relationship between SR and D is shown in Fig. 3. With the increase of dimension D from 2 to 60, the SR of Griewank first falls and then rises, and those of Ackley and Levy decrease gradually.

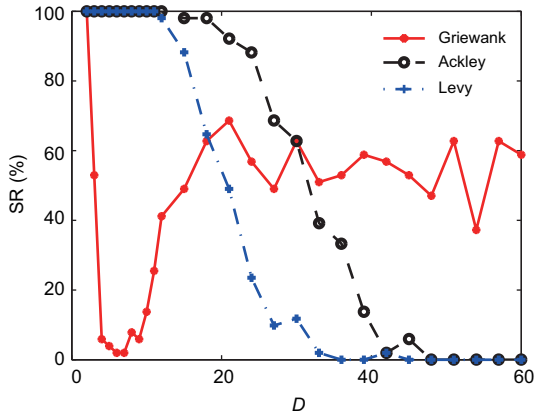

Fig. 3 Success rates (SRs) of Griewank, Ackley, and Levy, optimized with SPSO2011

3 Methodology

In this section, the phenomenon is analyzed in three ways: structural analysis, mathematical analysis, and quantum analysis.

3.1 Structural analysis

To explain the cause of the phenomenon mentioned above, detailed structural analysis of the Griewank function is described. We separately project 2-, 3-, 5-, and 10-dimensional Griewank functions to a plane (Fig. 4). One thousand sampling points are taken evenly in the solution space. For each sampling point, the coordinates of each dimension are the same. These sampling points can be projected to any plane, and the projections of these sampling points to any plane are the same.

In Fig. 4, each subfigure contains three curves of different colors. The green curve is the projection of $g(x)$ with a unique global minimum. The red curve is the projection of $h(x)$ with many equivalent global minima. The blue curve superimposed by red and green curves represents the projection of the Griewank function.

As can be seen from Fig. 4, the Griewank function is obviously of a two-scale structure in low dimensions. The two parts of the Griewank function in Eq. (2) can be considered as two scales of this function. $g(x)$ is the large scale of the Griewank function, representing the global information. $h(x)$ is the small scale of the Griewank function, representing the local information.

For a 2D Griewank function, the structure is relatively simple and is easy to optimize. The numbers of the Griewank function's minima for the two search space for up to three dimensions are listed in Table 2 (Cho et al., 2008). For $[-14, 14]^n$, the number of minima increases from 5 to 157, and for $[-28, 28]^n$, the number of minima increases from 9 to 1215. As for a low-dimensional Griewank function, with the increase of dimension, the number of local minima increases exponentially. The Griewank function becomes more difficult to optimize, which is similar to other functions.

When the number of dimensions is larger than six, the Griewank function becomes easier to optimize with the increase of dimension. By comparing the four subfigures of Fig. 4, we can see that the impact of $h(x)$ decreases with the increase of dimension. The local minima induced by $h(x)$ can be neglected in high-dimensional scenarios. The shapes of $f(x)$ and $h(x)$ are very similar for a high-dimensional Griewank function.

3.2 Mathematical analysis

As mentioned in Section 3.1, the Griewank function is composed of two major parts. The weight

Table 2 Number of minima for $[-14, 14]^n$ and $[-28, 28]^n$

n	$[-14, 14]^n$	$[-28, 28]^n$
1	5	9
2	31	111
3	157	1215

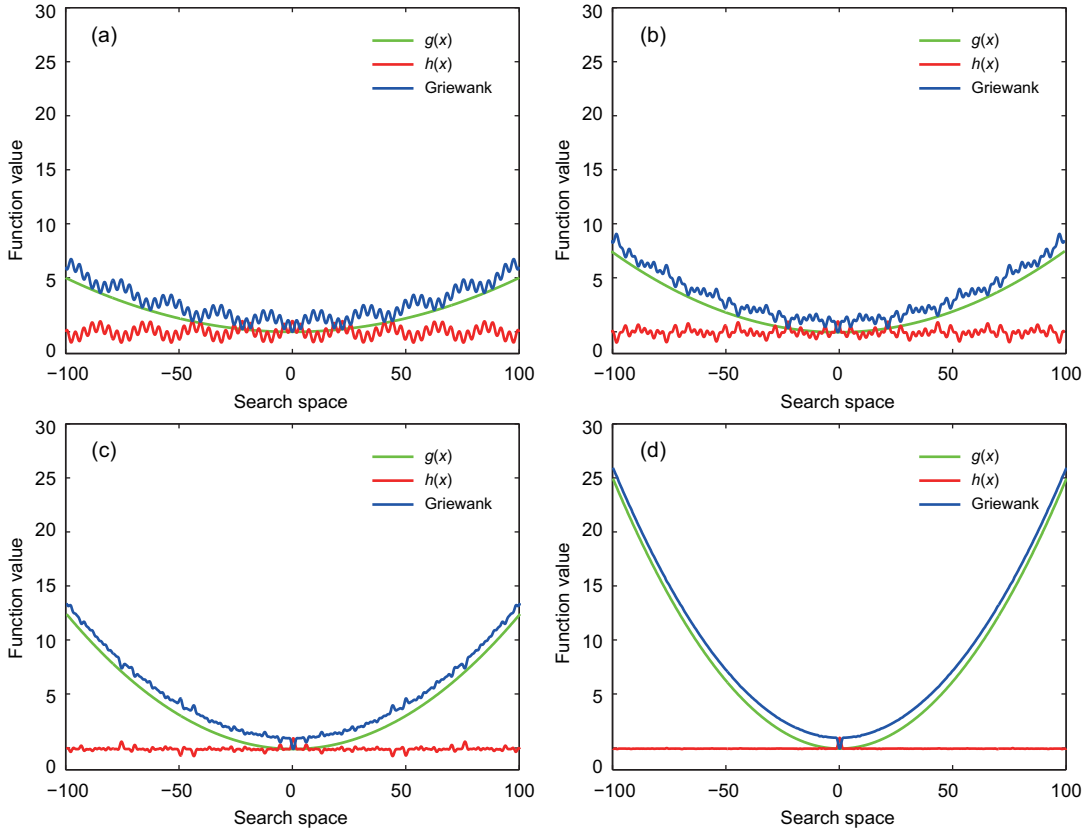

Fig. 4 Projecting an n-dimensional Griewank function to a plane: (a) $n=2$; (b) $n=3$; (c) $n=5$; (d) $n=10$. References to color refer to the online version of this figure

of each part determines the shape of the Griewank function.

The tangent slope represents the tilt degree of a curve. The higher the tangent slope of the current sampling point is, the harder it is to "jump out" of local optima. Eq. (7) represents the tangent slope of the Griewank function's curve at point x. The tangent slopes of the global minimum and local minima are all 0.

$$\begin{aligned} f'(x) =& g'(x) - h'(x) \\ =& \frac{1}{2000}\sum_{i=1}^{n} x_i - \sum_{i=1}^{n} \sin\left(1/\sqrt{i}\right) \\ & \cdot \prod_{i=1, j\neq i}^{n} \cos\left(x_i/\sqrt{i}\right). \end{aligned} \tag{7}$$

We can see that $f'(x)$ is composed of two parts, $g'(x)$ and $h'(x)$. $g'(x)$ is the sum of x_i, n-dimensional coordinates. The importance of $g'(x)$ and $h'(x)$ is respectively defined as $L_{g'(x)}$ and $L_{h'(x)}$:

$$L_{g'(x)} = \frac{g'(x)}{g'(x) + h'(x)}, \tag{8}$$

$$L_{h'(x)} = \frac{h'(x)}{g'(x) + h'(x)}. \tag{9}$$

With the increase of n, the absolute value of $g'(x)$ increases accordingly. $h'(x)$ is the product of n values belonging to the interval $[-1, 1]$. As n increases, the product of such values becomes so small and thus can be neglected with respect to $g'(x)$, as shown below:

$$\lim_{n\to+\infty} L_{g'(x)} = 1, \tag{10}$$

$$\lim_{n\to+\infty} L_{h'(x)} = 0. \tag{11}$$

This means that with the increase of dimension, the impact of $h(x)$ on $f(x)$ decreases. The local minima introduced by $h(x)$ have little impact on the global optimization process of a high-dimensional Griewank function.

3.3 Quantum analysis

In this subsection, the phenomenon is analyzed by quantum tunnel effect (Muthukrishnan et al., 2016).

For most optimization algorithms, to find the global optimal solution of the objective function, an iterative operation is executed from current sampling points towards the position of the optimal solution through various mechanisms. Falling into a local optimal solution is a common phenomenon in this process. Most global optimization algorithms have the solution generation mechanism to "jump out" of local optima, e.g., Gaussian mutation in FWA (Tan and Zhu, 2010), uniform random disturbance in SA (Kirkpatrick et al., 1983), the idea generation process in BSO (Zhou et al., 2012), and mutation in DE (Qin et al., 2009). These mechanisms can be explained by the tunnel effect.

MQHOA (Wang et al., 2013) is a newly proposed optimization algorithm based on the quantum theory, which has attained good performance on multimodal optimization (Wang et al., 2018a) and combination optimization (Wang et al., 2016), and can be applied to practical problems. The inspiration of MQHOA is the probability interpretation of a quantum wave function (Wang et al., 2018b), which provides MQHOA the quantum tunnel effect to "jump out" of local optima. This provides a new approach for analyzing the phenomenon.

3.3.1 Wave function

Wave function is an important concept in quantum theory, meaning the probability that a particle will appear at a certain location.

Gaussian function is used as the probability distribution function in MQHOA's sampling process. MQHOA's wave function $|\psi_k(x)|^2$ is defined as the superposition of k Gaussian functions, shown as follows:

$$\begin{aligned}|\psi_k(x)|^2 &= \sum_{i=1}^{k} N(x_i, \sigma_{\mathrm{s}}^2)\\ &= \sum_{i=1}^{k} \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{s}}} \mathrm{e}^{-(x-x_i)^2/(2\sigma_{\mathrm{s}}^2)}.\end{aligned} \tag{12}$$

Fig. 5 is the wave function image of the Griewank function at a high-level energy state with σ_{s}=0.625. Fig. 5 shows the probability distribution of the current sampling. The raised areas in this figure are sampling areas with high probability.

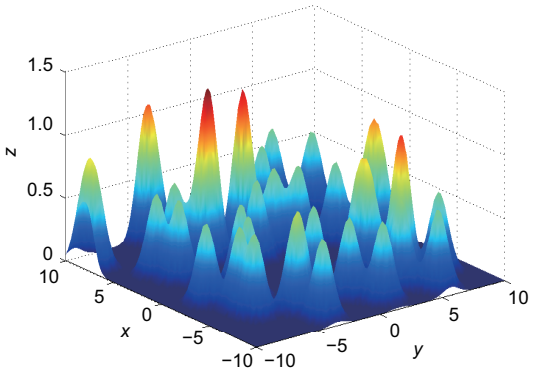

Fig. 5 Wave function of the Griewank function

3.3.2 Quantum tunnel effect based on MQHOA

Quantum tunnel effect is the phenomenon under the probability interpretation of MQHOA's wave function in the interval $[a, b]$. The sampling point can appear anywhere as shown in Fig. 5, once the new sampling point is superior to the current solution. The possibility of quantum tunnel effect is defined as follows:

$$\begin{aligned}P_{(a,b)} &= \int_a^b |\psi_k(x)|^2 \mathrm{d}x\\ &= \sum_{i=1}^{k} \int_a^b N(x_i, \sigma_{\mathrm{s}}^2) \mathrm{d}x\\ &= \sum_{i=1}^{k} \int_a^b \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{s}}} \mathrm{e}^{-(x-x_i)^2/(2\sigma_{\mathrm{s}}^2)} \mathrm{d}x\\ \text{s.t.}\quad & f(x) > f(a).\end{aligned} \tag{13}$$

Eq. (13) can be explained with the help of Fig. 6. In Fig. 6, the solid blue line represents the objective function $f(x)$. The global optimum of $f(x)$ is located at -2, and the local optimum is located at 2. The dashed pink line represents the wave function of $f(x)$ with reference to Eq. (12). Point A is one of the current sampling points, locating in the local optimal region. To find the global optimal solution, we should obtain the sampling point in the global optimal region. The dashed green line is the Gaussian sampling function with current σ_{s}. The possibility of sampling point A moving to the global optimal region $[a, b]$ is defined as follows:

$$\begin{aligned}P_1 &= \int_a^b N(c, \sigma_{\mathrm{s}}^2) \mathrm{d}x\\ &= \int_a^b \frac{1}{\sqrt{2\pi}\sigma_{\mathrm{s}}} \mathrm{e}^{-(x-c)^2/(2\sigma_{\mathrm{s}}^2)} \mathrm{d}x.\end{aligned} \tag{14}$$

P_1 decreases dramatically with the reduction of σ_{s}.

Quantum tunnel effect occurs when a better solution is generated in the dark red area, for example, point C. If a worse solution such as B is generated, this solution will be discarded and the quantum tunnel effect will not occur.

From Fig. 6, the two dashed green lines mean that with the iteration of MQHOA, the variance of Gaussian sampling decreases, and the possibility of quantum tunnel effect on the current sampling point decreases accordingly.

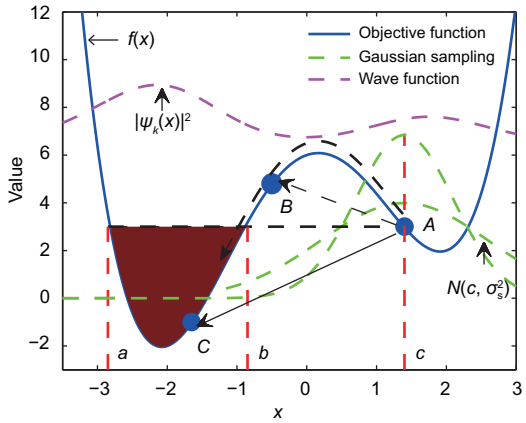

Fig. 6 Schematic of quantum tunnel effect. References to color refer to the online version of this figure

There are three important parameters in Fig. 6. With the increase of dimension, these parameters of the Griewank function will alter correspondingly: (1) the vertical distance between the global optimum and the local optimum; (2) the horizontal distance between the global optimum and the local optimum; (3) the oscillation amplitude of the local optimal region.

It can be found from Table 2 that with the increase of dimension, the number of local optima increases exponentially. When projected into one dimension, we can find that the horizontal distance between the global optimum and the local optimum stays stable (Fig. 4).

As can be seen from Eqs. (1) and (2), with the increase of dimension, $f(x)$ increases with the help of $g(x)$, and the decrease of $h(x)$ can be neglected. Combining Figs. 4 and 6, we can find that, with the increase of dimension, the vertical distance between the global optimum and the local optimum increases, and that the oscillation amplitude of the local optimal region decreases.

3.3.3 Experimental explanation of the quantum tunnel effect

What effect does these parameters have on optimizing the Griewank function? The double-well function, which includes all of the above parameters, is a suitable test function. The double-well function is defined as follows:

$$f(x) = V\frac{(x^2 - a^2)^2}{a^4} + bx, \tag{15}$$

where $V = f(0)$ represents the barrier height of the double-well function, a is the vertical distance between the global optimum and the local optimum, and b is the horizontal distance between the global optimum and the local optimum. We conduct experiments on parameters V and b of the double-well function to explain the phenomenon by quantum tunnel effect.

Fig. 7 shows the influence of V on the optimization of SR. With the increase of V, SR decreases gradually. This means that the probability of crossing over the barrier to the optimal solution area decreases with the increase of the barrier height. This is because the optimal solution area becomes steeper with the increase of V. The quantum tunnel effect decreases. By comparing the three curves in Fig. 7, we can see that with the increase of b, SR increases correspondingly.

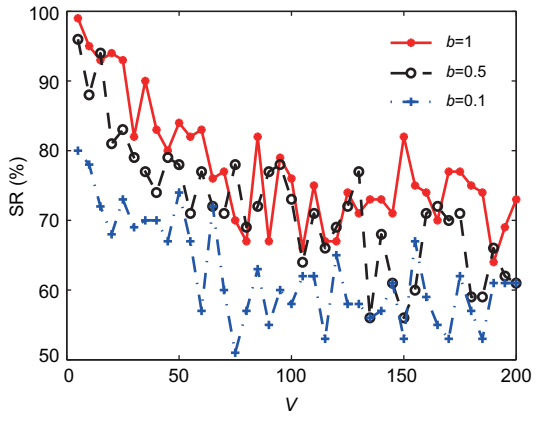

Fig. 7 An experiment on parameter V ($a = 6$, V increases from 5 to 200)

Fig. 8 shows the influence of b on the optimization of SR. With the increase of b, SR increases gradually. This means that with the increase of the horizontal distance between the global optimum and the local optimum, the probability of "jumping out" of

local optima increases. The quantum tunnel effect is enhanced. By comparing the three curves in Fig. 8, we can see that with the increase of V, SR decreases correspondingly. The same phenomenon can be observed in Fig. 7.

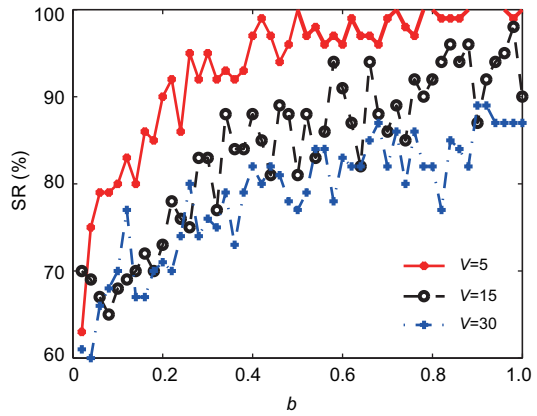

Fig. 8 An experiment on parameter b ($a = 6$, b increases from 0.02 to 1)

To better illustrate the quantum tunnel effect, we track MQHOA's wave function in the optimization process of the 2D Griewank function.

Three sub-experiments are done with different sampling group parameter (k) and initial sigma (σ_0). The snapshots of the wave function in the optimization process are presented in Fig. 9. Detailed descriptions are as follows:

1. Figs. 9a–9d: k=20, σ_0=2

In Fig. 9a, σ_0 is set to 2, 10% of the definition domain. This means that in the initial stage each point's sampling area is a small portion of the sampling area. The initial wave function image is shown in Fig. 9a. The sampling point can "jump" to the global optimal region due to quantum tunnel effect. The possibility of quantum tunnel effect decreases with the decrease of σ. In this sub-experiment, the global optimal solution is obtained accurately.

2. Figs. 9e–9h: k=20, σ_0=20

In Fig. 9e, σ_0 is equal to the value of the definition domain. This guarantees that the sampling area covers the entire domain more evenly. The initial wave function image is shown in Fig. 9e, which is very flat in shape. With the iteration of MQHOA, sampling points gather gradually, and σ shrinks. In Fig. 9f, several competitive sampling areas appear. Due to the quantum tunnel effect, the competitive areas are constantly changing. In this sub-experiment, the global optimal solution is obtained accurately.

3. Figs. 9i–9l: k=10, σ_0=2

In Fig. 9i, σ_0 is set to 2. The initial distribution of sampling points is more scattered. With the process of optimization, these points distribute more evenly, as shown in Fig. 9j. Then three competitive sampling areas appear and evolve into two more competitive sampling areas. The vertical distance between the two areas is very small. With the decrease of the vertical distance between the global optimum and local optimum, the quantum tunnel effect decreases. It becomes more difficult to obtain the global optimal solution. Eventually, the global optimal solution is not obtained accurately within a specified number of iterations.

4 Generalization of the Griewank function with experimental analysis

In Section 3, the phenomenon is researched from the methodology perspective. In this section, we focus on the generalization of the Griewank function to make an in-depth analysis of the phenomenon. The Griewank function is composed of two parts, which can be considered as two scales. The weights of the two parts represent the competitive significance of the two scales, which has an important influence on the optimization results of the Griewank function. Based on this structure, frequency transformation and amplitude transformation are implemented on the Griewank function to make a generalization. The phenomenon is further verified and explained by experimental analysis on a generalized Griewank function.

4.1 Definition of the generalized Griewank function

Based on the Griewank function's two-scale structure, a generalized Griewank function is defined as follows:

$$\begin{aligned} f_1(x) &= g_1(x) - h_1(x) + 1 \\ &= \frac{A}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n} \cos\left(F \cdot \frac{x_i}{\sqrt{i}}\right) + 1, \end{aligned} \quad (16)$$

where $-600 \le x \le 600$, $i = 1, 2, \ldots, n$.

Compared with Eq. (2), we introduce two new parameters: frequency parameter F and amplitude

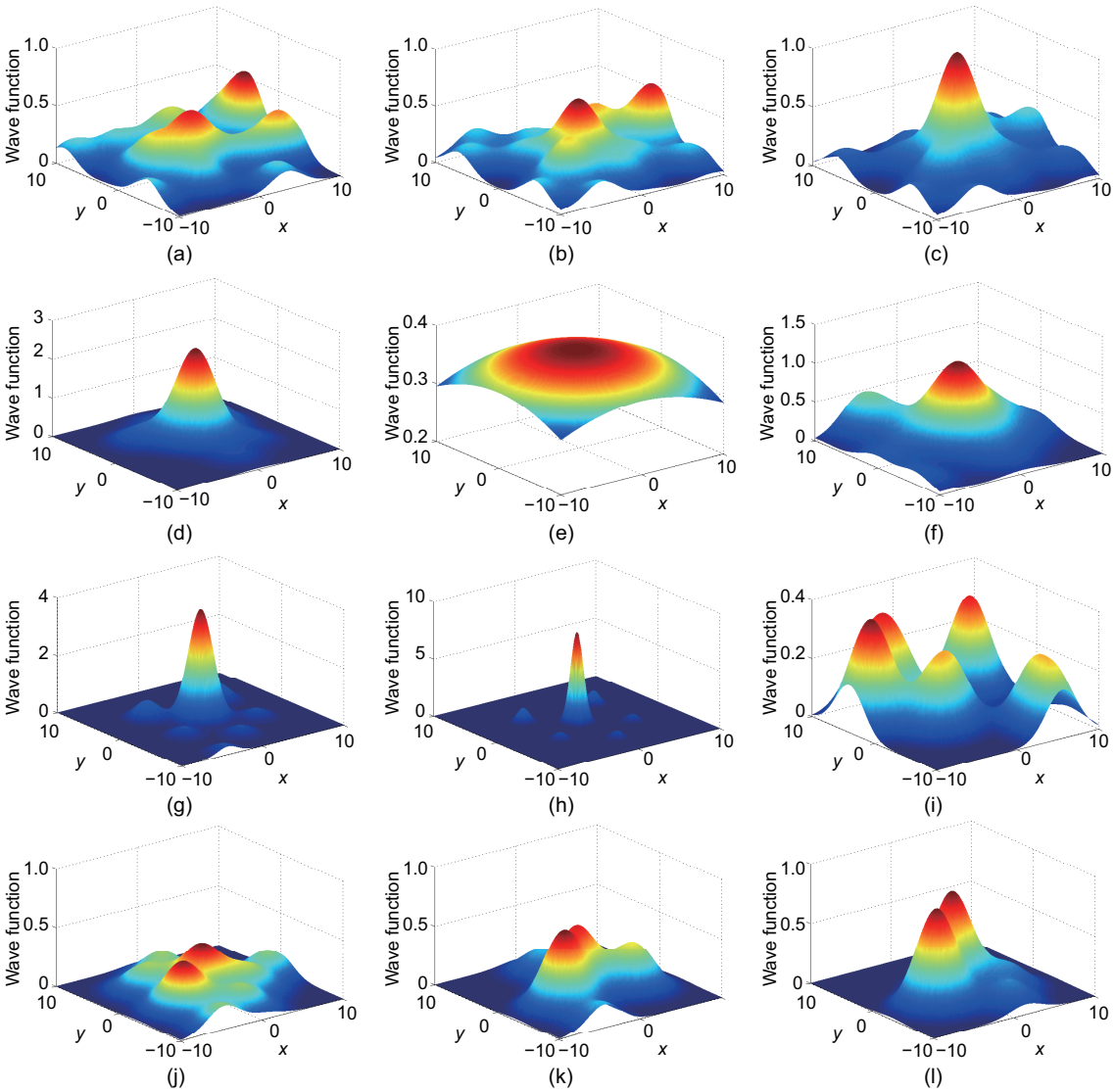

Fig. 9 Wave functions of the Griewank function optimized by MQHOA: (a) initial status (k=20, σ_0=2, i=1, σ=2); (b) competitive sampling areas appear (k=20, σ_0=2, i=400, σ=2); (c) aggregate and approach to the optimal solution (k=20, σ_0=2, i=450, σ=2); (d) the global optimal solution is obtained accurately (k=20, σ_0=2, i=20 000, σ=2); (e) initial status, very flat in shape (k=20, σ_0=20, i=1, σ=20); (f) competitive sampling areas appear (k=20, σ_0=20, i=505, σ=2.5); (g) the optimal area becomes obvious (k=20, σ_0=20, i=7611, σ=1.25); (h) the global optimal solution is obtained accurately (k=20, σ_0=20, i=7613, σ=0.625); (i) initial status, the sampling area is very scattered (k=10, σ_0=2, i=1, σ=2); (j) sampling points gather gradually (k=10, σ_0=2, i=250, σ=2); (k) three competitive sampling areas appear (k=10, σ_0=2, i=2500, σ=2); (l) two competitive sampling areas compete with each other, and the global optimal solution is not obtained accurately (k=10, σ_0=2, i=20 000, σ=2)

parameter A. With parameters F and A, the frequency and amplitude transformed Griewank functions are constructed as follows:

1. Frequency transformed Griewank function

F is used to adjust the frequency of the oscillatory nonconvex function. Fig. 10 shows the projection of a three-dimensional (3D) frequency transformed Griewank function to a plane, with the frequency parameter F separately setting to 0.5, 1, 3, and 10. In Fig. 10, the green curve remains unchanged. With the increase of F, the number of local minima increases rapidly. It is more likely to be trapped in local minima.

2. Amplitude transformed Griewank function

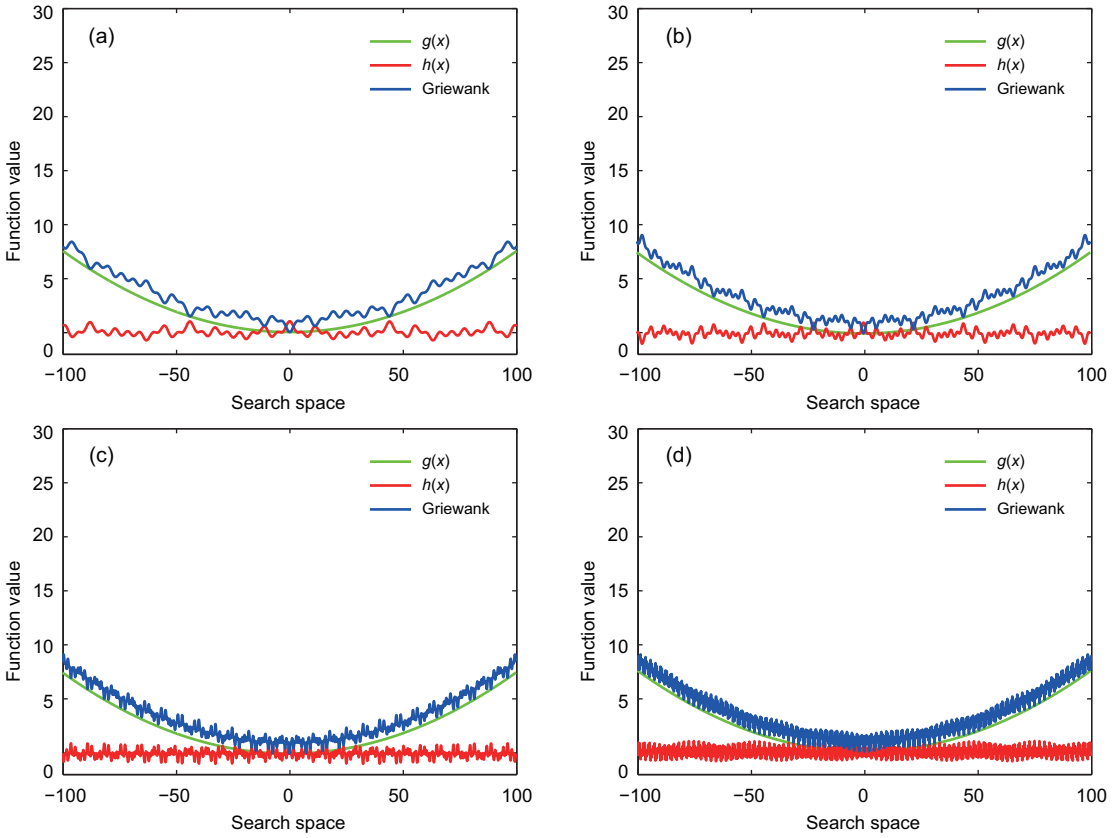

Fig. 10 Projecting a three-dimensional frequency transformed Griewank function to a plane: (a) F=0.5; (b) F=1; (c) F=3; (d) F=10. References to color refer to the online version of this figure

Parameter A is used to adjust the amplitude of the quadratic convex function. Fig. 11 shows the projection of a 3D amplitude transformed Griewank function to a plane. The amplitude parameter A is separately set to 0.1, 1, 3, and 5. The red curve is the projection of $h_1(x)$, which results in a large number of local minima. The green curve is the projection of $g_1(x)$. As can be seen from Fig. 11, with the increase of A, the amplitude of quadratic convex function $g_1(x)$ increases. The blue curve superimposed by red and green curves represents the projection of the amplitude transformed Griewank function. By comparing the four blue curves of Fig. 11, we can see that the impact of $h_1(x)$ decreases with the increase of amplitude, and it becomes easier to obtain the global optimal solution.

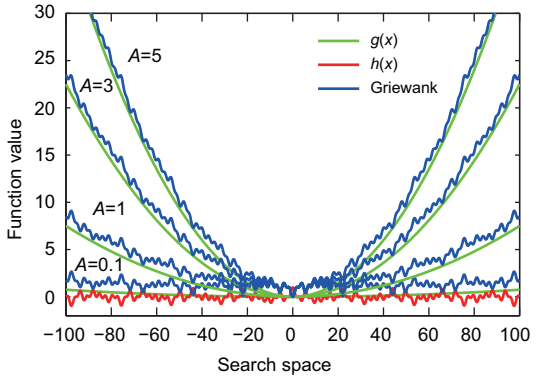

Fig. 11 Projecting a three-dimensional amplitude transformed Griewank function to a plane (amplitude parameter A=0.1, 1, 3, or 5). References to color refer to the online version of this figure

4.2 Algorithm flow of MQHOA

As we pointed out in Section 3.1, the Griewank function is a two-scale structure. To perform an in-depth study on the generalized Griewank function, we optimize the generalized Griewank function with MQHOA in this study.

Energy level stabilization, energy level decline, and scale adjustment are the main steps of MQHOA. The initial scale equals the search space of each dimension. After energy level decline, the current scale is cut to half. Then energy level stabilization

and energy level decline are executed repeatedly at a smaller scale. The details of the MQHOA were described in Wang et al. (2018c).

The pseudocode of MQHOA is listed in Algorithm 1. The notations used in this subsection are listed in Table 3. Parameters of MQHOA in this subsection are set as $\sigma_{\min} = 10^{-6}$ and $k = 20$.

Algorithm 1 MQHOA

Initialize k, $\sigma_{\min}$, MIN, MAX, and σ_{s}=MAX − MIN
Randomly generate x_i in [MIN, MAX], where $i = 1, 2, \ldots, k$
Calculate $F_i = f(x_i)$, F^{opt}=$f_{\min}(x_i)$, and $x^{\mathrm{opt}} = x_i$
while $\sigma_{\mathrm{s}} > \sigma_{\min}$ **do**
  Set $\mathrm{Flag}_{\mathrm{stable}} = 0$
  while $\mathrm{Flag}_{\mathrm{stable}} == 0$ **do**
    $\forall\, x_i$, generate $x_i' \sim N(x_i, \sigma_{\mathrm{s}}^2)$
    $\forall\, x_i$ and x_i', if $F(x_i') < F(x_i)$ then $x_i = x_i'$, $\mathrm{Flag}_{\mathrm{stable}} = 1$
  end while
  Update the worst solution $x^{\mathrm{worst}} = x^{\mathrm{mean}}$
  $\sigma_{\mathrm{s}} = \sigma_{\mathrm{s}}/2$
end while
Output: x^{opt} and F^{opt}

Table 3 Notations

Notation	Definition
MIN	Lower bound of the search space
MAX	Upper bound of the search space
MaxFES	Maximum number of function evaluations in each run
$\mathrm{Flag}_{\mathrm{stable}}$	Stop criterion of stabilization operation
k	Number of sampling points
$\sigma_{\min}$	Convergence accuracy
σ_{s}	Current scale
x_i	Current optimal solution, $i = 1, 2, ..., k$
x_i'	New solution generated by $N(x_i, \sigma_{\mathrm{s}}^2)$
x^{mean}	Mean value of x_i
x^{worst}	Worst solution in x_i
x^{opt}	Best solution in x_i

The following analysis of MQHOA is based on the pseudocode in Algorithm 1.

There are two while loops in the pseudocode, covering the three main stages of MQHOA.

1. Energy level stabilization

The inner layer loop corresponds to energy level stabilization, which is the basic arithmetic unit of MQHOA. After initialization, new solutions are generated in turn by Gaussian sampling based on every candidate solution. By this way, the nearby area of candidate solutions is fully exploited, searching for better solutions. A flag with an initial value of 0 is used to evaluate the status of energy level stabilization. If the function value of any sampling point is better than the corresponding candidate solution, the corresponding candidate solution will be replaced by the new sampling point, and the flag is set to 1 to end the energy level stabilization stage.

2. Energy level decline

Energy level decline is an important component of MQHOA's physical model, and is executed immediately after the energy level stabilization stage. A new sampling point is generated based on all of current candidate solutions to replace the current worst candidate solution. The coordinates of the new sampling point are the average of all the coordinates. Global information is introduced to the system in this stage, which can increase the diversity of the sampling points.

Different from the conventional MQHOA, we make changes to the energy level decline strategy in this study. Energy level decline will not be executed repeatedly at the current scale; it occurs once at each scale. This mechanism can effectively avoid invalid energy level declination in conventional MQHOA, speed up the convergence, and improve the accuracy of MQHOA.

3. Scale adjustment

Scale adjustment is executed immediately after the energy level decline stage. The current scale is cut to half to search intensively in a smaller area. If the variance of current candidate solutions is larger than the search accuracy ($\sigma_{\mathrm{s}} > \sigma_{\min}$), the program jumps to the energy level stabilization stage. The program terminates until $\sigma_{\mathrm{s}} \leq \sigma_{\min}$.

4.3 Experimental study on the generalized Griewank function

4.3.1 Optimizing the frequency transformed Griewank function

As for a frequency transformed Griewank function, frequency gradually increases from 1 to 10. Three typical dimensions are selected: D=2, D=5, and D=10. Experimental results of optimizing the frequency transformed Griewank function are listed in Table 4.

From the data displayed in Table 4, we have the following two observations:

1. For all of the three dimensions, as the frequency increases, it becomes more difficult to

Table 4 Best fitness value, mean fitness value, standard deviation of fitness values, and the success rate of MQHOA on the frequency transformed Griewank function with frequency gradually increasing from 1 to 10

Dimension	F	Best	Mean	Std	SR (%)	L95	U95
$D=2$	1	3.82×10^{-13}	1.02×10^{-3}	2.57×10^{-3}	86.27	3.15×10^{-4}	1.73×10^{-3}
	1.1	1.96×10^{-12}	1.32×10^{-3}	2.54×10^{-3}	78.43	6.23×10^{-4}	2.02×10^{-3}
	1.2	1.36×10^{-11}	1.31×10^{-3}	2.26×10^{-3}	74.51	6.90×10^{-4}	1.93×10^{-3}
	1.5	3.17×10^{-13}	1.53×10^{-3}	1.71×10^{-3}	54.90	1.01×10^{-3}	2.00×10^{-3}
	2	1.11×10^{-12}	1.06×10^{-3}	1.39×10^{-3}	52.94	6.80×10^{-4}	1.44×10^{-3}
	4	1.77×10^{-11}	1.11×10^{-3}	1.09×10^{-3}	54.90	8.11×10^{-4}	1.41×10^{-3}
	10	4.50×10^{-12}	1.50×10^{-3}	1.75×10^{-3}	47.06	1.02×10^{-3}	1.98×10^{-3}
$D=5$	1	1.31×10^{-13}	5.36×10^{-3}	5.23×10^{-3}	45.10	3.92×10^{-3}	6.80×10^{-3}
	1.1	9.30×10^{-14}	5.19×10^{-3}	4.77×10^{-3}	39.22	3.88×10^{-3}	6.50×10^{-3}
	1.2	1.97×10^{-13}	6.01×10^{-3}	5.12×10^{-3}	29.41	4.61×10^{-3}	7.42×10^{-3}
	1.5	2.73×10^{-13}	4.99×10^{-3}	3.84×10^{-3}	19.61	3.94×10^{-3}	6.04×10^{-3}
	2	1.91×10^{-12}	5.22×10^{-3}	4.00×10^{-3}	3.92	4.12×10^{-3}	6.32×10^{-3}
	4	6.17×10^{-4}	5.51×10^{-3}	3.43×10^{-3}	3.92	4.57×10^{-3}	6.45×10^{-3}
	10	4.93×10^{-4}	4.43×10^{-3}	2.65×10^{-3}	3.92	3.70×10^{-3}	5.16×10^{-3}
$D=10$	1	4.04×10^{-13}	1.45×10^{-4}	1.04×10^{-3}	98.04	-1.40×10^{-4}	4.30×10^{-4}
	1.1	6.03×10^{-13}	1.60×10^{-4}	1.14×10^{-3}	98.04	-1.53×10^{-4}	4.73×10^{-4}
	1.2	6.16×10^{-13}	1.34×10^{-4}	9.59×10^{-4}	98.04	-1.29×10^{-4}	3.98×10^{-4}
	1.5	1.27×10^{-12}	6.66×10^{-4}	1.52×10^{-3}	82.35	2.21×10^{-4}	1.11×10^{-3}
	2	2.68×10^{-12}	1.44×10^{-3}	2.17×10^{-3}	58.82	8.44×10^{-4}	2.04×10^{-3}
	4	2.77×10^{-11}	2.16×10^{-3}	1.29×10^{-3}	25.49	1.81×10^{-3}	2.51×10^{-3}
	10	6.42×10^{-4}	2.14×10^{-3}	1.04×10^{-3}	13.73	1.86×10^{-3}	2.42×10^{-3}

Best: best fitness value; Mean: mean fitness value; Std: standard deviation of fitness values; SR: success rate; L95: lower bound of 95% confidence limit; U95: upper bound of 95% confidence limit

optimize the frequency transformed Griewank function.

With the increase of F, the frequency of $\cos\left(F\cdot\frac{x_i}{\sqrt{i}}\right)$ increases to F times, and the frequency of $\prod_{i=1}^{n}\cos\left(F\cdot\frac{x_i}{\sqrt{i}}\right)$ increases to F^n times. This means that the frequency of the number of local minima in each dimension increases to F times, and the total number of local minima increases to F^n times. The probability of jumping out of the local optima decreases. Therefore, SR decreases with the increase of F. The best fitness value obtained by MQHOA remains a high precision for every F. The mean fitness and standard deviation of fitness values stay at around 10^{-3}.

2. With the increase of dimension, the SR of MQHOA first falls and then rises.

When the dimension is five, MQHOA's SR is lower than that with dimension 2 or 10. The 2D Griewank function is simple and is easy to optimize. With the increase of dimension, the number of local minima of the frequency transformed Griewank function increases exponentially, and it is more difficult to optimize. When the dimension is more than five and continues to grow, the frequency transformed Griewank function becomes easier to optimize with the increase of dimension. This corresponds with the experimental results in Table 1.

Global search and local search are the two simultaneous operations of swarm optimization algorithms. The aim of global search and local search is exploration and exploitation of the search space. Global search performs in a wide area to increase the diversity of solutions. Local search repeatedly tries to move from the current sampling point to a nearby sampling point.

When optimizing the Griewank function by MQHOA, the optimization process can be divided into global search and local search. At the beginning of optimization, the sampling points are distributed in a wide area, and have a large probability to move to a farther area. This stage can be considered as the global search. As the iteration proceeds, the scale of MQHOA shrinks. The sampling points are generated locally on a small scale, which guarantees a detailed optimization of the possible optimal area. This stage is local search.

The transition from global search to local search can be evaluated by the standard deviation of current

candidate solutions (σ_s):

$$\sigma_s = \sqrt{\sum_{i=1}^{k}(x_i - \overline{x})^2}. \quad (17)$$

When σ_s is large, the distribution of current candidate solutions is more dispersed. The algorithm is in the global search phase. As the iteration proceeds, σ_s continues to decrease, which causes the scale of MQHOA to shrink. With the decrease of σ_s, MQHOA gradually transits from global search to local search.

For the Griewank function, the initial scale of MQHOA is 200 and $\sigma_s = 10^{-6}$. The current scale of MQHOA is cut to half in each scale adjustment operation. We can calculate that MQHOA has 27 scales in the optimization process. For a frequency transformed Griewank function, we count the number of iterations in each scale. Experiments are done on 2-, 5-, and 10-dimensional frequency transformed Griewank functions. F is separately set to 1, 4, and 10. Fig. 12 shows the number of iterations in each scale.

As can be seen from Fig. 12, each subgraph has three curves, which represent F=1, F=4, and F=10, separately. With the increase of the scale number, the scale of MQHOA is cut to half continuously. MQHOA gradually transits from global to local search. The rectangle in each subgraph shows this transformation process. We can see that with the increase of F, the transition from global search to local search delays. With the increase of F, the number and the distribution density of local minima increase. Local search is done in a smaller area. As a result, MQHOA transits from global to local search at a smaller scale. This corresponds with the characteristic of the frequency transformed Griewank function mentioned in Section 4.1. The three subgraphs show the same trend.

4.3.2 Optimizing the amplitude transformed Griewank function

As for an amplitude transformed Griewank function, the amplitude parameter A ranges from 0.001 to 5. We select the same dimensions as in Table 4. Experimental results of optimizing the amplitude transformed Griewank function are listed in Table 5.

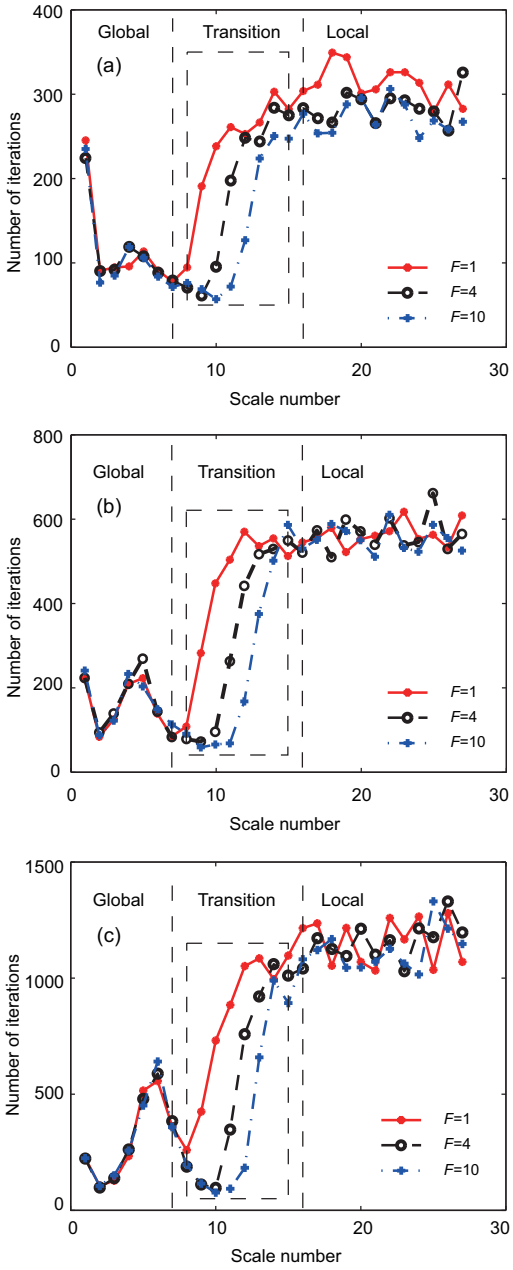

Fig. 12 Number of iterations in each scale (optimizing frequency transformed Griewank function with MQHOA): (a) 2D; (b) 5D; (c) 10D

From the data displayed in Table 5, we have the following two observations:

1. Similar to that in Section 4.3.1, with the increase of dimension, the SR of MQHOA first falls and then rises.

2. With the increase of A, the SR of MQHOA first falls and then rises.

When adjusting parameter A, the amplitude of the quadratic convex function changes accordingly.

Table 5 Best fitness value, mean fitness value, standard deviation of fitness values, and the success rate of MQHOA on the amplitude transformed Griewank function with amplitude ranging from 0.001 to 5

Dimension	A	Best	Mean	Std	SR (%)	L95	U95
$D=2$	0.001	1.99×10^{-9}	2.96×10^{-5}	3.10×10^{-5}	100.00	1.84×10^{-5}	3.54×10^{-5}
	0.01	5.62×10^{-13}	2.29×10^{-4}	2.11×10^{-4}	100.00	1.71×10^{-4}	2.87×10^{-4}
	0.1	9.79×10^{-13}	1.17×10^{-3}	1.37×10^{-3}	74.51	7.94×10^{-4}	1.55×10^{-3}
	0.2	1.65×10^{-13}	1.19×10^{-3}	1.38×10^{-3}	41.18	8.11×10^{-4}	1.57×10^{-3}
	0.5	5.14×10^{-12}	1.45×10^{-3}	1.82×10^{-3}	60.78	9.50×10^{-4}	1.95×10^{-4}
	1	3.82×10^{-13}	1.02×10^{-3}	2.57×10^{-3}	86.27	5.89×10^{-4}	1.45×10^{-3}
	2	5.02×10^{-13}	2.90×10^{-4}	2.07×10^{-3}	98.04	-2.78×10^{-4}	8.58×10^{-4}
	5	4.05×10^{-12}	5.32×10^{-9}	1.39×10^{-8}	100.00	1.51×10^{-9}	9.13×10^{-9}
$D=5$	0.001	1.97×10^{-5}	1.24×10^{-4}	6.16×10^{-5}	100.00	1.07×10^{-4}	1.41×10^{-4}
	0.01	1.48×10^{-4}	1.06×10^{-3}	7.36×10^{-4}	54.90	8.58×10^{-4}	1.26×10^{-3}
	0.1	7.40×10^{-4}	4.52×10^{-3}	2.52×10^{-3}	1.96	3.83×10^{-3}	5.21×10^{-3}
	0.2	1.71×10^{-13}	4.80×10^{-3}	3.24×10^{-3}	5.88	3.91×10^{-3}	5.69×10^{-3}
	0.5	5.98×10^{-14}	5.78×10^{-3}	4.57×10^{-3}	21.57	4.53×10^{-3}	7.03×10^{-3}
	1	1.31×10^{-13}	5.36×10^{-3}	5.23×10^{-3}	45.10	3.92×10^{-3}	6.80×10^{-3}
	2	4.74×10^{-14}	2.80×10^{-3}	6.70×10^{-3}	84.31	9.60×10^{-4}	4.64×10^{-3}
	5	3.01×10^{-14}	5.45×10^{-9}	1.28×10^{-8}	100.00	1.94×10^{-9}	8.96×10^{-9}
$D=10$	0.001	7.16×10^{-5}	2.90×10^{-4}	1.35×10^{-4}	100.00	2.53×10^{-4}	3.27×10^{-4}
	0.01	3.70×10^{-4}	1.45×10^{-3}	8.26×10^{-4}	35.29	1.22×10^{-3}	1.68×10^{-3}
	0.1	3.17×10^{-12}	2.60×10^{-3}	1.83×10^{-3}	23.53	2.10×10^{-3}	3.10×10^{-3}
	0.2	7.26×10^{-13}	1.88×10^{-3}	1.98×10^{-3}	37.25	1.34×10^{-3}	2.42×10^{-3}
	0.5	3.35×10^{-13}	6.04×10^{-4}	1.55×10^{-3}	86.27	1.79×10^{-4}	1.03×10^{-3}
	1	4.04×10^{-13}	1.45×10^{-4}	1.04×10^{-3}	98.04	-1.40×10^{-4}	4.30×10^{-4}
	2	1.92×10^{-13}	7.66×10^{-10}	2.62×10^{-9}	100.00	4.70×10^{-11}	1.48×10^{-9}
	5	6.49×10^{-13}	1.45×10^{-9}	2.02×10^{-9}	100.00	2.12×10^{-10}	1.32×10^{-9}

Best: best fitness value; Mean: mean fitness value; Std: standard deviation of fitness values; SR: success rate; L95: lower bound of 95% confidence limit; U95: upper bound of 95% confidence limit

When A is 0.001, the impact of $g_1(x)$ is very small and can be neglected. The problem of optimizing $f_1(x)$ is approximate to optimizing the oscillatory nonconvex function $h_1(x)+1$, which has many global minima. Theoretically speaking, the smaller A is, the more difficult it is to optimize $f_1(x)$, which does not match the experimental data. This is caused by the convergence accuracy of MQHOA. The local optimal solution of $f_1(x)$ can also achieve the convergence accuracy requirement of MQHOA, which meets MQHOA's termination condition. So, the SR does not reflect the real situations.

This situation changes when A is in the range [0.1, 0.2]. When A increases to 0.1, the local optimal solution of $f_1(x)$ cannot meet MQHOA's termination condition. When A is 0.1, the values of local minima near the global optimal solution and global minimum are similar. It is quite difficult to obtain the global optimal solution. In this situation, MQHOA easily falls into a local optimum while satisfying the convergence accuracy.

When A continues to increase, the impact of $h_1(x)$ decreases. The problem of optimizing $f_1(x)$ can gradually be approximate to optimizing the quadratic convex function $g_1(x)$. So, the SR of MQHOA gradually increases to 100%.

For the amplitude transformed Griewank function, we count the number of iterations in each scale. Experiments are done on 2-, 5-, and 10-dimensional amplitude transformed Griewank functions. A is separately set to 1, 5 and 10. Fig. 13 shows the number of iterations in each scale.

As can be seen from Fig. 13, three curves coincide well when the scale number is small. That is to say, as for the Griewank function, amplitude transformation has little effect on global search. When the scale number is larger than 15, we can find that the number of iterations increases with the increase of A. That is to say, with the increase of the amplitude parameter, local search becomes more difficult. This corresponds with the characteristic of the amplitude transformed Griewank function mentioned in Section 4.1. The three subfigures show the same trend.

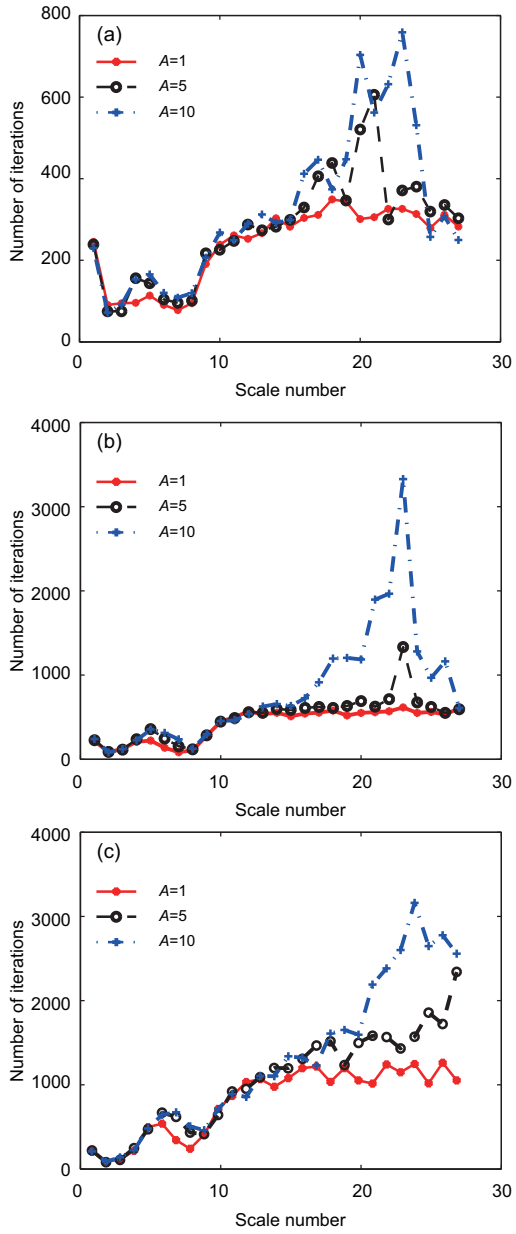

Fig. 13 Number of iterations in each scale (optimizing amplitude transformed Griewank function with MQHOA): (a) 2D; (b) 5D; (c) 10D

5 Conclusions

In this paper the characteristic of optimizing the Griewank function by swarm intelligence algorithms has been researched. The Griewank function first becomes more difficult and then becomes easier to optimize with the increase of dimension. Structural, mathematical, and quantum analyses have been conducted on the Griewank function to explain this phenomenon. MQHOA has been used to optimize the generalized Griewank function and it performs well. In the future research, we will continue to investigate the generalized model of the Griewank function and enrich the applications of MQHOA such as heat production optimization (Książek et al., 2017; Woźniak et al., 2018).

Compliance with ethics guidelines

Yan HUANG, Jian-ping LI, and Peng WANG declare that they have no conflict of interest.

References

Akay B, Karaboga D, 2012. A modified artificial bee colony algorithm for real-parameter optimization. *Inform Sci*, 192(1):120-142. https://doi.org/10.1016/j.ins.2010.07.015

Akbari R, Mohammadi A, Ziarati K, 2010. A novel bee swarm optimization algorithm for numerical function optimization. *Commun Nonl Sci Numer Simul*, 15(10):3142-3155. https://doi.org/10.1016/j.cnsns.2009.11.003

Chen DB, Zhao CX, 2009. Particle swarm optimization with adaptive population size and its application. *Appl Soft Comput*, 9(1):39-48. https://doi.org/10.1016/j.asoc.2008.03.001

Cho H, Olivera F, Guikema SD, 2008. A derivation of the number of minima of the Griewank function. *Appl Math Comput*, 204(2):694-701. https://doi.org/10.1016/j.amc.2008.07.009

Gao WF, Liu SY, Huang LL, 2012. A global best artificial bee colony algorithm for global optimization. *J Comput Appl Math*, 236(11):2741-2753. https://doi.org/10.1016/j.cam.2012.01.013

Griewank AO, 1981. Generalized descent for global optimization. *J Optim Theory Appl*, 34(1):11-39. https://doi.org/10.1007/BF00933356

Karaboga D, Basturk B, 2007. A powerful and efficient algorithm for numerical function optimization: artificial bee colony (ABC) algorithm. *J Glob Optim*, 39(3):459-471. https://doi.org/10.1007/s10898-007-9149-x

Kirkpatrick S, Gelatt CDJr, Vecchi MP, 1983. Optimization by simulated annealing. *Science*, 220(4598):671-680. https://doi.org/10.1126/science.220.4598.671

Książek K, Połap D, Woźniak M, et al., 2017. Radiation heat transfer optimization by the use of modified ant lion optimizer. Proc IEEE Symp Series on Computational Intelligence. https://doi.org/10.1109/SSCI.2017.8280853

Liang JJ, Qu BY, Suganthan PN, 2013. Problem Definitions and Evaluation Criteria for the CEC 2014 Special Session and Competition on Single Objective Real-parameter Numerical Optimization. Technical Report No. 201311, Zhengzhou University, Zhengzhou, China.

Locatelli M, 2003. A note on the Griewank test function. *J Glob Optim*, 25(2):169-174. https://doi.org/10.1023/A:1021956306041

Muthukrishnan S, Albash T, Lidar DA, 2016. Tunneling and speedup in quantum optimization for permutation-symmetric problems. *Phys Rev X*, 6(3):031010. https://doi.org/10.1103/PhysRevX.6.031010

Oftadeh R, Mahjoob MJ, Shariatpanahi M, 2010. A novel meta-heuristic optimization algorithm inspired by group hunting of animals: hunting search. *Comput Math Appl*, 60(7):2087-2098. https://doi.org/10.1016/j.camwa.2010.07.049

Połap D, Woźniak M, 2017. Polar bear optimization algorithm: meta-heuristic with fast population movement and dynamic birth and death mechanism. *Symmetry*, 9(10), Article 203. https://doi.org/10.3390/sym9100203

Qin AK, Huang VL, Suganthan PN, 2009. Differential evolution algorithm with strategy adaptation for global numerical optimization. *IEEE Trans Evol Comput*, 13(2):398-417. https://doi.org/10.1109/TEVC.2008.927706

Rao RV, Savsani VJ, Vakharia DP, 2012. Teaching-learning-based optimization: an optimization method for continuous non-linear large scale problems. *Inform Sci*, 183(1):1-15. https://doi.org/10.1016/j.ins.2011.08.006

Shi Y, Eberhart RC, 1999. Empirical study of particle swarm optimization. Proc Congress on Evolutionary Computation, p.1945-1950. https://doi.org/10.1109/CEC.1999.785511

Tan Y, Zhu YC, 2010. Fireworks algorithm for optimization. Proc 1st Int Conf on Advances in Swarm Intelligence, p.355-364. https://doi.org/10.1007/978-3-642-13495-1_44

Wang P, Huang Y, Ren C, et al., 2013. Multi-scale quantum harmonic oscillator for high-dimensional function global optimization algorithm. *Acta Electron Sin*, 41(12):2468-2473 (in Chinese). https://doi.org/10.3969/j.issn.0372-2112.2013.12.023

Wang P, Huang Y, An JX, et al., 2016. Performance analysis of multi-scale quantum harmonic oscillator global optimization algorithm in combinatorial optimization problems. *J Univ Electron Sci Technol China*, 45(3):469-474 (in Chinese). https://doi.org/10.3969/j.issn.1001-0548.2016.02.027

Wang P, Cheng K, Huang Y, et al., 2018a. Multiscale quantum harmonic oscillator algorithm for multimodal optimization. *Comput Intell Neurosci*, 2018:8430175. https://doi.org/10.1155/2018/8430175

Wang P, Ye XG, Li B, et al., 2018b. Multi-scale quantum harmonic oscillator algorithm for global numerical optimization. *Appl Soft Comput*, 69:655-670. https://doi.org/10.1016/j.asoc.2018.05.005

Wang P, Li B, Jin J, et al., 2018c. Multi-scale quantum harmonic oscillator algorithm with individual stabilization strategy. Proc 9th Int Conf on Swarm Intelligence, p.624-633. https://doi.org/10.1007/978-3-319-93815-8_59

Wang PP, Chen DS, 1996. Continuous optimization by a variant of simulated annealing. *Comput Optim Appl*, 6(1):59-71. https://doi.org/10.1007/BF00248009

Woźniak M, Książek K, Marciniec J, et al., 2018. Heat production optimization using bio-inspired algorithms. *Eng Appl Artif Intell*, 76:185-201. https://doi.org/10.1016/j.engappai.2018.09.003

Zambrano-Bigiarini M, Clerc M, Rojas R, 2013. Standard particle swarm optimisation 2011 at CEC-2013: a baseline for future PSO improvements. Proc IEEE Congress on Evolutionary Computation, p.2337-2344. https://doi.org/10.1109/CEC.2013.6557848

Zhang LP, Yu HJ, Hu SX, 2003. A new approach to improve particle swarm optimization. Proc Genetic and Evolutionary Computation Conf, p.134-139. https://doi.org/10.1007/3-540-45105-6_12

Zhou DD, Shi YH, Cheng S, 2012. Brain storm optimization algorithm with modified step-size and individual generation. Proc 3rd Int Conf on Advances in Swarm Intelligence, p.243-252. https://doi.org/10.1007/978-3-642-30976-2_29