Quantum-Behaved Particle Swarm Optimization: Analysis of Individual Particle Behavior and Parameter Selection

Jun Sun sunjun_wx@hotmail.com
Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi, Jiangsu 214122, China

Wei Fang wxfangwei@hotmail.com
Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi, Jiangsu 214122, China

Xiaojun Wu wu_xiaojun@yahoo.com.cn
Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi, Jiangsu 214122, China

Vasile Palade vasile.palade@cs.ox.ac.uk
Department of Computer Science, University of Oxford, Oxford, OX1 3QD, United Kingdom

Wenbo Xu xwb@jiangnan.edu.cn
Key Laboratory of Advanced Process Control for Light Industry (Ministry of Education), Jiangnan University, Wuxi, Jiangsu 214122, China

Abstract

Quantum-behaved particle swarm optimization (QPSO), motivated by concepts from quantum mechanics and particle swarm optimization (PSO), is a probabilistic optimization algorithm belonging to the bare-bones PSO family. Although it has been shown to perform well in finding the optimal solutions for many optimization problems, there has so far been little analysis on how it works in detail. This paper presents a comprehensive analysis of the QPSO algorithm. In the theoretical analysis, we analyze the behavior of a single particle in QPSO in terms of probability measure. Since the particle's behavior is influenced by the contraction-expansion (CE) coefficient, which is the most important parameter of the algorithm, the goal of the theoretical analysis is to find out the upper bound of the CE coefficient, within which the value of the CE coefficient selected can guarantee the convergence or boundedness of the particle's position. In the experimental analysis, the theoretical results are first validated by stochastic simulations for the particle's behavior. Then, based on the derived upper bound of the CE coefficient, we perform empirical studies on a suite of well-known benchmark functions to show how to control and select the value of the CE coefficient, in order to obtain generally good algorithmic performance in real world applications. Finally, a further performance comparison between QPSO and other variants of PSO on the benchmarks is made to show the efficiency of the QPSO algorithm with the proposed parameter control and selection methods.

Keywords

Particle swarm optimization, quantum behavior, probabilistic boundedness, convergence, parameter selection.

1 Introduction

The particle swarm optimization (PSO) algorithm is a population-based optimization technique, originally developed by Kennedy and Eberhart in 1995. It was motivated by the social behavior of bird flocking or fish schooling and shares many similarities with evolutionary computation techniques. A PSO system is initialized with a population of random solutions and searches for optima by updating generations. However, unlike evolutionary algorithms, PSO has no evolution operators such as crossover and mutation. In PSO, the potential solutions, called particles, fly through the problem space by following their own experiences and the current best particles. It has been shown that the PSO algorithm is comparable in performance with and may be considered as an alternative method to evolutionary algorithms (Angeline, 1998a).

During the last decade, PSO gained increasing popularity since it can get better results for optimization problems in a faster and cheaper way compared with other methods, but has fewer parameters to be adjusted. In order to investigate in detail the mechanism of PSO, a lot of theoretical analyses have been done on the algorithm (Kennedy, 1998; Ozcan and Mohan, 1999; Clerc and Kennedy, 2002; van den Bergh, 2002; Shi and Eberhart, 1998b; Trelea, 2003; Emara and Fattah, 2004; Gavi and Passino, 2003; Kadirkamanathan et al., 2006; Jiang et al., 2007; Poli, 2008). These theoretical analyses were focused on the behavior of the individual particle, which is essential to the understanding of the search mechanism of the algorithm and to parameter selection. For example, Kennedy (1998) carried out an analysis of simplified particle behavior and showed the different trajectories of particles for a range of design choices aiming to gain some insights into the behavior of particles through simulations, and Clerc and Kennedy (2002) undertook the first formal analysis of the particle trajectory and of the stability properties of the algorithm.

Besides the theoretical analyses, there has been a considerable amount of work done in developing the original version of PSO through empirical simulations. In order to accelerate the convergence of the particle, Shi and Eberhart (1998a) introduced the concept of an inertia weight into the original PSO, and Clerc (1999) proposed an alternative version of PSO incorporating a parameter known as the constriction factor which should replace the restriction on velocities. Some researchers employed the operations of other evolutionary algorithms in PSO to enhance its performance, the most important work being done by Angeline (1998b). Another general form of particle swarm, referred to as the lbest model, was first proposed by Eberhart and Kennedy (1995), and then was studied in depth by many other researchers in order to find other topologies to improve the performance of PSO (Suganthan, 1999; Kennedy, 1999, 2002; Mendes et al., 2004; van den Bergh and Engelbrecht, 2004; Janson and Middendorf, 2005; Liang and Suganthan, 2005; Mohais et al., 2005; Parrott and Li, 2006). Bratton and Kennedy defined a standard PSO, which is an extension of the original PSO algorithm while taking into account the previous developments that can improve the performance of the algorithm (Bratton and Kennedy, 2007).

Some researchers have attempted to experiment with various ways to simulate the particle trajectory by directly sampling, using a random number generator, from a distribution of some theoretical interest, and thus have proposed many probabilistic

PSO algorithms (Kennedy, 2003, 2004, 2006; Secrest and Lamont, 2003; Krohling, 2004; Sun, Feng, et al., 2004; Sun, Xu, et al., 2004; Sun et al., 2005; Krohling and Coelho, 2006; Richer and Blackwell, 2006), the most popular being the bare-bones PSO (BBPSO) family. In BBPSO, each particle has no velocity vector and its new position is sampled around a supposed good one according to a probability distribution. In the original BBPSO, this distribution is Gaussian (Kennedy, 2003). A moment analysis was proposed to determine the characteristics of the sampling distribution of BBPSO and other PSO algorithms (Poli, 2008).

The focus of this paper is on a probabilistic algorithm, quantum-behaved particle swarm optimization (QPSO), which was proposed by Sun, Feng et al. (2004). The inspiration of QPSO came from quantum mechanics and the trajectory analysis of PSO (Clerc and Kennedy, 2002). The trajectory analysis showed that each particle in PSO oscillates around and converges to its local attractor, or to put it in other words, each particle is in a bound state. In QPSO, the particle is assumed to have quantum behavior and to be in a bound state, and is further assumed to be attracted by a quantum potential well centered on its local attractor, thus having a new stochastic update equation for its position (Sun, Feng et al., 2004). Later, a global point known as the mean best position was introduced into the algorithm in order to enhance the global search ability of the QPSO algorithm (Sun, Xu et al., 2004; Sun et al., 2005).

The QPSO algorithm essentially belongs to the BBPSO family, but samples the new position with a double exponential distribution. Besides, its update equation uses an adaptive strategy and has fewer parameters to be adjusted, leading to a good performance of the algorithm as an overall result. The QPSO algorithm has aroused the interest of many researchers from different communities. It has been shown to successfully solve a wide range of continuous optimization problems. Among these applications, it has been used to tackle the problems of constrained optimization (Sun et al., 2007), multiobjective optimization (Omkara et al., 2009), neural network training (Li et al., 2007), electromagnetic design (Mikki and Kishk, 2006; Coelho and Alotto, 2008), semiconductor design (Sabata et al., 2009), clustering (Sun et al., 2006), system identification (Gao, 2008), engineering design (Coelho, 2008, 2010), image processing (Lei and Fu, 2008), power systems (Coelho and Mariani, 2008; Sun et al., 2009; Zhang, 2010; Sun and Lu, 2010), bioinformatics (Chen et al., 2008; Cai et al., 2008), to name only a few.

In addition to the applications, many efficient strategies have been proposed to improve the performance of QPSO. For example, Liu et al. (2005) introduced the mutation operation into QPSO to improve the search ability of the algorithm. Wang and Zhou (2007) proposed a local QPSO (LQPSO) as a generalized local search operator and incorporated LQPSO into a main QPSO algorithm, which leads to a hybrid QPSO scheme QPSO-LQPSO, with enhanced searching qualities. In Coelho (2008), it was shown that the chaotic mutation operation could diversify the population of QPSO and thus improve the performance of the algorithm. Pant et al. developed a new variant of QPSO, which used an interpolation-based recombination operator for generating a new solution vector in the search space (Pant et al., 2008). They also proposed a new mutation operator called the Sobal mutation to improve the performance of the QPSO algorithm (Pant et al., 2009). Xi et al. (2008) accelerated the convergence speed of QPSO by introducing a weighted mean best position into the algorithm. Huang et al. (2009) proposed an improved QPSO, employing a selection operation on the particles in order to filter the particle swarm and accelerate its convergence.

While empirical evidence has shown that the QPSO algorithm works well, thus far there has been little insight into how it works. In this paper, we make a comprehensive

analysis of individual particle behavior for QPSO from the perspective of probability measure and show how to select the contraction-expansion (CE) coefficient, which is the most important algorithmic parameter. To achieve this goal, the work presented in the paper is mainly divided into the following two parts.

The first part includes the theoretical analysis of the algorithm. Since the behavior of the individual particle in QPSO exerts great influence on the convergence of the particle swarm and in turn on the convergence of the algorithm, we analyze the individual particle behavior for two different versions of QPSO. In QPSO, the CE coefficient controls the behavior of the individual particle, just as the inertia weight and acceleration coefficients influence the behavior of the particle in PSO. Therefore, this part of the work comes down to deriving the upper bound of the value of the CE coefficient that guarantees the convergence or boundedness of the particle.

The upper bound of the CE coefficient only provides the condition that leads to the convergence of the particle swarm, not the condition that ensures the efficiency of QPSO in practical applications. Thus, the second part of the paper involves empirical studies on how to select the value of the CE coefficient within the upper bound to lead the QPSO algorithm to good performance in general. With regard to this, we firstly perform stochastic simulations in order to visualize the influence of the value of the CE coefficient on particle convergence speed or the expected range of particle random vibration, as well as to verify the derived theoretical results of the parameter's upper bound. Secondly, we execute the QPSO algorithm with two parameter control methods on some well-known benchmark functions to find the parameters' values resulting in good solutions in general. Finally, the QPSO algorithms with the parameters' values found for both parameter control methods are tested on a suite of benchmark functions proposed by Suganthan at CEC 2005 (Suganthan et al., 2005), and the performances are compared with those of other PSO variants.

The rest of the paper is organized as follows. In Section 2, the basic principles of QPSO are introduced. Section 3 presents a theoretical analysis of individual particle behavior. Section 4 provides the experimental analysis of the algorithm. Some concluding remarks are given in the last section.

2 Quantum-Behaved Particle Swarm Optimization

In the PSO with M individuals, each individual is treated as a volumeless particle in an N-dimensional space, with the current position vector and the velocity vector of particle i ($1 \le i \le M$) at the nth iteration represented as $X_{i,n} = (X_{i,n}^1, X_{i,n}^2, \ldots, X_{i,n}^N)$ and $V_{i,n} = (V_{i,n}^1, V_{i,n}^2, \ldots, V_{i,n}^N)$, respectively. The particle moves according to the following equations:

$$V_{i,n+1}^j = V_{i,n}^j + c_1 r_{i,n}^j (P_{i,n}^j - X_{i,n}^j) + c_2 R_{i,n}^j (G_n^j - X_{i,n}^j), \tag{1}$$

$$X_{i,n+1}^j = X_{i,n}^j + V_{i,n+1}^j, \tag{2}$$

for $j = 1, 2, \ldots, N$, where c_1 and c_2 are known as the acceleration coefficients. Vector $P_{i,n} = (P_{i,n}^1, P_{i,n}^2, \ldots, P_{i,n}^N)$ is the best previous position (the position giving the best objective function value or fitness value) of particle i, called the personal best (pbest) position, and vector $G_n = (G_n^1, G_n^2, \ldots, G_n^N)$ is the position of the best particle among all the particles in the population, called the global best (gbest) position. Without loss of

generality, we consider the following minimization problem:

$$\text{Minimize } f(X), \quad s.t. \quad X \in S \subseteq R^N, \tag{3}$$

where $f(X)$ is an objective function continuous almost everywhere and S is the feasible space. Accordingly, $P_{i,n}$ can be updated by

$$P_{i,n} = \begin{cases} X_{i,n} & \text{if} \quad f(X_{i,n}) < f(P_{i,n}\text{-}1) \\ P_{i,n}\text{-}1 & \text{otherwise} \end{cases}, \tag{4}$$

and G_n can be found by $G_n = P_{g,n}$, where $g = \arg\min_{1 \le i \le M}\{f(P_{i,n})\}$. The parameters $r_{i,n}^j$ and $R_{i,n}^j$, varying with n for each i and j, are two different random numbers distributed uniformly on $(0, 1)$, which is denoted by $r_{i,n}^j$, $R_{i,n}^j \sim U(0, 1)$. Generally, the value of $V_{i,n}^j$ is restricted in the interval $[-V_{\max}, V_{\max}]$.

Trajectory analysis (Clerc and Kennedy, 2002) demonstrated the fact that convergence of the PSO algorithm may be achieved if each particle converges to its local attractor, $p_{i,n} = (p_{i,n}^1, p_{i,n}^2, \ldots, p_{i,n}^N)$, defined at the coordinates

$$p_{i,n}^j = \frac{c_1 r_{i,n}^j P_{i,n}^j + c_2 R_{i,n}^j G_n^j}{c_1 r_{i,n}^j + c_2 R_{i,n}^j}, \tag{5}$$

or

$$p_{i,n}^j = \varphi_{i,n}^j P_{i,n}^j + (1 - \varphi_{i,n}^j) G_n^j, \tag{6}$$

for $1 \le j \le N$, where $\varphi_{i,n}^j = \frac{c_1 r_{i,n}^j}{c_1 r_{i,n}^j + c_2 R_{i,n}^j}$, with regard to the random numbers $r_{i,n}^j$ and $R_{i,n}^j$ in Equations (1) and (5). In PSO, the acceleration coefficients c_1 and c_2 are generally set to be equal, that is, $c_1 = c_2$, and thus $\varphi_{i,n}^j$ is a sequence of random numbers uniformly distributed on $(0, 1)$. As a result, Equation (6) can be restated as

$$p_{i,n}^j = \varphi_{i,n}^j P_{i,n}^j + (1 - \varphi_{i,n}^j) G_n^j, \quad \varphi_{i,n}^j \sim U(0, 1). \tag{7}$$

The above equation indicates that $p_{i,n}$, the stochastic attractor of particle i, lies in the hyper-rectangle with $P_{i,n}$ and G_n being the two ends of its diagonal so that it moves following $P_{i,n}$ and G_n. In fact, as the particles are converging to their own local attractors, their current position, personal best positions, local attractors, and the global best positions are all converging to one point, leading the PSO algorithm to convergence. From the point view of Newtonian dynamics, in the process of convergence, the particle moves around and careens toward $p_{i,n}$ with its kinetic energy (or velocity) declining to zero, like a returning satellite orbiting the earth. As such, the particle in PSO can be considered as the one flying in an attraction potential field centered at $p_{i,n}$ in the Newtonian space. It has to be in a bound state for the sake of avoiding explosion and guaranteeing convergence. If these conditions are generalized to the case in which the particle in PSO has quantum behavior moving in an N-dimensional Hilbert space, it is also indispensable that the particle moves in a quantum potential field to ensure the bound state. From the perspective of quantum mechanics, the bound state in the

quantum space, however, is entirely different from that in the Newtonian space, which may lead to a very different form of PSO. This is the motivation of the proposed QPSO algorithm which is described below (Sun, Feng et al., 2004).

In the quantum time-space framework, the quantum state of a particle is described by the wave function $\Psi(X,t)$ (Cohen-Tannoudji et al., 1997). In a three-dimensional space, the wave function $\Psi(X,t)$ of a particle satisfies the relation

$$|\Psi|^2\,dxdydz = Q\,dxdydz, \tag{8}$$

where $Q\,dxdydz$ is the probability that the particle will appear in the infinitesimal element about the point (x,y,z). In other words, $|\Psi|^2 = Q$ represents the probability density function satisfying

$$\int_{-\infty}^{+\infty} |\Psi|^2\,dxdydz = \int_{-\infty}^{+\infty} Q\,dxdydz = 1. \tag{9}$$

Equation (8) or (9) gives the statistical interpretation for the wave function. Generally, $\Psi(X,t)$ varies in time according to the following equation

$$i\hbar\frac{\partial}{\partial t}\Psi(X,t) = \hat{H}\Psi(X,t), \tag{10}$$

where $\hbar$ is Planck's constant and $\hat{H}$ is the Hamiltonian operator defined by

$$\hat{H} = -\frac{\hbar^2}{2m}\nabla^2 + V(X) \tag{11}$$

for a single particle of mass m in a potential field $V(X)$. Equation (10) is known as the time-dependent Schrödinger equation.

We assume that each single particle in QPSO is treated as a spin-less particle moving in an N-dimensional Hilbert space with a given energy, and thus its state is characterized by a wave function which only depends on the position of the particle. Inspired by the convergence analysis of the particle in PSO (Clerc and Kennedy, 2002), we further assume that, at the nth iteration, particle i flies in the N-dimensional Hilbert space with a δ potential well centered at $p_{i,n}^j$ on the jth dimension ($1 \le j \le N$). In order to facilitate the description, we consider a particle in an one-dimensional space firstly and denote the position of the particle as X and $p_{i,n}$ as p. With point p being the center of the potential well, the potential energy of the particle in the one-dimensional δ potential well is represented as

$$V(X) = -\gamma\delta(X-p) = -\gamma\delta(Y), \tag{12}$$

where $Y = X - p$ and γ is the intensity of the potential well. Hence, for this bound state problem, the Hamiltonian operator is

$$\hat{H} = -\frac{\hbar^2}{2m}\frac{d^2}{dY^2} - \gamma\delta(Y). \tag{13}$$

The particle's state is subject to the following stationary Schrödinger equation:

$$\frac{d^2\psi}{dY^2} + \frac{2m}{\hbar^2}[E + \gamma\delta(Y)]\psi = 0, \tag{14}$$

where E is the energy of the particle and ψ is the wave function of the particle which only depends on its position. As presented in Theorem 1 below, we can obtain the normalized wave function of the particle in the bound state by solving the above equation with the bound condition: $\psi \to 0$, as $|Y| \to +\infty$.

THEOREM 1: *For the particle that moves in the one-dimensional δ potential well formulated by Equation (12), its normalized wave function in the bound state is given by*

$$\psi(Y) = \frac{1}{\sqrt{L}} e^{-\frac{|Y|}{L}}, \tag{15}$$

where $L = \frac{\hbar^2}{m\gamma}$.

PROOF: Integrating Equation (14) with respect to Y from $-\varepsilon$ to ε and taking $\varepsilon \to 0^+$ leads to

$$\psi'(0^+) - \psi'(0^-) = -\frac{2m\gamma}{\hbar^2}\psi(0). \tag{16}$$

For $Y \neq 0$, Equation (14) can be written as

$$\frac{d^2\psi}{dY^2} - \beta^2\psi = 0, \tag{17}$$

where $\beta = \sqrt{\frac{-2mE}{\hbar}}$ $(E < 0)$. To satisfy the bound condition

$$\psi \to 0, \quad \text{as} \quad |Y| \to +\infty, \tag{18}$$

the solution of Equation (17) must be

$$\psi(Y) \propto e^{-\beta|Y|}, \quad (Y \neq 0). \tag{19}$$

Note that it can be proved that only an even wave function satisfies the bound condition in Equation (18), and then the solution of Equation (17) can be written as

$$\psi(Y) = \begin{cases} Ae^{-\beta Y} & Y > 0 \\ Ae^{\beta Y} & Y < 0 \end{cases}, \tag{20}$$

where A is the normalization constant. According to Equation (16), we have

$$-2A\beta = -\frac{2m\gamma}{\hbar^2}A.$$

Thus

$$\beta = \frac{m\gamma}{\hbar^2}, \tag{21}$$

and

$$E = E_0 = -\frac{\hbar^2\beta^2}{2m} = -\frac{m\gamma^2}{2\hbar^2}. \tag{22}$$

The function $\psi(Y)$ satisfies normalization condition

$$\int_{-\infty}^{+\infty} |\psi(Y)|^2 \, dY = \frac{|A|^2}{\beta} = 1, \tag{23}$$

leading to $|A| = \sqrt{\beta}$. $L = \frac{1}{\beta} = \frac{\hbar^2}{m\gamma}$ is the characteristic length of the δ potential well. Inserting $|A| = \sqrt{\beta} = \frac{1}{L}$ and $\beta = \frac{1}{L}$ into Equation (20), the normalized wave function can then be written as

$$\psi(Y) = \frac{1}{\sqrt{L}} e^{-\frac{|Y|}{L}}. \tag{24}$$

This completes the proof of the theorem. □

In terms of the statistical interpretation of the wave function, the probability density function of Y is given by

$$Q(Y) = |\psi(Y)|^2 = \frac{1}{L} e^{\frac{-2|Y|}{L}}, \tag{25}$$

and the corresponding probability distribution function is

$$F(Y) = 1 - e^{\frac{-2|Y|}{L}}. \tag{26}$$

With the given probability distribution function, we can measure the position of the particle using Monte Carlo inverse transformation. This technique is described in the proof of Theorem 2. Such a process of measuring the particle's position in quantum mechanics is essentially achieved by collapsing the quantum state to the classical state.

THEOREM 2: *If a particle moves in the bound state in the one-dimensional δ potential well as described by Equation (12), its position can be determined by using the stochastic equation*

$$X = p \pm \frac{L}{2} \ln\left(\frac{1}{u}\right), \tag{27}$$

where u is a random number uniformly distributed on $(0, 1)$, *that is,* $u \sim U(0, 1)$.

PROOF: Let v be a random number uniformly distributed on $(0, 1)$, that is

$$v \sim U(0, 1).$$

Substituting v for $F(Y)$ in Equation (26) and following Monte Carlo inverse transformation, we obtain

$$1 - v = e^{\frac{-2|Y|}{L}}. \tag{28}$$

Since $1 - v \sim U(0, 1)$, putting $u = 1 - v$ leads to $u \sim U(0, 1)$. Thus Equation (28) can be written as

$$u = e^{\frac{-2|Y|}{L}}, \tag{29}$$

from which we can immediately obtain

$$Y = \pm\frac{L}{2}\ln(1/u).$$

Because $Y = X - p$, we have

$$X = p \pm \frac{L}{2}\ln(1/u), \quad u \sim U(0, 1). \tag{30}$$

This completes the proof of the theorem. □

Now we generalize Equation (27) to the case in the N-dimensional Hilbert space where each dimension of the particle's position is bounded in a δ potential well and updated independently. Considering that the particle's position, its local attractor, the characteristic length of the δ potential well and the random variable u develop with the iteration number n, we can use the following equation to measure the jth $(1 \le j \le N)$ component of the position of particle i $(1 \le i \le M)$ at the $(n + 1)$th iteration.

$$X^j_{i,n+1} = p^j_{i,n} \pm \frac{L^j_{i,n}}{2}\ln(1/u^j_{i,n+1}), \tag{31}$$

where $u^j_{i,n+1}$ is a sequence of random numbers uniformly distributed on $(0, 1)$, varying with n for each i and j.

The value of $L^j_{i,n}$ in Equation (31) can be determined by either of the following two equations (Sun, Feng et al., 2004; Sun, Xu et al., 2004):

$$L^j_{i,n} = 2\alpha|X^j_{i,n} - p^j_{i,n}|, \tag{32}$$

or

$$L^j_{i,n} = 2\alpha|X^j_{i,n} - C^j_n|, \tag{33}$$

where $C_n = (C^1_n, C^2_n, \ldots, C^N_n)$ is known as the mean best (mbest) position which is defined by the average of the pbest positions of all particles, that is, $C^j_n = \frac{1}{M}\sum_{i=1}^{M} P^j_{i,n}$ $(1 \le$

$j \leq N$) . Therefore, the position of the particle is updated by using either of the following two equations:

$$X_{i,n+1}^{j} = p_{i,n}^{j} \pm \alpha |X_{i,n}^{j} - p_{i,n}^{j}| \ln(1/u_{i,n+1}^{j}), \tag{34}$$

or

$$X_{i,n+1}^{j} = p_{i,n}^{j} \pm \alpha |X_{i,n}^{j} - C_{n}^{j}| \ln(1/u_{i,n+1}^{j}). \tag{35}$$

The parameter α in Equations (32)–(35) is a positive real number, the CE coefficient, which can be adjusted to balance the local and global search of the algorithm during the search process. The PSO with Equation (34) or (35) is the QPSO. To distinguish them, we denote the QPSO with Equation (34) as QPSO-Type 1 and that with Equation (35) as QPSO-Type 2.

The QPSO algorithm starts with the initialization of the particles' current positions and their pbest positions (setting $P_{i,0} = X_{i,0}$), followed by the iterative update of the particle swarm. At each iteration of the procedure, the mbest position of the particle swarm is computed (for QPSO-Type 2) and the current position of each particle is updated according to Equation (34) or (35) with the coordinates of its local attractor evaluated by Equation (7). Before each particle updates its current position, its fitness value is evaluated together with an update of its pbest position and the current gbest position. In Equation (34) or (35), the probability of using either the $+$ operation or the $-$ operation is equal to 0.5. The iterative process continues until the termination condition is met.

The procedure of the QPSO algorithm is outlined in Algorithm 1. Note that rand$i(\cdot)$, $i = 1, 2, 3$, is used to denote the random numbers that are separately generated and uniformly distributed on (0, 1).

3 Theoretical Analysis

The weighting of the CE coefficient α in the QPSO algorithm may result in a kind of explosion as the position coordinates careen toward infinity. This section demonstrates that properly selected α can prevent explosion, and further, this coefficient can induce the particle to converge to its local attractor (in QPSO-Type 1) or to be probabilistic bounded (in QPSO-Type 2).

As in the PSO algorithm, an important source of the swarm's search capability is the interactions among particles as they react to one another's findings. However, the theoretical analysis in this paper does not involve the analysis of interparticle effects but is focused on the stochastic movements of single particles.

3.1 Preliminaries

This section provides a brief introduction to mathematical preliminaries on probability measure and sequences of random variables, which are essential to the theoretical analysis of a single particle's behavior. To study this knowledge in depth, one may refer to the related references such as Dudley (2003) and Shiryayev (1984).

Algorithm 1 The QPSO algorithm

begin
  Initialize the current positions and the pbest positions of all the particles;
  Set $n = 0$;
  while *the termination condition is not met* **do**
    Compute the mean best position C_n (for QPSO-Type 2);
    Select a suitable value for α;
    for $i = 1$ *to* M **do**
      Evaluate the objective function value $f(X_{i,n})$;
      Update $P_{i,n}$ and G_n;
      for $j = 1$ *to* N **do**
        $\varphi_{i,n}^j = \text{rand1}(\cdot)$;
        $p_{i,n}^j = \varphi_{i,n}^j P_{i,n}^j + (1 - \varphi_{i,n}^j) G_n^j$;
        $u_{i,n+1} = \text{rand2}(\cdot)$;
        if $\text{rand3}(\cdot) < 0.5$ **then**
          $X_{i,n+1}^j = p_{i,n}^j + \alpha |X_{i,n}^j - p_{i,n}^j| \ln(1/u_{i,n+1}^j)$ (for QPSO-Type 1);
          (or $X_{i,n+1}^j = p_{i,n}^j + \alpha |X_{i,n}^j - C_n^j| \ln(1/u_{i,n+1}^j)$ (for QPSO-Type 2));
        else
          $X_{i,n+1}^j = p_{i,n}^j - \alpha |X_{i,n}^j - p_{i,n}^j| \ln(1/u_{i,n+1}^j)$ (for QPSO-Type 1);
          (or $X_{i,n+1}^j = p_{i,n}^j - \alpha |X_{i,n}^j - C_n^j| \ln(1/u_{i,n+1}^j)$ (for QPSO-Type 2));
        end
      end
    end
    Set $n = n + 1$;
  end
end

DEFINITION 1: *The set Ω containing all possible outcomes of a random experiment is called the space of elementary events or the sample space. Each outcome or a point ω in Ω is known as an elementary event or a sample point.*

DEFINITION 2: *The space Ω together with a σ-algebra $\mathcal{F}$ of its subsets is a measurable space which is denoted by $(\Omega, \mathcal{F})$.*

DEFINITION 3: *An ordered triple $(\Omega, \mathcal{F}, P)$ where*
(a) *Ω is a set of points ω,*
(b) *$\mathcal{F}$ is a σ-algebra of subsets of Ω,*
(c) *P is a probability measure on $\mathcal{F}$,*
is called a probabilistic model or a probability space. Here Ω is the sample space or the space of elementary events, each set A in $\mathcal{F}$ is known as an event, and $P(A)$ is the probability of the event A.

Let $R = (-\infty, +\infty)$ be the real line. $(R, \mathcal{B}(R))$ or $(R, \mathcal{B})$ is a measure space. $\mathcal{B}(R)$ or $\mathcal{B}$ is called the Borel algebra of subsets.

DEFINITION 4: *A real function $\xi = \xi(\omega)$ defined on $(\Omega, \mathcal{F}, P)$ is an $\mathcal{F}$-measurable function or a random variable, if $\{\omega : \xi(\omega) \in B) \in \mathcal{F}$ for every $B \in \mathcal{B}(R)$; or equivalently, if the inverse image $\xi^{-1}(B) \equiv \{\omega : \xi(\omega) \in B\}$ is a measurable set in Ω.*

DEFINITION 5: *Let $\xi_1, \xi_2, \ldots$ be random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. The sequence $\{\xi_n\}$ is called a sequence of independent identically distributed (i.i.d.) random variables, if each random variable ξ_i has the same probability distribution as the others and all are mutually independent.*

Just as in analysis, there are various kinds of convergence of random variables in probability theory. Four of these are particularly important: *with probability one*, *in probability*, *in distribution*, *and in mean of order r* (Shiryayev, 1984).

DEFINITION 6: *Let $\{\xi_n\}$ and ξ be random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. The sequence $\{\xi_n\}$ converges almost surely (with probability one, almost everywhere) to the random variable ξ if*

$$P\left\{\lim_{n\to\infty} \xi_n(\omega) = \xi\right\} = 1, \tag{36}$$

that is, if the set of sample points ω for which $\xi_n(\omega)$ does not converge to ξ has probability zero. This convergence is denoted by $\xi_n \to \xi$ (P-a.s.), or $\xi_n \xrightarrow{a.s.} \xi$ or $\xi_n \xrightarrow{a.e.} \xi$.

DEFINITION 7: *Let $\{\xi_n\}$ and ξ be random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. The sequence $\{\xi_n\}$ converges in probability to the random variable ξ (notation: $\xi_n \xrightarrow{P} \xi$) if for every $\varepsilon > 0$,*

$$\lim_{n\to\infty} P\{|\xi_n(\omega) - \xi| > \varepsilon\} = 0, \tag{37}$$

or

$$\lim_{n\to\infty} P\{|\xi_n(\omega) - \xi| < \varepsilon\} = 1. \tag{38}$$

DEFINITION 8: *Let $\{\xi_n\}$ and ξ be random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ with distribution function $F_n(x)$ and $F(x)$ respectively. The sequence $\{\xi_n\}$ converges to a distribution to the random variable ξ(notation: $\xi_n \xrightarrow{d} \xi$) if for every continuous point x,*

$$\lim_{n\to\infty} F_n(x) = F(x) \tag{39}$$

is satisfied.

DEFINITION 9: *Let $\{\xi_n\}$ and ξ be random variables defined on a probability space $(\Omega, \mathcal{F}, P)$ with $E|\xi_n|^r < \infty$, $0 < r < \infty$. The sequence $\{\xi_n\}$ converges to a mean of order r to the random variable ξ (notation: $\xi_n \xrightarrow{r} \xi$) if*

$$\lim_{n\to\infty} E|\xi_n - \xi|^r = 0. \tag{40}$$

The following theorem indicates the relationships between the four kinds of convergence. It should be noted that the converses of Equations (41), (42), and (43) are false in general.

THEOREM 3: *We have the following implications:*

$$\xi_n \xrightarrow{a.s.} \xi \Rightarrow \xi_n \xrightarrow{P} \xi, \tag{41}$$

$$\xi_n \xrightarrow{P} \xi \Rightarrow \xi_n \xrightarrow{d} \xi, \tag{42}$$

$$\xi_n \xrightarrow{r} \xi \Rightarrow \xi_n \xrightarrow{P} \xi, \quad r > 0, \tag{43}$$

The following two theorems present the strong and weak laws of large numbers, respectively.

THEOREM 4 (KOLMOGOROV'S STRONG LAW OF LARGE NUMBERS): *Let $\xi_1, \xi_2, \ldots$ be a sequence of independent identically distributed random variables with $E|\xi_1| < \infty$. Then*

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i \xrightarrow{a.s.} E|\xi_1|. \tag{44}$$

THEOREM 5 (KHINTCHINE'S WEAK LAW OF LARGE NUMBERS): *Let $\xi_1, \xi_2, \ldots$ be a sequence of independent identically distributed random variables with $E|\xi_1| < \infty$. Then for every $\varepsilon > 0$,*

$$\lim_{n\to\infty} P\left\{\left|\frac{1}{n}\sum_{i=1}^{n} \xi_i - E|\xi_1|\right| < \epsilon\right\} = 1, \tag{45}$$

or

$$\frac{1}{n}\sum_{i=1}^{n} \xi_i \xrightarrow{P} E|\xi_1|. \tag{46}$$

3.2 Simplification of the Iterative Equations

It is obvious that the convergence or boundedness of the position of an individual particle is consistent with the convergence or boundedness of its component in each dimension. In detail, for QPSO-Type 1, the necessary and sufficient condition for $X_{i,n}$ to converge to $p_{i,n}$ in any kind of convergence is that $X_{i,n}^j$ converges to $p_{i,n}^j$ for each $1 \le j \le N$ in that kind of convergence; for QPSO-Type 2, the necessary and sufficient condition for $X_{i,n}$ to be probabilistic bounded is that $X_{i,n}^j$ is probabilistic bounded for each $1 \le j \le N$.

Furthermore, it appears from Equation (34) or (35) that each dimension of the particle's position is updated independently following the same equation. The only link between the dimensions of the problem space relies on the objective function, and in turn, through the locations of the personal and global best positions found so far, or the mean best position among particles. As a result, to investigate the convergence

or boundedness of an individual particle, we only need to analyze the convergence or boundedness of the component in any dimension. Without loss of universality, the issue of the convergence or boundedness for a particle in the N-dimensional space can be reduced to the one for a single particle in the one-dimensional space, and we only need to investigate the behavior of the particle in the one-dimensional space using the iterative equation given by

$$X_{n+1} = p \pm \alpha |X_n - p| \ln(1/u_{n+1}), \quad u_{n+1} \sim U(0,1), \tag{47}$$

or

$$X_{n+1} = p \pm \alpha |X_n - C| \ln(1/u_{n+1}), \quad u_{n+1} \sim U(0,1). \tag{48}$$

In the above equations, the local attractor of the particle and the mean best position are replaced by p and C, which are treated as probabilistic bounded random variables, instead of constants as in Clerc and Kennedy (2002). Here, the probabilistic boundedness of p and C means that $P\{\sup |p| < +\infty\} = 1$ and $P\{\sup |C| < +\infty\} = 1$. The position sequence $\{X_n\}$ is a sequence of random variables and $\{u_n\}$ is a sequence of independent random variables with $u_n \sim U(0,1)$ for all $n > 0$. For convenience, we denote a particle moving according to Equation (34) or (47) as a Type-1 particle, and a particle moving according to Equation (35) or (48) as a Type-2 particle. The remainder of this section focuses on the behaviors of the two types of particles.

Moreover, it is obvious that the sequence $\{u_n\}$ in Equation (47) or (48) is a sequence of independent identically distributed random variables, since each random variable u_i is generated independently and has the same probability distribution, that is, the uniform distribution on (0, 1), as the others.

3.3 Convergence of the Type-1 Particle

Rewriting Equation (47) as

$$|X_{n+1} - p| = \alpha |X_n - p| \ln(1/u_{n+1}) = \lambda_{n+1} |X_n - p|, \quad u_{n+1} \sim U(0,1), \tag{49}$$

where $\lambda_{n+1} = \alpha \ln(1/u_{n+1})$, we obtain

$$|X_n - p| = |X_0 - p| \prod_{i=1}^{n} \lambda_i, \tag{50}$$

where X_0 is the initial position of the particle. As such, the convergence analysis of the sequence of random variables $\{X_n\}$ can be reduced to that of the infinite product

$$\beta_n = \prod_{i=1}^{n} \lambda_i = \alpha^n \prod_{i=1}^{n} \ln(1/u_i). \tag{51}$$

The following theorem gives an integral which is of significance to the succedent analysis of the convergence of $\{\beta_n\}$ and $\{X_n\}$.

THEOREM 6: *The improper integral*

$$\int_0^1 \ln\left[\ln\left(\frac{1}{x}\right)\right] dx = -\gamma, \tag{52}$$

holds with the Euler-Mascheroni constant $\gamma \approx 0.5772156649$.

PROOF: Let $s = \frac{1}{x}$, and thus we have

$$\int_0^1 \ln\left[\ln\left(\frac{1}{x}\right)\right] dx = \int_0^{+\infty} e^{-s} \ln s \, ds.$$

Since

$$\Gamma(m) = \int_0^1 x^{m-1} e^{-x} \, dx,$$

where $\Gamma(\cdot)$ is the Gamma function,

$$\Gamma'(m) = \int_0^1 x^{m-1} e^{-x} \ln x \, dx.$$

From Courant (1989) we have that

$$\Gamma'(1) = \int_0^1 e^{-x} \ln x \, dx = -\gamma,$$

which implies that

$$\int_0^1 \ln\left[\ln\left(\frac{1}{x}\right)\right] dx = \int_0^{+\infty} e^{-s} \ln s \, ds = \Gamma'(1) = -\gamma.$$

This completes the proof of the theorem. □

With the above preliminaries, in the following part of this section, we derive the sufficient and necessary condition for $\{X_n\}$ to converge to random variable p in each kind of convergence. The detailed proof of the almost sure convergence is provided in the text, while those of other kinds of convergence are presented in the Appendix.

3.3.1 Almost Sure Convergence

LEMMA 1: *If* $\{u_n\}$ *is a sequence of independent identically distributed random variables with* $u_n \sim U(0, 1)$ *for all* $n > 0$ *and* $\zeta_n = \ln[\ln(1/u_n)]$, *then*

$$\frac{1}{n} \sum_{i=1}^{n} \zeta_i \xrightarrow{a.s.} -\gamma. \tag{53}$$

PROOF: Since $\{u_n\}$ is a sequence of independent identically distributed (*i.i.d.*) random variables, $\{\xi_n\}$ is also a sequence of *i.i.d.* random variables. Theorem 6 implies that

$$E(\zeta_1) = E\left\{\ln\left[\ln\left(\frac{1}{u_1}\right)\right]\right\} = \int_0^1 \ln\left[\ln\left(\frac{1}{x}\right)\right] dx = -\gamma.$$

Thus, by Theorem 4 (Kolmogorov's strong law of large numbers), we have

$$\frac{1}{n}\sum_{i=1}^{n}\zeta_i \xrightarrow{a.s.} E(\xi_1) = -\gamma.$$

This completes the proof of the lemma. □

THEOREM 7: *The necessary and sufficient condition for the position sequence of a Type-1 particle* $\{X_n\}$ *to converge almost surely to* p (*i.e.*, $X_n \xrightarrow{a.s.} p$) *is that* $\alpha < e^{\gamma}$.

PROOF: Before the proof, we provide the following two groups of equivalent propositions at first.

(a) If $X_n \xrightarrow{a.s.} p$, then

$$P\left\{\lim_{n\to\infty}|X_n - p| = 0\right\} = 1,$$

which is equivalent to

$$\begin{aligned}
&P\left\{\lim_{n\to\infty}|X_n - p| = 0\right\} = 1\\
&\Leftrightarrow P\left\{\lim_{n\to\infty}\beta_n = 0\right\} = 1\\
&\Leftrightarrow P\left\{\lim_{n\to\infty}\ln\left[\alpha^n\prod_{i=1}^{n}\ln\left(\frac{1}{u_i}\right)\right] = -\infty\right\} = 1\\
&\Leftrightarrow P\left\{\lim_{n\to\infty}\left(n\ln\alpha + \sum_{i=1}^{n}\ln\left[\ln\left(\frac{1}{u_i}\right)\right]\right) = -\infty\right\} = 1\\
&\Leftrightarrow P\left\{\lim_{n\to\infty}\left(n\ln\alpha + \sum_{i=1}^{n}\zeta_i\right) = -\infty\right\} = 1\\
&\Leftrightarrow \forall m \in Z^+,\ \exists K_1 \in Z^+ \text{ such that } P\left\{k\ln\alpha + \sum_{i=1}^{k}\zeta_i < -m\right\} = 1, \text{ whenever } k \geq K_1.\\
&\Leftrightarrow \forall m \in Z^+,\ \exists K_1 \in Z^+ \text{ such that } P\left\{\ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < -\frac{m}{k}\right\} = 1, \text{ whenever } k \geq K_1.\\
&\Leftrightarrow P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left(\ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < -\frac{m}{k}\right)\right\} = 1, \qquad (54)
\end{aligned}$$

where Z^+ is the set of all positive integers.

(b) Lemma 1 has the following equivalences

$$\frac{1}{n}\sum_{i=1}^{n}\zeta_i \xrightarrow{a.s.} -\gamma$$

$$\Leftrightarrow P\left\{\lim_{n\to\infty}\left|\frac{1}{n}\sum_{i=1}^{n}\zeta_i+\gamma\right|=0\right\}=1$$

$$\Leftrightarrow \forall m\in Z^+,\ \exists K_2\in Z^+ \text{ such that } P\left\{\left|\frac{1}{k}\sum_{i=1}^{k}\zeta_i+\gamma\right|<\frac{1}{m}\right\}=1,\ \text{whenever } k\geq K_2.$$

$$\Leftrightarrow \forall m\in Z^+,\ \exists K_2\in Z^+ \text{ such that } P\left\{-\gamma-\frac{1}{m}<\frac{1}{k}\sum_{i=1}^{k}\zeta_i<-\gamma+\frac{1}{m}\right\}=1,$$

whenever $k\geq K_2$.

$$\Leftrightarrow \forall m\in Z^+,\ \exists K_2\in Z^+ \text{ such that } P\left\{\ln\alpha-\gamma-\frac{1}{m}<\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\ln\alpha-\gamma+\frac{1}{m}\right\}=1,$$

whenever $k\geq K_2$.

PROOF OF NECESSITY: If $X_n \xrightarrow{a.s.} p$, the propositions in (a) hold. Since the propositions in (b) also hold due to Lemma 1, for any positive integer m, there exists $K=\max(K_1, K_2)$ such that whenever $k\geq K$,

$$P\left\{\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<-\frac{m}{k}\right\}=1$$

and

$$P\left\{\ln\alpha-\gamma-\frac{1}{m}<\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\ln\alpha-\gamma+\frac{1}{m}\right\}=1 \tag{55}$$

holds simultaneously, and thus it follows that $P\{\ln\alpha-\gamma-\frac{1}{m}<-\frac{m}{k}\}=1$. That is, $\forall m\in Z^+$, $\exists K=\max(K_1, K_2)$, such that whenever $k\geq K$,

$$P\left\{\ln\alpha<\gamma+\frac{1}{m}-\frac{m}{k}\right\}=1.$$

This proposition is equivalent to

$$
\begin{aligned}
& P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left(\ln\alpha < \gamma + \frac{1}{m} - \frac{m}{k}\right)\right\} = 1 \\
& \Leftrightarrow P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\left(\ln\alpha < \gamma + \frac{1}{m} - \frac{m}{n}\right)\right\} = 1 \\
& \Leftrightarrow P\left\{\bigcap_{m=1}^{\infty}\left(\ln\alpha < \gamma + \frac{1}{m}\right)\right\} = 1 \\
& \Leftrightarrow P\{\ln\alpha < \gamma\} = 1 \Leftrightarrow \ln\alpha < \gamma \\
& \Leftrightarrow \alpha < e^{\gamma}.
\end{aligned}
$$

This ends the proof of necessity.

PROOF OF SUFFICIENCY

(i) From the equivalences of Lemma 1, we have that $\forall m \in Z^+$, $\exists K_1 \in Z^+$ such that whenever $k \geq K_1$

$$
P\left\{\ln\alpha - \gamma - \frac{1}{m} < \ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < \ln\alpha - \gamma + \frac{1}{m}\right\} = 1. \tag{56}
$$

Since $\alpha < e^{\gamma}$, $\ln\alpha < \gamma$, we have

$$
\ln\alpha - \gamma + \frac{1}{m} < \frac{1}{m}, \tag{57}
$$

and therefore

$$
\begin{aligned}
& \left\{\ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < \frac{1}{m}\right\} \\
& \supset \left\{\ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < \ln\alpha - \gamma + \frac{1}{m}\right\} \\
& \supset \left\{\ln\alpha - \gamma - \frac{1}{m} < \ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < \ln\alpha - \gamma + \frac{1}{m}\right\}.
\end{aligned}
$$

From Equation (56), we have

$$
\begin{aligned}
& P\left\{\ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < \frac{1}{m}\right\} \\
& \geq P\left\{\ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < \ln\alpha - \gamma + \frac{1}{m}\right\} \\
& \geq P\left\{\ln\alpha - \gamma - \frac{1}{m} < \ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < \ln\alpha - \gamma + \frac{1}{m}\right\} = 1,
\end{aligned}
$$

and thus

$$P\left\{\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\}=1. \tag{58}$$

Since $-\frac{m}{k}<\frac{1}{m}$, we find that

$$\left\{\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<-\frac{m}{k}\right\}=\left\{\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\}-\left\{-\frac{m}{k}\le\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\},$$

resulting in the fact that

$$\begin{aligned}&P\left\{\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<-\frac{m}{k}\right\}\\&=P\left\{\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\}-P\left\{-\frac{m}{k}\le\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\}\\&=1-P\left\{-\frac{m}{k}\le\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\}.\end{aligned} \tag{59}$$

(ii) $\forall m\in Z^+$, $\exists K_2=m^2$ such that whenever $k\ge K_2$, $-\frac{m}{k}\ge-\frac{1}{m}$, from which we have

$$\left\{-\frac{m}{k}\le\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\}\subset\left\{-\frac{1}{m}<\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\},$$

and thus have

$$P\left\{-\frac{m}{k}\le\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\}\le P\left\{-\frac{1}{m}<\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\}. \tag{60}$$

From Equations (59) and (60), we have that $\forall m\in Z^+$, $\exists K=\max(K_1,K_2)$ such that whenever $k\ge K$,

$$\begin{aligned}&P\left\{\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<-\frac{m}{k}\right\}\\&=1-P\left\{-\frac{m}{k}\le\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\}\ge1-P\left\{-\frac{1}{m}<\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right\},\end{aligned}$$

which is equivalent to

$$\begin{aligned} & P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left(\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<-\frac{m}{k}\right)\right\} \\ & \geq 1-P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left(-\frac{1}{m}<\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<\frac{1}{m}\right)\right\} \\ & = 1-P\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\zeta_i=-\ln\alpha\right\}. \end{aligned} \tag{61}$$

Since $\frac{1}{n}\sum_{i=1}^{n}\zeta_i \xrightarrow{a.s.} -\gamma$, we obtain

$$P\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\zeta_i=-\ln\alpha\right\}=P\{\ln\alpha=-\gamma\}.$$

The condition that $\alpha < e^{\gamma}$ implies that $P\{\ln\alpha=\gamma\}=0$, so

$$P\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\zeta_i=-\ln\alpha\right\}=0. \tag{62}$$

From the inequality in Equation (61) and from Equation (62), we obtain

$$\begin{aligned} & P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left(\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<-\frac{m}{k}\right)\right\} \\ & \geq 1-P\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\zeta_i=-\ln\alpha\right\}=1-0=1, \end{aligned}$$

and thus

$$P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left(\ln\alpha+\frac{1}{k}\sum_{i=1}^{k}\zeta_i<-\frac{m}{k}\right)\right\}=1.$$

Considering Equation (54), we find that

$$P\left\{\lim_{n\to\infty}|X_n-p|=0\right\}=1. \tag{63}$$

This ends the proof of sufficiency.
This completes the proof of the theorem. □

3.3.2 Convergence in Probability

THEOREM 8: *The necessary and sufficient condition for the position sequence of a Type-1 particle* $\{X_n\}$ *to converge to* p *in probability* ($X_n \xrightarrow{P} p$) *is that* $\alpha < e^{\gamma}$.

PROOF: See the proof in the Appendix. □

3.3.3 Convergence in Distribution

THEOREM 9: *The necessary and sufficient condition for the position sequence of a Type-1 particle* $\{X_n\}$ *to converge to* p *in distribution* ($X_n \xrightarrow{d} p$) *is that* $\alpha < e^{\gamma}$.

PROOF: See the proof in the Appendix. □

3.3.4 Convergence in Mean of Order r

THEOREM 10: *The necessary and sufficient condition for the position sequence of a Type-1 particle* $\{X_n\}$ *to converge to* p *in mean of order* r ($X_n \xrightarrow{r} p,\ 0 < r < \infty$) *is that* $\alpha < e^{\gamma}$.

PROOF: See the proof in the Appendix. □

3.3.5 Some Discussion

The above analysis indicates that the sufficient and necessary condition for the position sequence of a Type-1 particle to converge to p in any of four kinds of convergence is that $\alpha < e^{\gamma}$. In other words, the four kinds of convergence in this case are equivalent conditionally. This conclusion does not conflict with Theorem 3. It is shown by Equation (41) that almost sure convergence implies convergence in probability, while the converse is in general not true. Under some special circumstances, the converse of Equation (41) may hold. Kolmogorov's strong law and Khintchine's weak law of large numbers reveal that a convergence in probability may imply almost sure convergence for certain cases, leading to the equivalence between the two kinds of convergence of the Type-1 particle.

In any case, it is due to the probabilistic boundedness of p and its uniform distribution on (P, G) or (G, P) that the converse of Equation (42) holds for the Type-1 particle, as shown in the proof of Theorem 9. Owing to the equivalence between the convergence in probability and the convergence in distribution, the point p can be treated as a constant as we do in the stochastic simulations in the next section. Furthermore, as shown by the proof of Theorem 10 in the Appendix, since $\{X_n\}$ is uniformly integrable, the sufficient and necessary condition for its convergence in the mean of order r is the same as for its convergence in probability.

It is shown in the proof of Theorem 9 that when $\alpha = e^{\gamma}$, the position of the Type-1 particle can be any real number but infinity, which means that the position is probabilistically bounded as n tends to infinity. Therefore, we may conclude that if $\alpha \leq e^{\gamma}$, the position of the particle is probabilistically bounded; otherwise, it is divergent.

3.4 Boundedness of the Type-2 Particle

Now we turn our attention to the behavior of a Type-2 particle, beginning by rewriting Equation (48) as

$$X_{n+1} - C = p - C \pm \alpha |X_n - C| \ln(1/u_{n+1}) = p - C \pm \lambda_{n+1} |X_n - C|, \tag{64}$$

where $\lambda_{n+1} = \alpha \ln(1/u_{n+1})$ and $u_{n+1} \sim U(0, 1)$. From the above equation, we have the following two inequalities:

$$|X_n - C| \leq |C - p| + \lambda_n |X_{n\text{-}1} - C|, \tag{65}$$

and

$$|X_n - C| \geq -|C - p| + \lambda_n |X_{n\text{-}1} - C|. \tag{66}$$

Based on the inequalities in Equations (65) and (66), the proofs of the following theorems derive the necessary and sufficient condition for the probabilistic boundedness of the Type-2 particle's position, which is also related to the probabilistic boundedness of $\beta_n = \prod_{i=1}^{n} \lambda_i$.

THEOREM 11: *The necessary and sufficient condition for $\beta_n = \prod_{i=1}^{n} \lambda_i$ to be probabilistically bounded (i.e., $P\{\sup_{n>0} \beta_n < +\infty\} = 1$) is that $\alpha \leq e^{\gamma}$.*

PROOF: The proof is essentially given by the first two parts of Theorem 10 in the Appendix. □

THEOREM 12: *The necessary and sufficient condition for the position sequence of a Type-2 particle $\{X_n\}$ to be probabilistically bounded (i.e., $P\{\sup_{n>0} X_n < +\infty\} = 1$), is that $\alpha \leq e^{\gamma}$.*

PROOF: Since λ_n is a continuous random variable, it is evident that $P\{\lambda_n = 1,\ n > 0\} = 0$. Therefore, $P\{\sup |C - p| < +\infty\} = 1$ implies that

$$P\left\{\sup_{n>0}\left(\frac{|C-p|}{1-\lambda_n}\right) < +\infty\right\} = 1.$$

Denoting $\sup_{n>0}\left(\frac{|C-p|}{1-\lambda_n}\right) = r$, where $0 < r < +\infty$, we have that for every $n > 0$, $\frac{|C-p|}{1-\lambda_n} \leq r$, namely

$$|C - p| \leq r(1 - \lambda_n). \tag{67}$$

PROOF OF SUFFICIENCY: By replacing $|C - p|$ in Equation (65) by that in the inequality in Equation (67), we obtain the following inequality:

$$|X_n - C| - r \leq \lambda_n(|X_{n-1} - C| - r), \quad \text{(for every } n > 0),$$

from which we find that

$$|X_n - C| - r \le \lambda_n(|X_{n\text{-}1} - C| - r) \le \lambda_n\lambda_{n-1}(|X_{n\text{-}2} - C| - r)$$
$$\le \lambda_n\lambda_{n\text{-}1}\lambda_{n\text{-}2}(|X_{n\text{-}3} - C| - r) \le \cdots \le (|X_0 - C| - r)\prod_{i=1}^{n}\lambda_i.$$

Thus the following inequality holds:

$$|X_n - C| \le r + (|X_0 - C| - r)\prod_{i=1}^{n}\lambda_i. \tag{68}$$

Since $\beta_n = \prod_{i=1}^{n}\lambda_i > 0$,

$$\sup_{n>0}|X_n - C| \le \sup_{n>0}\left[r + (|X_0 - C| - r)\prod_{i=1}^{n}\lambda_i\right]$$
$$\le r + \sup(|X_0 - C| - r)\sup_{n>0}\left(\prod_{i=1}^{n}\lambda_i\right) = r + \sup(|X_0 - C| - r)\sup_{n>0}(\beta_n).$$

By Theorem 11, we have that, whenever $\alpha \le e^{\gamma}$,

$$P\left\{\sup_{n>0}\beta_n < +\infty\right\} = P\left\{\sup_{n>0}\left(\prod_{i=1}^{n}\lambda_i\right) < +\infty\right\} = 1.$$

Considering that $0 < r < +\infty$, we have $P\{\sup(|X_0 - C| - r) < +\infty\} = 1$. Therefore

$$P\{\sup_{n>0}|X_n - C| < +\infty\}$$
$$\ge P\left\{r + \sup(|X_0 - C| - r)\sup_{n>0}\left(\prod_{i=1}^{n}\lambda_i\right) < +\infty\right\}$$
$$= P\left\{\sup(|X_0 - C| - r)\sup_{n>0}\left(\prod_{i=1}^{n}\lambda_i\right) < +\infty\right\}$$
$$\ge P\left\{[\sup(|X_0 - C| - r) < +\infty] \cap \left(\sup_{n>0}\beta_n < +\infty\right)\right\}$$
$$= P\left\{\sup(|X_0 - C| - r) < +\infty\right\} + P\left\{\sup_{n>0}\beta_n < +\infty\right\}$$
$$-P\left\{[\sup(|X_0 - C| - r) < +\infty] \cup \left[\sup_{n>0}\beta_n < +\infty\right]\right\}$$
$$= 1 + 1 - P\left\{\left[\sup(|X_0 - C| - r) < +\infty\right] \cup \left[\sup_{n>0}\beta_n < +\infty\right]\right\} \ge 1,$$

from which we immediately have that $P\{\sup_{n>0}|X_n - C| < +\infty\} = 1$, namely, $|X_n - C|$ is probabilistically bounded. This implies that X_n is also probabilistically bounded, that is, $P\{\sup_{n>0} X_n < +\infty\} = 1$. This ends the proof of sufficiency.

PROOF OF NECESSITY: Replacing $|C - p|$ in Equation (66) by that in Equation (67), we find that

$$|X_n - C| + r \geq \lambda_n(|X_{n\text{-}1} - C| + r), \quad \text{(for every } n > 0\text{)}, \tag{69}$$

from which we obtain

$$\begin{aligned} |X_n - C| + r &\geq \lambda_n(|X_{n\text{-}1} - C| + r) \geq \lambda_n\lambda_{n\text{-}1}(|X_{n\text{-}2} - C| + r) \\ &\geq \lambda_n\lambda_{n\text{-}1}\lambda_{n\text{-}2}(|X_{n\text{-}3} - C| + r) \geq \cdots \geq (|X_0 - C| + r)\prod_{i=1}^{n}\lambda_i. \end{aligned} \tag{70}$$

Thus we have that

$$\begin{aligned} \sup_{n>0}(|X_n - C| + r) &\geq \sup_{n>0}\left[(|X_0 - C| + r)\prod_{i=1}^{n}\lambda_i\right] \\ &= \sup_{n>0}[(|X_0 - C| + r)\beta_n] = \sup(|X_0 - C| + r)\sup_{n>0}\beta_n. \end{aligned}$$

As a result, it follows that

$$\begin{aligned} &P\left\{\sup_{n>0}(|X_n - C| + r) < +\infty\right\} \\ &\leq P\left\{\sup(|X_0 - C| + r)\sup_{n>0}\beta_n < +\infty\right\} \\ &= P\left\{[\sup(|X_0 - C| + r) < +\infty] \cap \left[\sup_{n>0}\beta_n < +\infty\right]\right\}. \end{aligned} \tag{71}$$

If X_n is probabilistically bounded, $|X_n - C|$ is also probabilistically bounded, that is, $P\{\sup_{n>0}|X_n - C| < +\infty\} = 1$. Since $r < +\infty$, $|X_n - C| + r$ is probabilistically bounded, that is,

$$P\left\{\sup_{n>0}|X_n - C| + r < +\infty\right\} = 1.$$

Considering the inequality in Equation (71), we have

$$P\left\{[\sup(|X_0 - C| + r) < +\infty] \cap \left[\sup_{n>0}\beta_n < +\infty\right]\right\} \geq 1,$$

which immediately results in

$$P\left\{[\sup(|X_0 - C| + r) < +\infty] \cap \left[\sup_{n>0} \beta_n < +\infty\right]\right\} = 1. \tag{72}$$

Due to the probabilistic boundedness of $|X_0 - C| + r$, we have

$$P\{\sup(|X_0 - C| + r) < +\infty\} = 1. \tag{73}$$

From Equations (72) and (73), we obtain $P\{\sup_{n>0} \beta_n < +\infty\} = 1$ which means $\alpha \leq e^{\gamma}$ as shown by Theorem 11. This implies that $\alpha \leq e^{\gamma}$ is the necessary condition for the probabilistic boundedness of X_n. This ends the proof of necessity.

This completes the proof of the theorem. □

Theorem 12 reveals that the sufficient and necessary condition for the probabilistic boundedness of a Type 2 particle is the same as that for the probabilistic boundedness of a Type 1 particle, since the behaviors of both types of particles are related to the probabilistic boundedness of β_n. Besides, the behavior of the Type 2 particle is also influenced by point C. In practice, when the QPSO algorithm is running, the personal best positions of all the particles converge to the same point. This implies that $|C - p|$ converges to zero almost surely, and thus $\sup_{n>0}[|C - p|/(1 - \lambda_n)] = r$ also converges to zero almost surely. Hence, if and only if $\alpha < e^{\gamma}$, $P\{\lim_{n\to\infty} \beta_n = 0\} = 1$. According to the inequality in Equation (68), we can find that $P\{\lim_{n\to\infty} |X_n - C| = 0\} = 1$ or $P\{\lim_{n\to\infty} |X_n - p| = 0\} = 1$.

4 Experimental Analysis

4.1 Stochastic Simulations on the Particle's Behavior

It has been demonstrated that the behavior of an individual particle in QPSO is subject to how the value of α is selected. In order to verify the correctness of the previous analysis, we carried out stochastic simulations for the behavior of the particle. Two groups of simulations were implemented, one for a Type-1 particle and the other for a Type-2 particle.

The value of α for either group of simulations was selected from a series of numbers in [0.5, 2.0] and the corresponding maximum number of iterations ($n_{\max}$) executed was sufficiently large for the purpose of convergence, boundedness, and divergence. When the stochastic simulation was executed, the logarithmic value of the distance between the current position X_n and the point p was recorded as the ordinate, and the number of iterations was the abscissa. Without loss of generality, in the simulations for a Type-1 particle, p was fixed at the origin, and the initial position of the particle was set as $X_0 = 1000$; in the simulations for a Type-2 particle, C was fixed at $X = 0.001$, and p and the initial position were the same as those in the simulations for the Type 1 particle. The simulations were performed on Matlab 7.0 and the results are shown in Figures 1 and 2.

Figure 1 shows the simulation results for the position of the Type-1 particle. When $\ln|X_n - p|$ is smaller than -700 or is larger than 700, $|X_n - p|$ reaches the minimum or maximum positive value that the computer can identify, and thus we can consider that $|X_n - p|$ converges to zero (X_n converges to p) or diverges to infinity. It can be

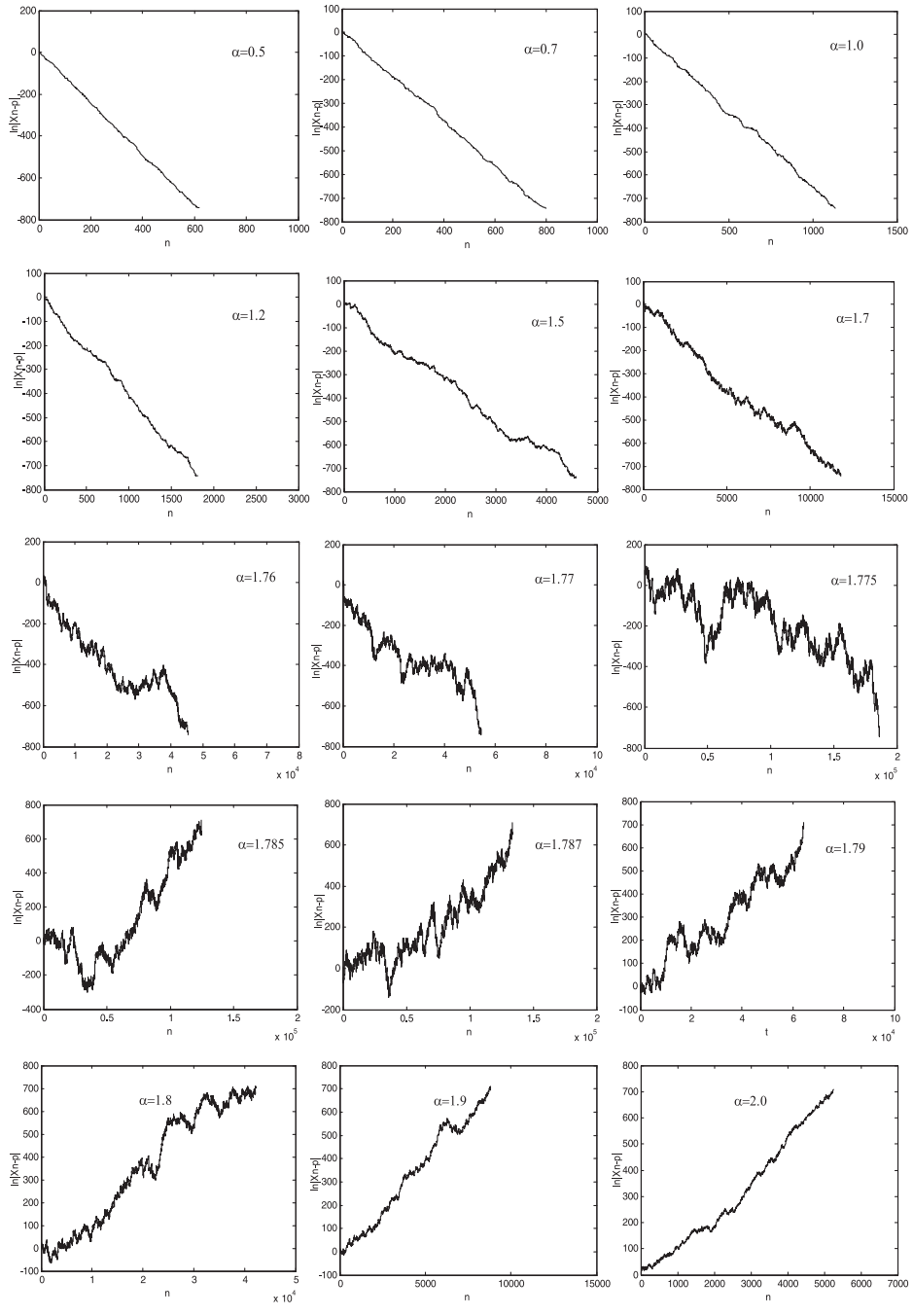

Figure 1: Simulation results for a Type-1 particle. When $\alpha \leq 1.775$, $\ln|X_n - p| \to -\infty$ (i.e., $|X_n - p| \to 0$) as n increased, and the smaller α resulted in a faster convergence speed. When $\alpha \geq 1.785$, $\ln|X_n - p| \to +\infty$ (i.e., $|X_n - p| \to +\infty$) as n increased, and the larger α led to a faster divergence speed.

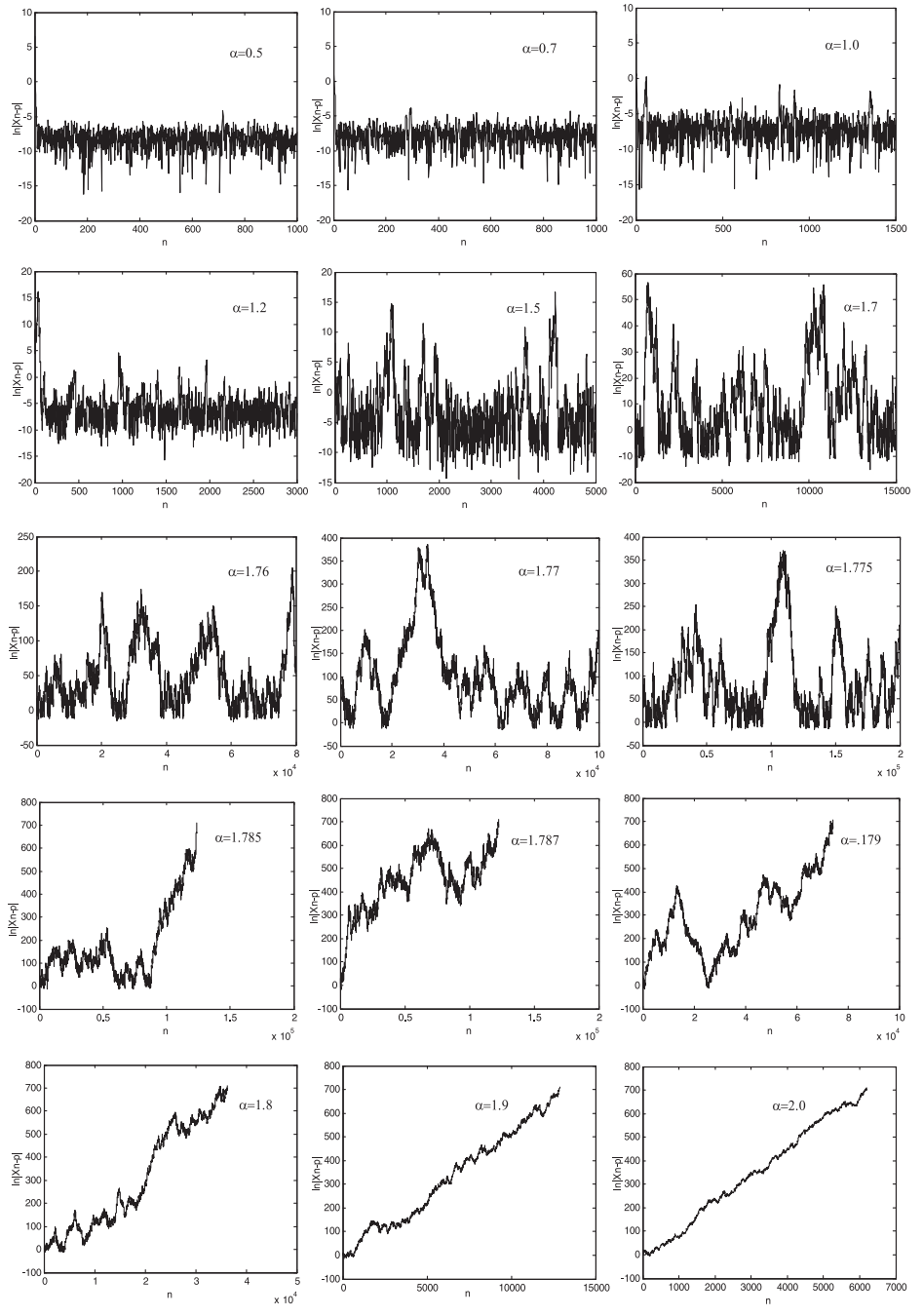

Figure 2: Simulation results for a Type-2 particle. When $\alpha \leq 1.775$, $\ln|X_n - p|$ was bounded as n increased, and the smaller α resulted in a narrower oscillation range. When $\alpha \geq 1.785$, $\ln|X_n - p| \to +\infty$ (i.e., $|X_n - p| \to +\infty$) as n increased, and the larger α led to a faster divergence speed.

seen from Figure 1 that the particle's position X_n converged to p when $\alpha \leq 1.775$, and diverged when $\alpha \geq 1.785$. When X_n converged to p, the smaller the value of α, the faster X_n converged. On the other hand, when X_n diverged, the larger value of α resulted in a faster speed at which the position careened to infinity.

As for a Type-2 particle, according to Figure 2, its position oscillated around p and C when $\alpha \leq 1.775$, while the position exploded when $\alpha \geq 1.785$. The results also show that when X_n was probabilistically bounded, the smaller value of α resulted in a narrower oscillation range of the particle's position. When X_n diverged, the larger value of α led to a faster divergence speed.

It can be concluded that there exists an α_0 within (1.775, 1.785) such that whenever $\alpha \leq \alpha_0$, the position of the Type-1 particle converges (when $\alpha < \alpha_0$) or is probabilistic bounded (when $\alpha = \alpha_0$), and that of the Type-2 particle is probabilistic bounded; otherwise the position of either type of particles diverges. Therefore, it is obvious that the simulation results are consistent with the theoretical results in the previous section that setting $\alpha \leq e^{\gamma} \simeq 1.781$ can prevent explosion of the particle's position.

The fact that the smaller value of α results in the faster convergence speed of a Type-1 particle and the narrower oscillation range of a Type-2 particle can be illustrated as follows. For a Type-1 particle, according to Equation (47), we define the convergence speed at the nth iteration by

$$c_n = E\left[\frac{|X_n - p|}{|X_{n\text{-}1} - p|}\middle| X_{n\text{-}1}\right] = E[\alpha \ln(1/u_n)] = -\alpha E[\ln u_n] = -\alpha \int_0^1 \ln u \, du = \alpha. \quad (74)$$

Thus we have

$$E\left[|X_n - p| \middle| X_{n\text{-}1}\right] = c_n |X_{n\text{-}1} - p| = \alpha |X_{n\text{-}1} - p|, \quad (75)$$

and note that the velocity at which $|X_n - p|$ declines, increases with the smaller c_n, namely, the smaller α. For a Type-2 particle, according to Equation (48), we find that

$$E[|X_n - p| \big| X_{n\text{-}1}] = \alpha |X_{n\text{-}1} - C|, \quad (76)$$

implying that a smaller value of α results in a smaller conditional expected value of $|X_n - p|$ and thus a narrower oscillation range of X_n. It is obvious that the slower convergence of a Type-1 particle or the wider oscillation range of a Type-2 particle means a stronger global search ability of the particle, while the faster convergence or the narrower oscillating scope implies a stronger local search ability. Excessive global search may result in slow convergence of the QPSO algorithm, and on the other hand, excessive local search may cause premature convergence of the algorithm. It is significant, therefore, to balance the global search (exploration) and local search (exploitation) to obtain a generally good performance of the algorithm. The parameter α, whose value selection is studied empirically in the rest of this section, plays a major role in balancing the exploration and exploitation of the particles in QPSO.

4.2 Methods of Parameter Control

When an algorithm is employed to solve a problem at hand, one of the most important issues is how to select its parameters. It is evident that α is the most influential parameter on the convergence properties of QPSO, except for the population size. Although it

was demonstrated in Section 3 that it is necessary and sufficient to set $\alpha \leq e^{\gamma}$ to prevent the individual from explosion and guarantee the convergence of the particle swarm, this does not mean that any value of α less than or equal to e^{γ} can lead to a satisfactory performance of QPSO in practical applications. In the remainder of this section, we aim at finding out through empirical study how to control and select the value of α so that QPSO may yield good performance in general. We also provide a performance comparison with other forms of PSO. To do so, we tested the algorithms on 12 benchmark functions F_1 to F_{12} presented by Suganthan et al. (2005), which are not listed here due to space limitations.

When QPSO is applied to practical problems, there are several control methods for the parameter α. A simple approach is to set α to be a fixed value when the algorithm is running. Another efficient method is to decrease the value of α linearly in the course of search. That is, α is determined by

$$\alpha = \frac{(\alpha_1 - \alpha_2)(n_{\max} - n)}{n_{\max}} + \alpha_2,$$

where α_1 and α_2 ($\alpha_1 > \alpha_2$) are the initial and final values of α, respectively, n is the current iteration number, and $n_{\max}$ is the maximum number of allowable iterations.

We also propose a variant of QPSO-Type 2 in which the mean best position C_n in Equation (15) is replaced by the pbest position of a randomly selected particle in the swarm at each iteration. To distinguish between them, we denote the QPSO-Type 2 with C_n as QPSO-Type 2-I and that with a randomly selected pbest position as QPSO-Type 2-II.

4.3 Empirical Studies on Parameter Selection

When either control strategy for α is used in QPSO, the value of α in the fixed-value method or α_1 and α_2 in the linear time-varying approach must be determined. To select the values that can yield generally good algorithmic performance, we tested QPSO-Type 1, QPSO-Type 2-I and QPSO-Type 2-II on three frequently used functions in Suganthan's benchmark suite: Shifted Rosenbrock Function (F_6), Shifted Rotated Griewank's Function (F_7), and Shifted Rastrigin's Function (F_9), using the two methods of controlling α. Rastrigin and Griewank are two difficult multimodal problems and Rosenbrock is a unimodal problem. For each of these functions, $N = 30$. The expressions and bounds of the functions are shown in Suganthan et al. (2005). For each α setting, each algorithm using 20 particles was tested for 100 runs on each problem. In each trial run, the initial positions of the particles were determined randomly within the search bounds. To determine the effectiveness of each algorithm for the α setting of each control method on each problem, the best objective function value (i.e., the best fitness value) found after 3,000 iterations was averaged over 100 runs of the algorithm for the same α setting and the same benchmark function. The results were compared by α settings across the three benchmarks. The best α setting of each control method was selected by ranking the averaged best objective function values for each problem, adding the ranks, and taking the value that had the lowest summed rank, provided that the performance is acceptable (in the top half of the rankings) in all the tests for a particular setting of α.

The results for QPSO-Type 1 are presented in Table 1. When the fixed-value method was used, α was set to different values smaller than $e^{\gamma} \approx 1.781$ in each case. The results obtained for α outside the range [1.2, 0.85] are very poor and are not listed in the table. It

Table 1: Mean best fitness values obtained by QPSO-Type 1 on three benchmark functions.

Fixed α							
α	Rosenbrock	Rastrigin	Griewank	α	Rosenbrock	Rastrigin	Griewank
1.2	1.9761e+03	73.2189	75.2146	0.97	140.7474	65.2704	0.0216
1.10	207.2907	51.3072	0.7604	0.96	**128.861**	63.8414	0.0229
1.05	179.4073	**48.2471**	0.0676	0.95	173.8644	69.3123	0.0211
1.02	178.5613	53.0128	0.0259	0.94	157.0963	71.4457	0.0254
1.01	151.0301	55.9349	0.0246	0.93	168.8516	70.5280	0.0208
1.00	138.0746	56.4232	0.0218	0.92	193.6964	75.5213	0.0281
0.99	188.3713	52.7332	0.0217	0.90	152.7539	84.4062	0.1007
0.98	170.1586	60.7105	**0.0196**	0.85	282.2530	96.1226	51.0799

Linearly-Decreasing α							
$\alpha_1 \to \alpha_2$	Rosenbrock	Rastrigin	Griewank	$\alpha_1 \to \alpha_2$	Rosenbrock	Rastrigin	Griewank
$1.2 \to 0.8$	195.5114	60.3980	0.0688	$1.0 \to 0.8$	174.0003	59.4232	0.0390
$1.2 \to 0.7$	180.0561	64.5759	0.2481	$1.0 \to 0.7$	200.9611	65.3739	0.2538
$1.2 \to 0.6$	178.0053	63.8211	0.4468	$1.0 \to 0.6$	204.2796	61.2373	0.6978
$1.2 \to 0.5$	170.1478	63.3740	0.7066	$1.0 \to 0.5$	261.7130	71.5227	0.9640
$1.1 \to 0.9$	143.5103	50.7536	0.0276	$0.9 \to 0.8$	208.1228	81.9640	2.1316
$1.1 \to 0.8$	181.5607	57.9190	0.0346	$0.9 \to 0.7$	262.8041	81.9872	7.3829
$1.1 \to 0.7$	145.4834	55.5282	0.1123	$0.9 \to 0.6$	311.9669	87.7742	12.0388
$1.1 \to 0.6$	165.9128	61.5130	0.3720	$0.9 \to 0.5$	506.0865	87.1547	21.0660
$1.1 \to 0.5$	218.6484	58.8480	0.5683	$0.8 \to 0.7$	7.6355e+06	119.1037	259.0206
$1.0 \to 0.9$	**139.9815**	**54.4278**	**0.0209**	$0.8 \to 0.6$	3.3683e+06	111.3555	282.3997

can be observed from Table 1 that if the value of α was fixed during the search process, QPSO-Type 1 generated quite different objective function values for different values of α, particularly for the Shifted Rosenbrock Function, although the difference between two adjacent values of α is not significant. The results indicate that the performance of QPSO-Type 1 was somewhat sensitive to the value of α. The best results were obtained when $\alpha = 1.0$. When time-varying α was used, α_1 and α_2 ($\alpha_1 > \alpha_2$) were selected from a series of different values less than $e^{\gamma} \approx 1.781$. Only acceptable results are listed in the table. It can be seen that the algorithmic performance was sensitive to both α_1 and α_2. It is also found that decreasing α linearly from 1.0 to 0.9 led to the best performance in general.

Table 2 records the results for QPSO-Type 2-I. For the case in which the fixed-value method was used, the results of the algorithm for α outside the range $[0.6, 1.2]$ are not listed because of poor quality. It can be observed from the results that the performance of QPSO-Type 2-I was less sensitive to α than that of QPSO-Type 1. The results seemed fairly stable when α was valued in the interval $[0.8, 0.7]$. The value of α that generated the best results in this case was 0.75. When α decreased linearly in the course of search, the objective function values obtained by QPSO-Type 2-I with different α settings were not so different from each other as those obtained by QPSO-Type 1. It has been identified that varying α linearly from 1.0 to 0.5 could yield the best quality results.

The results for QPSO-Type 2-II are summarized in Table 3. It is clear from the results that the value of α, whether it was fixed or time-varying, should be set to be relatively smaller so that the algorithm was comparable in performance with the other types of

Table 2: Mean best fitness values obtained by QPSO-Type 2-I on three benchmark functions.

Fixed α							
α	Rosenbrock	Rastrigin	Griewank	α	Rosenbrock	Rastrigin	Griewank
1.20	5.0457e+007	209.4031	35.3916	0.75	**82.9908**	39.0991	0.0203
1.00	2.9384e+004	164.7224	1.5544	0.74	119.3931	41.9011	0.0205
0.95	1.5836e+003	150.1191	0.9811	0.73	167.4110	45.8569	0.0196
0.90	153.7730	127.1872	0.1448	0.72	162.2084	49.3278	0.0251
0.85	129.7591	97.7650	0.0193	0.71	275.5061	49.9125	0.0745
0.80	115.6558	50.7660	0.0216	0.70	217.7180	58.0127	1.8027
0.78	138.0215	39.3672	**0.0184**	0.65	1.0587e+005	73.1374	322.2796
0.76	108.5730	**35.5272**	0.0228	0.60	1.8762e+007	102.6281	877.2902

Linearly-Decreasing α							
$\alpha_1 \to \alpha_2$	Rosenbrock	Rastrigin	Griewank	$\alpha_1 \to \alpha_2$	Rosenbrock	Rastrigin	Griewank
$1.2 \to 0.6$	199.7309	39.4630	0.0243	$1.0 \to 0.4$	113.5787	30.6210	0.0391
$1.2 \to 0.5$	157.9285	33.7100	0.0422	$1.0 \to 0.3$	138.9362	31.6721	0.0667
$1.2 \to 0.4$	192.3246	31.8872	0.0709	$0.9 \to 0.6$	106.3916	30.2193	**0.0186**
$1.2 \to 0.3$	159.6439	34.1352	0.1024	$0.9 \to 0.5$	124.7271	31.0468	0.0246
$1.1 \to 0.6$	210.3639	35.8151	0.0233	$0.9 \to 0.4$	149.4981	31.4059	0.0442
$1.1 \to 0.5$	165.3690	30.8869	0.0343	$0.9 \to 0.3$	177.1897	35.1173	0.0885
$1.1 \to 0.4$	157.2430	32.1224	0.0491	$0.8 \to 0.6$	179.4460	32.9276	0.0226
$1.1 \to 0.3$	184.4184	33.5440	0.0697	$0.8 \to 0.5$	169.4544	36.3398	0.0471
$1.0 \to 0.6$	74.5490	33.7418	0.0190	$0.8 \to 0.4$	206.0009	38.5983	0.1958
$1.0 \to 0.5$	**88.0494**	**29.9218**	0.0208	$0.8 \to 0.3$	400.7041	37.8844	0.5162

QPSO. The results for α outside the range $[0.4, 0.8]$ were of poor quality and are not listed in the table. As shown in Table 3, α were set in the range of 0.5 to 0.6 when the fixed-value method was used. The best results were obtained by setting $\alpha = 0.54$. On the other hand, the algorithm exhibited the best performance when α decreased linearly from 0.6 to 0.5 for the linear time-varying method.

4.4 Performance Comparison

To determine whether QPSO can be as effective as other variants of PSO, PSO with inertia weight (PSO-In; Shi and Eberhart, 1998a, 1998b, 1999), PSO with constriction factor (PSO-Co; Clerc, 1999; Clerc and Kennedy, 2002), the standard PSO (Bratton and Kennedy, 2007), Gaussian PSO (Secrest and Lamont, 2003), Gaussian bare-bones PSO (Gaussian BBPSO; Kennedy, 2003, 2004), PSO with the exponential distribution (PSO-E; Krohling and Coelho, 2006), Lévy PSO (Richer and Blackwell, 2006), and QPSO were all compared by running a series of experiments on the first 12 functions from the CEC2005 benchmark suite (Suganthan et al., 2005). Functions F_1 to F_6 are unimodal, while functions F_7 to F_{12} are multimodal. Each algorithm was run 100 times on each problem using 20 particles to search the global best fitness value. In each run, the particles in the algorithm started in new and randomly-generated positions, which were uniformly distributed within the search bounds. Each run of each algorithm lasted 3,000 iterations and the best fitness value (objective function value) for each run was recorded.

Table 3: Mean best fitness values obtained by QPSO-Type 2-II on three benchmark functions.

Fixed α							
α	Rosenbrock	Rastrigin	Griewank	α	Rosenbrock	Rastrigin	Griewank
0.80	7.7628e+05	195.3315	2.8068	0.55	167.7840	45.7357	0.0193
0.70	176.5589	165.4325	0.2060	0.54	105.7474	**42.4817**	0.0163
0.65	**93.9793**	132.4897	**0.0162**	0.53	111.5862	47.1062	0.0198
0.60	107.5629	71.4177	0.0206	0.52	182.5722	53.4841	0.0225
0.59	120.8616	64.2002	0.0195	0.51	176.2885	51.7469	0.0296
0.58	145.5502	47.8130	0.0196	0.50	173.9438	56.2267	0.0217
0.57	145.3949	45.4865	0.0177	0.45	1.1344e+05	85.6050	169.9241
0.56	126.0142	46.9918	0.0232	0.40	9.3422e+07	113.9682	1.2831e+03

Linearly-Decreasing α							
$\alpha_1 \to \alpha_2$	Rosenbrock	Rastrigin	Griewank	$\alpha_1 \to \alpha_2$	Rosenbrock	Rastrigin	Griewank
$1.0 \to 0.5$	270.1094	123.8036	0.0366	$0.8 \to 0.3$	**81.5586**	44.3686	0.0234
$1.0 \to 0.4$	205.8190	54.4520	**0.0152**	$0.8 \to 0.2$	293.5975	47.4461	0.0202
$1.0 \to 0.3$	285.8976	48.9965	0.0549	$0.7 \to 0.5$	131.1352	69.5491	0.0166
$1.0 \to 0.2$	554.4213	46.6750	0.0341	$0.7 \to 0.4$	176.7990	42.3857	0.0193
$0.9 \to 0.5$	185.8477	109.5310	0.0173	$0.7 \to 0.3$	131.4735	45.0470	0.0164
$0.9 \to 0.4$	206.1129	49.5677	0.0165	$0.7 \to 0.2$	234.9833	43.7979	0.0343
$0.9 \to 0.3$	142.1582	49.8265	0.0315	$0.6 \to 0.5$	89.6543	43.8327	0.0150
$0.9 \to 0.2$	264.8597	44.7424	0.0579	$0.8 \to 0.7$	157.4971	**41.0110**	0.0208
$0.8 \to 0.5$	187.7583	92.9885	0.0118	$0.6 \to 0.3$	209.7387	42.6532	0.0190
$0.8 \to 0.4$	127.4620	44.5099	0.0143	$0.6 \to 0.2$	306.1035	46.4197	0.0394

For the three types of QPSO, both methods of controlling α were used and the parameters for each case were set as those values that generated the best results in the previous experiments. The parameter configurations of other PSO variants were the same as those recommended by the existing publications. For PSO-In, Shi and Eberhart (1998b) showed that the PSO-In with linearly decreasing inertia weight performs better than the one with fixed inertia weight. They varied the inertia weight linearly from 0.9 to 0.4 over the course of the run and fixed the acceleration coefficients (c_1 and c_2) at 2.0 in their empirical study (Shi and Eberhart, 1999). For PSO-Co, Clerc and Kennedy found that the values of the constriction factor χ and acceleration coefficients (c_1 and c_2) need to satisfy some constraints in order for the particle's trajectory to converge without a restriction on velocity (Clerc, 1999; Clerc and Kennedy, 2002). They recommended using a value of 4.1 for the sum of c_1 and c_2, which results in a value of the constriction factor $\chi = 0.7298$ and $c_1 = c_2 = 2.05$. Eberhart and Shi (2000) also used these values of the parameters when comparing the performance of PSO-Co with that of PSO-In. In our experiments, we also employed these parameter values for PSO-In and PSO-Co, although they may not be optimal. For the standard PSO, the LBEST ring topology was used with other parameters set as those in PSO-Co (Bratton and Kennedy, 2007). For PSO-E, except for the population size and the maximum number of iterations, all the other parameters were configured as those in Krohling and Coelho (2006). The parameter configurations for Gaussian PSO, Gaussian BBPSO, and Lévy PSO were the same as those in Secrest and Lamont (2003), in Kennedy (2003), and in Richer and Blackwell (2006), respectively.

The mean best fitness value and standard deviation out of 100 runs of each algorithm on each problem is presented in Table 4. To investigate whether the differences in mean best fitness values between algorithms were significant, the mean values for each problem were analyzed using a multiple comparison procedure, an ANOVA with 0.05 as the level of significance. Unlike Tukey's honestly significant (THS) difference test used in Richer and Blackwell (2006), the procedure employed in this work is called a stepdown procedure which takes into account that all but one of the comparisons are less extreme than the range. When doing all pairwise comparisons, this approach is the best available if confidence intervals are not needed and sample sizes are equal (Day and Quinn, 1989).

The algorithms were ranked to determine which algorithm could reliably be said to be the most effective for each problem. The algorithms that were not statistically different from each other were given the same rank; those that were not statistically different from more than one other group of algorithms were ranked with the best-performing of these groups. For each algorithm, the resulting rank for each problem and the total rank are shown in Table 5.

For the Shifted Sphere Function (F_1), QPSO-Type 2-II with either fixed or time-varying α generated better results than other methods. The results for Shifted Schwefel's Problem 1.2 (F_2) show that QPSO-Type 2-II with fixed α and PSO-Co got the best results but the performances of PSO-In and QPSO-Type 2-I with linearly decreasing α were inferior to those of the other competitors. For the Shifted Rotated High Conditioned Elliptic Function (F_3), the QPSO-Type 2-I with fixed α seemed to outperform the other methods at a level of statistical significance. QPSO-Type 2-II showed to be the winner among all the tested algorithms for Shifted Schwefel's Problem 1.2 with Noise in Fitness (F_4). F_5 is Schwefel's Problem 2.6 with Global Optimum on the Bounds, and for this benchmark, QPSO-Type 2-I with linearly decreasing α yielded the best results. For benchmark F_6, Shifted Rosenbrock Function, the performance of PSO-In and PSO-E was inferior to those of the other algorithms, among which there was no statistically significant difference. The results for the Shifted Rotated Griewank's Function without Bounds (F_7) suggest that QPSO-Type 2-II, either with fixed or with time-varying α, was able to find the solution for the function with the best quality among all the algorithms. Benchmark F_8 is Shifted Rotated Ackley's Function with Global Optimum on the Bounds. All the QPSO methods except QPSO-Type 1 with time-varying α, along with PSO-Co, showed a better performance for this problem than the others. As far as Shifted Rastrigin's Function (F_9) is concerned, the QPSO-Type 2-I with linearly decreasing α yielded the best result, and the results obtained by PSO-In and all other QPSO-based methods rank tied for the second best. F_{10} is Shifted Rotated Rastrigin's Function which appears to be a more difficult problem than F_9. For this benchmark, QPSO-Type 2-II, QPSO-Type 1 with time-varying α and the standard PSO outperformed the other competitors in a statistically significant manner. The best results for Shifted Rotated Weierstrass Function (F_{11}) were obtained by QPSO-Type 2 when α was fixed during the search process. When searching the optima of Schewefel's Problem 2.13 (F_{12}), QPSO-Type 2-II, whether it employed fixed α or time-varying α, was found to have the best performance.

As shown by the total ranks listed in Table 5, the methods based on QPSO-Type 2 obtained a better overall performance than all the other tested algorithms. For each of the benchmark functions, their performance were as good as or significantly better than that of the other algorithms. It is also revealed by the total ranks that QPSO-Type 2-II performed slightly better than QPSO-Type 2-I. However, more detailed comparisons

Table 4: Mean (*SD*) of the best fitness values over 100 trial runs of different algorithms.

Algorithms	F_1	F_2	F_3	F_4	F_5	F_6
PSO-In	3.8773e-13	785.0932	3.9733e+7	1.1249e+4	6.0547e+3	263.7252
	(1.6083e-12)	(661.2154)	(4.6433e+7)	(5.4394e+3)	(2.0346e+3)	(437.4145)
PSO-Co	1.5713e-26	0.1267	8.6472e+6	1.3219e+4	7.6892e+3	123.0243
	(1.4427e-25)	(0.3796)	(9.1219e+6)	(6.0874e+3)	(2.3917e+3)	(266.2520)
Standard PSO	8.2929e-26	78.2831	6.6185e+6	1.3312e+4	6.2884e+3	153.5178
	(1.2289e-25)	(52.3272)	(3.0124e+6)	(4.1076e+3)	(1.4318e+3)	(246.1049)
Gaussian PSO	7.3661e-26	0.0988	1.1669e+7	2.3982e+4	8.0279e+3	150.7872
	(5.9181e-25)	(0.3362)	(2.5153e+7)	(1.2512e+4)	(2.3704e+3)	(303.3368)
Gaussian	1.7869e-25	16.8751	7.7940e+6	1.1405e+4	9.5814e+3	144.1377
BBPSO	(8.4585e-25)	(16.2021)	(4.3240e+6)	(6.7712e+3)	(3.0227e+3)	(165.2616)
PSO-E	5.2531e-24	20.2750	6.2852e+6	8.2706e+3	7.2562e+3	189.8292
	(2.2395e-23)	(15.2414)	(2.8036e+6)	(3.6254e+3)	(1.8666e+3)	(375.8636)
Lévy PSO	1.1880e-24	36.9986	1.7366e+07	7.4842e+3	8.2543e+3	133.9526
	(1.1455e-23)	(29.1360)	(1.9001e+7)	(6.6588e+3)	(2.2297e+3)	(293.8460)
QPSO-Type 1	3.5936e-28	40.2282	4.8847e+6	6.2397e+3	8.0749e+3	138.0746
($\alpha = 1.00$)	(1.5180e-28)	(23.3222)	(2.1489e+6)	(2.4129e+3)	(1.7099e+3)	(209.1735)
QPSO-Type 1	5.0866e-29	4.5003	3.2820e+6	6.4303e+3	7.8471e+3	139.9815
($\alpha = 1.0 \rightarrow 0.9$)	(4.4076e-29)	(2.9147)	(1.9953e+6)	(2.9744e+3)	(1.7878e+3)	(206.8138)
QPSO-Type 2-I	1.9838e-27	0.1771	**1.6559e+6**	3.1321e+3	5.7853e+3	**82.9908**
(α =0.75)	(5.2716e-28)	(0.1137)	**(7.1264e+5)**	(2.0222e+3)	(1.2483e+3)	**(119.836)**
QPSO-Type 2-I	1.2672e-27	120.6051	4.4257e+6	4.0049e+3	**3.3684e+3**	88.0494
($\alpha = 1.0 \rightarrow 0.5$)	(3.7147e-28)	(62.2340)	(2.3302e+6)	(2.7218e+3)	**(975.6551)**	(159.7481)
QPSO-Type 2-II	**3.1554e-36**	**0.0715**	1.8544e+6	3.1443e+3	5.7144e+3	105.7474
($\alpha = 0.54$)	**(2.3913e-36)**	**(0.0530)**	(6.4710e+5)	(3.8785e+3)	(1.4898e+3)	(155.4583)
QPSO-Type 2-II	2.6728e-35	1.4099	2.1737e+6	**2.1835e+3**	4.3398e+3	89.6543
($\alpha = 0.6 \rightarrow 0.5$)	(6.5932e-35)	(7.8582)	(1.0089e+6)	**(2.8487e+3)**	(1.4313e+3)	(151.6908)
Algorithms	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}
PSO-In	0.9907	0.0414	39.5528	239.5814	41.0529	3.6785e+4
	(4.7802)	(0.2393)	(16.1654)	(72.2521)	(6.0318)	(4.0943e+4)
PSO-Co	0.0255	5.1120	96.7296	171.6488	36.0339	9.9648e+3
	(0.0327)	(4.5667)	(28.0712)	(58.5713)	(7.2659)	(1.6158e+4)
Standard PSO	0.0218	0.2744	79.1219	128.9865	30.3424	1.8178e+4
	(0.0165)	(0.6795)	(20.2619)	(32.3662)	(2.7409)	(1.4866e+4)
Gaussian PSO	0.0224	2.7722	103.6245	184.2657	33.5448	6.8875e+4
	(0.0178)	(1.4603)	(28.6113)	(57.3675)	(6.5823)	(6.5610e+4)
Gaussian	0.0205	3.5460	80.9496	164.2914	29.8088	3.4327e+4
BBPSO	(0.0208)	(6.1929)	(22.0621)	(72.8542)	(3.2671)	(6.2435e+4)
PSO-E	0.0493	3.5881	66.5112	163.7187	29.2666	1.7161e+4
	(0.0538)	(5.5286)	(20.9853)	(55.0921)	(3.2083)	(1.0862e+4)
Lévy PSO	0.0446	2.2168	74.0446	154.3838	28.9923	1.6282e+4
	(0.1182)	(1.3575)	(21.6913)	(76.3070)	(5.0212)	(2.5184e+4)
QPSO-Type 1	0.0218	0.1217	56.4232	137.0334	28.2096	1.2145e+4
($\alpha = 1.00$)	(0.0204)	(0.4504)	(16.7090)	(38.5269)	(3.0216)	(9.7844e+3)
QPSO-Type 1	0.0209	0.0916	54.4278	126.1298	29.4137	1.0576e+4
($\alpha = 1.0 \rightarrow 0.9$)	(0.0203)	(0.3166)	(16.6044)	(44.9531)	(2.8907)	(9.0572e+3)
QPSO-Type 2-I	0.0203	0.0683	39.0991	128.5351	19.8616	7.2794e+3
($\alpha = 0.75$)	(0.0164)	(0.3080)	(12.4904)	(57.6255)	(7.0620)	(8.2210e+3)
QPSO-Type 2-I	0.0208	2.0961e-14	**29.9218**	**118.4549**	28.1887	1.2938e+4
($\alpha = 1.0 \rightarrow 0.5$)	(0.0130)	(1.9099e-14)	**(10.5736)**	**(53.0216)**	(6.2233)	(1.3787e+4)
QPSO-Type 2-II	0.0163	0.0762	42.4817	185.6351	**19.0976**	**4.6519e+3**
($\alpha = 0.54$)	(0.0134)	(0.3075)	(12.1384)	(46.6356)	**(6.7920)**	**(4.3177e+3)**
QPSO-Type 2-II	**0.0150**	**7.5318e-15**	43.8327	207.0548	23.0303	5.6134e+3
($\alpha = 0.6 \rightarrow 0.5$)	**(0.0119)**	**(1.7046e-15)**	(17.881)	(14.4658)	(8.7874)	(4.6511e+3)

Table 5: Ranking by algorithms and problems.

Algorithms	F_1	F_2	F_3	F_4	F_5	F_6	F_7	F_8	F_9	F_{10}	F_{11}	F_{12}	Total rank
PSO-In	13	13	13	=9	=3	=12	13	=1	=2	13	13	=11	116
PSO-Co	=6	=1	=9	=11	=7	=1	=3	13	=12	=9	12	=3	87
Standard PSO	=6	11	=7	=11	=3	=1	=3	=7	=9	=1	10	=9	78
Gaussian PSO	=6	=1	=9	13	10	=1	=3	=9	=12	=9	11	13	97
Gaussian BBPSO	=10	=7	=9	=9	13	=1	=3	=9	=9	=6	=4	=11	91
PSO-E	12	=7	=7	8	=7	=12	=11	=9	8	=6	=4	=9	100
Lévy PSO	=10	=9	12	=5	=10	=1	=11	=9	=9	=6	=4	=5	91
QPSO-Type 1 ($\alpha = 1.00$)	4	=9	=5	=5	=10	=1	=3	=7	=6	5	=4	=5	64
QPSO-Type 1 ($\alpha = 1.0 \rightarrow 0.9$)	3	6	4	=5	=7	=1	=3	=1	=6	=1	=4	=5	46
QPSO-Type 2-I ($\alpha = 0.75$)	=6	=4	1	=2	=3	=1	=3	=1	=2	=1	=1	=3	28
QPSO-Type 2-I ($\alpha = 1.0 \rightarrow 0.5$)	5	12	=5	=2	1	=1	=3	=1	1	=1	=4	=5	41
QPSO-Type 2-II ($\alpha = 0.54$)	1	=1	2	=2	=3	=1	=1	=1	=2	=9	=1	=1	25
QPSO-Type 2-II ($\alpha = 0.6 \rightarrow 0.5$)	2	=4	3	1	2	=1	=1	=1	=2	12	3	=1	33

reveal that for F_{10}, the results obtained by QPSO-Type 2-II had poorer quality than those by QPSO-Type 2-I even than those by other PSO variants, and that the QPSO-Type 2-I with fixed α had the most stable performance across all of the benchmark functions with the worst rank being 6 for F_6. It can also be observed that for either version of QPSO-Type 2, time-varying α led the algorithm to an overall performance no better than for fixed α, particularly for QPSO-Type 2-I.

The second best performing algorithm was the QPSO-Type 1 algorithm as indicated by the total ranks. It can be seen that there is a great deal of difference between the performance of the two parameter control methods. However, in contrast with QPSO-Type 2, QPSO-Type 1 appeared to work better by using the time-varying control method. The standard PSO was the next best algorithm, and it enhances the search ability of PSO by incorporating the lbest ring topology into PSO-Co, which was ranked fourth in overall performance. It is evident from the ranking list that the standard PSO and PSO-Co were both a great improvement over the PSO-In algorithm. The other four probabilistic algorithms did not work so effectively as the QPSO algorithms and the standard PSO. Among the four methods, Gaussian BBPSO and Lévy PSO showed better search ability than PSO-E and Gaussian PSO.

5 Conclusion

In this paper, after describing the background of the QPSO algorithm, which belongs to the bare-bones PSO family, we theoretically analyzed the behavior of the individual particle in QPSO. Then, through empirical studies on a well-known benchmark suite, we provided guidelines for parameter selection for the algorithm, as well as the performance comparison between QPSO and some other forms of PSO.

Two types of particles, corresponding to two search strategies of the QPSO algorithm, were analyzed by using the theory of probability measure. We derived the sufficient and necessary condition for the position of a single particle to be convergent or probabilistically bounded. Since the behavior of the particle is determined by the CE coefficient α, the derivation of the condition is reduced to find out how to select the value of α to guarantee the convergence or probabilistic boundedness of the particle's position. For a Type-1 particle, if $\alpha < e^{\gamma}$, the position sequence of the particle converges to its local attractor in probability (or almost surely, in distribution, in mean of order r); if $\alpha = e^{\gamma}$, the position is probabilistically bounded; otherwise, the position diverges. For a Type-2 particle, if $\alpha \leq e^{\gamma}$, the position is probabilistically bounded, or else it diverges. Therefore, $e^{\gamma} \approx 1.781$ provides an upper bound for the value of α selected when the QPSO algorithm is used to solve real-world applications.

For further investigation of the parameter selection issue in QPSO, based on the presented theoretical analysis, two methods of controlling α were proposed, one employing fixed-value α and the other a time-varying one. The methods were tested by using different parameter settings on several benchmark functions from a well-known set of test problems proposed at CEC2005. The parameter settings resulting in the best performance were identified by ranking the generated results and thus were used in QPSO for performance comparison with other variants of PSO on 12 test problems from the CEC2005 benchmark. It was shown that QPSO, particularly QPSO-Type 2, is comparable with or even better than other forms of PSO in finding the optimal solutions of the tested benchmark functions.

The search mechanism of the individual particle in QPSO has been theoretically investigated in this paper, and it provides a solid foundation for different applications of QPSO. However, how to control and determine the values of parameters of QPSO to further improve the performance of the algorithm is still a challenging problem. Furthermore, investigation into the interaction between particles is an immediate task of our future research on the mechanism of QPSO.

Acknowledgments

This work is supported by the National Natural Science Foundation of China (Project Number: 61170119, 60973094), the Natural Science Foundation of Jiangsu Province, China (Project Number: BK2010143), by the Fundamental Research Funds for the Central Universities (Project Number: JUSRP21012, JUSRP11023), and by the Foundation of Key Laboratory of Advanced Process Control for Light Industry (Jiangnan University), Ministry of Education, Peoples' Republic of China.

References

Angeline, P. J. (1998a). Evolutionary optimization versus particle swarm optimization: Philosophy and performance differences. In *Proceedings of the 7th International Conference on Evolutionary Programming*, pp. 601–610.

Angeline, P. J. (1998b). Using selection to improve particle swarm optimization. In *Proceedings of IEEE International Conference on Evolutionary Computation*, pp. 84–89.

Bratton, D., and Kennedy, J. (2007). Defining a standard for particle swarm optimization. In *Proceedings of IEEE Swarm Intelligence Symposium*, pp. 120–127.

Cai, Y., Sun, J., Wang, J., Ding, Y., Tian, N., Liao, X., and Xu, W. (2008). Optimizing the codon usage of synthetic gene with QPSO algorithm. *Journal of Theoretical Biology*, 254(1):123–127.

Chen, W., Sun, J., Ding, Y., Fang, W., and Xu, W. (2008). Clustering of gene expression data with quantum-behaved particle swarm optimization. In *Proceedings of the Twenty-First International Conference on Industrial, Engineering & Other Applications of Applied Intelligent Systems (IEA/AIE 2008)*, pp. 388–396.

Clerc, M. (1999). The swarm and the queen: Towards a deterministic and adaptive particle swarm optimization. In *Proceedings of Congress on Evolutionary Computation*, pp. 1951–1957.

Clerc, M., and Kennedy, J. (2002). The particle swarm-explosion, stability and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, 6(2):58–73.

Coelho, L. (2008). A quantum particle swarm optimizer with chaotic mutation operator. *Chaos, Solitons & Fractals*, 37(5):1409–1418.

Coelho, L. (2010). Gaussian quantum-behaved particle swarm optimization approaches for constrained engineering design problems. *Expert Systems with Applications*, 37(2):1676–1683.

Coelho, L., and Alotto, P. (2008). Global optimization of electromagnetic devices using an exponential quantum-behaved particle swarm optimizer. *IEEE Transactions on Magenetics*, 44(6):1074–1077.

Coelho, L., and Mariani, V. (2008). Particle swarm approach based on quantum mechanics and harmonic oscillator potential well for economic load dispatch with valve-point effects. *Energy Conversion and Management*, 49(11):3080–3085.

Cohen-Tannoudji, C., Diu, B., and Laloe, F. (1997). *Quantum mechanics*, Vol. 1. New York: John Wiley.

Courant, R. (1989). *Introduction to calculus and analysis*. Berlin: Springer-Verlag.

Day, R. W., and Quinn, G. P. (1989). Comparisons of treatments after an analysis of variance in ecology. *Ecological Monographs*, 59:433–463.

Dudley, R. M. (2003). *Real analysis and probability* (2nd ed.). Cambridge, UK: Cambridge University Press.

Eberhart, R. C., and Kennedy, J. (1995). A new optimizer using particle swarm theory. In *Proceedings of the Sixth International Symposium on Micro Machine and Human Science*, pp. 39–43.

Eberhart, R. C., and Shi, Y. (2000). Comparing inertia weights and constriction factors in particle swarm optimization. In *Proceedings of the 2000 Congress on Evolutionary Computation (CEC '00)*, Vol. 1, pp. 84–88.

Emara, H. M., and Fattah, H. A. A. (2004). Continuous swarm optimization technique with stability analysis. In *Proceedings of American Control Conference*, pp. 2811–2817.

Gao, F. (2008). Parameters estimation on-line for Lorenz system by a novel quantum-behaved particle swarm optimization. *Chinese Physics B*, 17(4):1196–1201.

Gavi, V., and Passino, K. M. (2003). Stability analysis of social foraging swarms. *IEEE Transactions on Systems, Man, and Cybernetics*, 34(1):539–557.

Huang, Z., Wang, Y., Yang, C., and Wu, C. (2009). A new improved quantum-behaved particle swarm optimization model. In *Proceedings of the 5th IEEE Conference on Industrial Electronics and Applications*, pp. 1560–1564.

Janson, S., and Middendorf, M. (2005). A hierarchical particle swarm optimizer and its adaptive variant. *IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics*, 35(6):1272–1282.

Jiang, M., Luo, Y., and Yang, S. (2007). Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm. *Information Processing Letters*, 102:8–16.

Kadirkamanathan, V., Selvarajah, K., and Fleming, P. J. (2006). Stability analysis of the particle dynamics in particle swarm optimizer. *IEEE Transactions on Evolutionary Computation*, 10:245–255.

Kennedy, J. (1998). The behavior of particles. In *Proceedings of 7th Annual Conference on Evolutionary Programming*, pp. 581–589.

Kennedy, J. (1999). Small worlds and mega-minds: Effects of neighborhood topology on particle swarm performance. In *Proceedings of Congress on Evolutionary Computation*, pp. 1931–1938.

Kennedy, J. (2002). Stereotyping: Improving particle swarm performance with cluster analysis. In *Proceedings of Congress on Computational Intelligence*, pp. 1671–1676.

Kennedy, J. (2003). Bare bones particle swarms. In *Proceedings of IEEE Swarm Intelligence Symposium*, pp. 80–87.

Kennedy, J. (2004). Probability and dynamics in the particle swarm. In *Proceedings of Congress on Evolutionary Computation*, pp. 340–347.

Kennedy, J. (2006). In search of the essential particle swarm. In *Proceedings of IEEE World Congress on Computational Intelligence*, pp. 1694–1701.

Kennedy, J., and Eberhart, R. C. (1995). Particle swarm optimization. In *Proceedings of IEEE International Conference on Neural Networks*, pp. 1942–1948.

Krohling, R. A. (2004). Gaussian swarm: A novel particle swarm optimization algorithm. In *Proceedings of IEEE Conference on Cybernetics and Intelligent Systems*, pp. 372–376.

Krohling, R. A., and Coelho, L. (2006). PSO-E: Particle swarm with exponential distribution. In *Proceedings of IEEE Congress on Evolutionary Computation*, pp. 1428–1433.

Lei, X., and Fu, A. (2008). Two-dimensional maximum entropy image segmentation method based on quantum-behaved particle swarm optimization algorithm. In *Proceedings of the Fourth International Conference on Nature Computation*, pp. 692–696.

Li, S.-Y., Wang, R.-G., Hu, W.-W., and Sun, J.-Q. (2007). A new QPSO based BP neural network for face detection. In B.-Y. Cao (Ed.), *Fuzzy information and engineering* (pp. 355–363). Berlin: Springer-Verlag.

Liang, J., and Suganthan, P. (2005). Dynamic multiswarm particle swarm optimizer (DMS-PSO). In *Proceedings of IEEE Swarm Intelligence Symposium*, pp. 124–129.

Liu, J., Xu, W., and Sun, J. (2005). Quantum-behaved particle swarm optimization with mutation operator. In *Proceedings of the 17th IEEE International Conference on Tools with Artificial Intelligence*, pp. 237–240.

Mendes, R., Kennedy, J., and Neves, J. (2004). The fully informed particle swarm: Simpler, maybe better. *IEEE Transactions on Evolutionary Computation*, 8(3):204–210.

Mikki, S., and Kishk, A. (2006). Quantum particle swarm optimization for electromagnetics. *IEEE Transactions on Antennas and Propagation*, 54(10):2764–2775.

Mohais, A., Mendes, R., Ward, C., and Postoff, C. (2005). Neighborhood re-structuring in particle swarm optimization. In *Proceedings of 18th Australian Joint Conference on Artificial Intelligence*, pp. 776–785.

Omkara, S., Khandelwala, R., Ananthb, T., Naika, G. N., and Gopalakrishnana, S. (2009). Quantum behaved particle swarm optimization (QPSO) for multi-objective design optimization of composite structures. *Expert Systems with Applications*, 36(8):11312–11322.

Ozcan, E., and Mohan, C. K. (1999). Particle swarm optimization: Surfing the waves. In *Proceedings of 1999 IEEE Congress on Evolutionary Computation*, pp. 1939–1944.

Pant, M., Thangaraj, R., and Abraham, A. (2008). A new quantum behaved particle swarm optimization. In *Proceedings of the 10th Annual Conference on Genetic and Evolutionary Computation*, pp. 87–94.

Pant, M., Thangaraj, R., and SinghSobol, V. P. (2009). Sobal mutated quantum particle swarm optimization. *International Journal of Recent Trends in Engineering*, 1(1): 95–99.

Parrott, D., and Li, X. (2006). Locating and tracking multiple dynamic optima by a particle swarm model using speciation. *IEEE Transactions on Evolutionary Computation*, 10(4):440–458.

Poli, R. (2008). Dynamics and stability of the sampling distribution of particle swarm optimisers via moment analysis. *Journal of Artificial Evolution and Applications*, Article ID 761459.

Richer, T. J., and Blackwell, T. M. (2006). The Levy particle swarm. In *Proceedings of Congress on Evolutionary Computation*, pp. 808–815.

Sabata, S., Coelho, L. S., and Abrahamc, A. (2009). MESFET DC model parameter extraction using quantum particle swarm optimization. *Microelectronics Reliability*, 49(6):660–666.

Secrest, B., and Lamont, G. (2003). Visualizing particle swarm optimization-Gaussian particle swarm optimization. In *Proceedings of IEEE Swarm Intelligence Symposium*, pp. 198–204.

Shi, Y., and Eberhart, R. C. (1998a). A modified particle swarm optimizer. In *Proceedings of IEEE International Conference on Evolutionary Computation*, pp. 69–73.

Shi, Y., and Eberhart, R. C. (1998b). Parameter selection in particle swarm optimization. In *Proceedings of 7th Conference on Evolutionary Programming VII (EP '98)*, pp. 591–600.

Shi, Y., and Eberhart, R. C. (1999). Empirical study of particle swarm optimization. In *Proceedings of the 1999 Congress on Evolutionary Computation (CEC '99)*, Vol. 3, pp. 1945–1950.

Shiryayev, A. (1984). *Probability*. Berlin: Springer-Verlag.

Suganthan, P. (1999). Particle swarm optimizer with neighborhood operator. In *Proceedings of Congress on Evolutionary Computation*, pp. 1958–1962.

Suganthan, P., Hansen, N., Liang, J. J., Deb, K., Chen, Y.-P., Auger, A., and Tiwari, S. (2005). Problem definitions and evaluation criteria for the CEC 2005 special session on real-parameter optimization. Technical report, Nanyang Technological University, Singapore, May 2005 and KanGAL Report #2005005, IIT Kanpur, India.

Sun, C., and Lu, S. (2010). Short-term combined economic emission hydrothermal scheduling using improved quantum-behaved particle swarm optimization. *Expert Systems with Applications*, 37(6):4232–4241.

Sun, J., Fang, W., Wang, D., and Xu, W. (2009). Solving the economic dispatch problem with a modified quantum-behaved particle swarm optimization method. *Energy Conversion and Management*, 50(12):2967–2975.

Sun, J., Feng, B., and Xu, W. (2004). Particle swarm optimization with particles having quantum behavior. In *Proceedings of Congress on Evolutionary Computation*, pp. 326–331.

Sun, J., Liu, J., and Xu, W. (2007). Using quantum-behaved particle swarm optimization algorithm to solve non-linear programming problems. *International Journal of Computer Mathematics*, 84(2):261–272.

Sun, J., Xu, W., and Feng, B. (2004). A global search strategy of quantum-behaved particle swarm optimization. In *Proceedings of IEEE Conference on Cybernetics and Intelligent Systems*, pp. 111–116.

Sun, J., Xu, W., and Feng, B. (2005). Adaptive parameter control for quantum-behaved particle swarm optimization on individual level. In *Proceedings of IEEE International Conference on Systems, Man, and Cybernetics*, pp. 3049–3054.

Sun, J., Xu, W., and Ye, B. (2006). Quantum-behaved particle swarm optimization clustering algorithm. In *Proceedings of the International Conference on Advanced Data Mining and Applications*, pp. 340–347.

Trelea, I. C. (2003). The particle swarm optimization algorithm: Convergence analysis and parameter selection. *Information Processing Letters*, 85(6):317–325.

van den Bergh, F. (2002). *An analysis of particle swarm optimizers*. Ph.D. dissertation, University of Pretoria, Pretoria, South Africa.

van den Bergh, F., and Engelbrecht, A. P. (2004). A cooperative approach to particle swarm optimization. *IEEE Transactions on Evolutionary Computation*, 8(3):225–239.

Wang, J., and Zhou, Y. (2007). Quantum-behaved particle swarm optimization with generalized local search operator for global optimization. In *Proceedings of International Conference on Intelligent Computing*, pp. 344–352.

Xi, M., Sun, J., and Xu, W. (2008). An improved quantum-behaved particle swarm optimization algorithm with weighted mean best position. *Applied Mathematics and Computation*, 205(2):751–759.

Zhang, Z. (2010). Quantum-behaved particle swarm optimization algorithm for economic load dispatch of power system. *Expert Systems with Applications*, 37(2):1800–1803.

Appendix

LEMMA 2: *If $\{u_n\}$ is a sequence of independent identically distributed random variables with $u_n \sim U(0,1)$ for all $n > 0$ and $\zeta_n = \ln[\ln(1/u_n)]$, then*

$$\frac{1}{n}\sum_{i=1}^{n}\zeta_i \xrightarrow{P} -\gamma. \tag{A1}$$

PROOF: Similar to the proof of Lemma 1, by Theorem 5 (Khintchine's weak law of large numbers), we find that the lemma holds. □

THEOREM 9: *The necessary and sufficient condition for the position sequence of a Type-1 particle $\{X_n\}$ to converge to p in probability ($X_n \xrightarrow{P} p$) is that $\alpha < e^{\gamma}$.*

PROOF: PROOF OF NECESSITY

(i) If $X_n \xrightarrow{P} p$, then for every $\varepsilon > 0$, $\lim_{n\to\infty} P\{X_n - p < \varepsilon\} = 1$ or $\lim_{n\to\infty} P\left\{\alpha^n \prod_{i=1}^{n} \ln(1/u_i) < \varepsilon\right\} = 1$, which is equivalent to the proposition that $\forall m \in Z^+$,

$$\lim_{n\to\infty} P\left\{\ln\alpha + \frac{1}{n}\sum_{i=1}^{n}\zeta_i < -\frac{m}{n}\right\} = 1,$$

that is, $\forall m \in Z^+$ and $\forall \delta > 0, \exists K_1 \in Z^+$ such that whenever $k \geq K_1$,

$$P\left\{\ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < -\frac{m}{k}\right\} > 1 - \frac{\delta}{2}. \tag{A2}$$

(ii) Lemma 2 indicates that $\frac{1}{n}\sum_{i=1}^{n}\zeta_i \xrightarrow{P} -\gamma$, so $\forall m \in Z^+$ and $\forall \delta > 0, \exists K_2 \in Z^+$ such that whenever $k \geq K_2$,

$$P\left\{\left|\frac{1}{k}\sum_{i=1}^{k}\zeta_i + \gamma\right| < \frac{1}{m}\right\} > 1 - \frac{\delta}{2}$$

or

$$P\left\{\ln\alpha - \gamma - \frac{1}{m} < \ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < \ln\alpha - \gamma + \frac{1}{m}\right\} > 1 - \frac{\delta}{2}. \tag{A3}$$

Hence, by conditions (i) and (ii), $\forall m \in Z^+$ and $\forall \delta > 0, \exists K = \max(K_1, K_2)$ such that whenever $k \geq K$, Equations (A2) and (A3) are satisfied simultaneously. Let

$$A = \left\{\ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < -\frac{m}{k}\right\}$$

and

$$B = \left\{\ln\alpha - \gamma - \frac{1}{m} < \ln\alpha + \frac{1}{k}\sum_{i=1}^{k}\zeta_i < \ln\alpha - \gamma + \frac{1}{m}\right\},$$

and thus we have from Equations (A2) and (A3) that

$$P(AB) = P(A) + P(B) - P(A \cup B) > 2 - \delta - P(A \cup B) \geq 1 - \delta. \tag{A4}$$

Since $P(AB) = P\{\ln\alpha - \gamma - \frac{1}{m} < -\frac{m}{k}\}$, we obtain that $P\{\ln\alpha - \gamma - \frac{1}{m} < -\frac{m}{k}\} > 1 - \delta$, which means that $\forall \delta > 0$,

$$P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left(\ln\alpha < \gamma + \frac{1}{m} - \frac{m}{k}\right)\right\} > 1 - \delta,$$

from which we immediately have $P\{\ln\alpha < \gamma\} > 1 - \delta$. Due to the arbitrariness of δ, we have that $P\{\ln\alpha < \gamma\} = 1$, that is, $\alpha < e^{\gamma}$. This ends the proof of necessity.

PROOF OF SUFFICIENCY: If $\alpha < e^{\gamma}$, we can infer from Theorem 3 that $X_n \xrightarrow{a.s.} p$. Since $X_n \xrightarrow{a.s.} p$ implies $X_n \xrightarrow{P} p$, this ends the proof of sufficiency.

This completes the proof of the theorem. □

THEOREM 10: *The necessary and sufficient condition for the position sequence of a Type-1 particle* $\{X_n\}$ *to converge to* p *in distribution* ($X_n \xrightarrow{d} p$) *is that* $\alpha < e^{\gamma}$.

PROOF: Consider the following three possible cases.

(i) If $\alpha < e^{\gamma}$, by Theorem 8, we have $X_n \xrightarrow{P} p$. Thus by Theorem 3, we have that $X_n \xrightarrow{d} p$ and $\beta_n = \prod_{i=1}^{n} \zeta_i \xrightarrow{P} p$.

(ii) If $\alpha = e^{\gamma}$, then $\ln \alpha = \gamma$. Since $\frac{1}{n}\sum_{i=1}^{n} \zeta_i \xrightarrow{a.s.} -\gamma$,

$$P\left\{\lim_{n\to\infty}\left|\frac{1}{n}\sum_{i=1}^{n}\zeta_i + \gamma\right| = 0\right\} = 1$$

and consequently

$$P\left\{\lim_{n\to\infty}\left|\frac{1}{n}\sum_{i=1}^{n}\zeta_i + \ln\alpha\right| = 0\right\} = 1,$$

which means that $\forall m \in Z^+$, $\exists K \in Z^+$, such that whenever $k \geq K$,

$$P\left\{\left|\frac{1}{k}\sum_{i=1}^{k}\zeta_i + \ln\alpha\right| < \frac{1}{m}\right\} = 1,$$

namely

$$P\left\{\left|\sum_{i=1}^{k}\zeta_i + k\ln\alpha\right| < \frac{k}{m}\right\} = P\left\{\left|\ln\left[\alpha^k\prod_{i=1}^{k}\ln\left(\frac{1}{u_i}\right)\right]\right| < \frac{k}{m}\right\} = P\left\{|\ln\beta_k < \frac{k}{m}|\right\} = 1. \tag{A5}$$

The above proposition is equivalent to the following one that

$$P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left(|\ln\beta_k| < \frac{k}{m}\right)\right\}$$
$$= P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left(-\frac{k}{m} < \ln\beta_k < \frac{k}{m}\right)\right\} = P\left\{-\infty < \sup_n \ln\beta_n < +\infty\right\} = 1,$$

scilicet that $P\left\{0 < \sup_{n>0}\beta_n < +\infty\right\} = 1$. According to Equations (50) and (51), we have

$$P\left\{0 < \sup_{n>0}|X_n - p| < +\infty\right\} = 1, \tag{A6}$$

which means that when $n \to \infty$, X_n neither converges to p nor diverges. That means the particle can appear at any location except point p and the infinite points. Noting that the distribution function of the particle is $F_n(x) = 1 - \exp[-2|x-p|/\alpha|X_{n-1} - p|]$, we can find that the limit of $F_n(x)$ is uncertain. However, the random variable p distributes

uniformly on the interval (P, G) or (P, G), where P and G are the pbest position of the particle and the gbest position in the one-dimensional case. Thus $\lim_{n\to\infty} F_n(x) \neq F_p(x)$, implying that X_n does not converge to p in distribution.

(iii) If $\alpha > e^{\gamma}$, $\ln\alpha > \gamma$. Since $\frac{1}{n}\sum_{i=1}^{n}\zeta_i \xrightarrow{a.s.} -\gamma$, we have

$$P\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\zeta_i + \ln\alpha > 0\right\} = 1,$$

implying that $\exists b > 0$, such that

$$P\left\{\lim_{n\to\infty}\frac{1}{n}\sum_{i=1}^{n}\zeta_i + \ln\alpha = b\right\} = 1. \tag{A7}$$

And it is then easy to deduce that $\forall m \in Z^+$, $\exists K \in Z^+$, such that whenever $k \geq K$,

$$P\left\{\left|\frac{1}{k}\sum_{i=1}^{k}\zeta_i + \ln\alpha - b\right| < \frac{1}{m}\right\} = 1,$$

namely,

$$P\left\{b - \frac{1}{m} < \frac{1}{k}\sum_{i=1}^{k}\zeta_i + \ln\alpha < \frac{1}{m} + b\right\} = 1.$$

Thus we have that

$$\begin{aligned}
&P\left\{kb - \frac{k}{m} < \sum_{i=1}^{k}\zeta_i + k\ln\alpha < kb + \frac{k}{m}\right\} \\
&= P\left\{kb - \frac{k}{m} < \ln\left[\alpha^k\prod_{i=1}^{k}\ln\left(\frac{1}{u_i}\right)\right] < kb + \frac{k}{m}\right\} \\
&= P\left\{kb - \frac{k}{m} < \ln\beta_k < kb + \frac{k}{m}\right\} = 1.
\end{aligned}$$

Since both $kb - \frac{k}{m}$ and $kb + \frac{k}{m}$ are arbitrarily large real numbers, the above proposition leads to

$$P\left\{\bigcap_{m=1}^{\infty}\bigcup_{n=1}^{\infty}\bigcap_{k=n}^{\infty}\left(kb - \frac{k}{m} < \ln\beta_k < kb + \frac{k}{m}\right)\right\} = P\left\{\sup_{n>0}\beta_n = +\infty\right\} = 1.$$

Thus, according to Equations (50) and (51), we have

$$P\left\{\sup_{n>0}|X_n - p| = +\infty\right\}. \tag{A8}$$

This proves that when $n \to \infty$, $|X_n - p|$ is divergent. Since p is probabilistically bounded, we have

$$\lim_{n\to\infty} F_n(x) \neq F_p(x). \tag{A9}$$

From the above three cases, we find that the theorem holds. This completes the proof of the theorem. □

DEFINITION 10: *Let $\{\xi_n\}$ be a sequence of random variables defined on a probability space $(\Omega, \mathcal{F}, P)$. The sequence $\{\xi_n\}$ is said to be uniformly integrable if*

$$\lim_{a\to\infty} \sup_{n>0} \int_{|\xi_n|\geq a} |\xi_n|\, dP = 0. \tag{A10}$$

LEMMA 3: *A necessary and sufficient condition for the sequence of random variables $\{\xi_n\}$ to be uniformly integrable is that:*
(i) *for every $\varepsilon > 0$, $\exists \delta = \delta(\varepsilon)$, such that for the every $A \in \mathcal{F}$, whenever $P(A) < \delta$, $\int_A |\xi_n|\, dP < \varepsilon$ for all $n \geq 1$;*
(ii) $\sup_{n\geq 1} E|\xi_n| < +\infty$.

In Lemma 3, the first condition is known as the uniform absolute continuity of $\{\xi_n\}$, and the second one is the uniform boundedness of $\{\xi_n\}$.

LEMMA 4: *The sequence of random variables $\{\xi_n\}$ is uniformly integrable if there exists $\theta > 0$ such that $\sup_{n\geq 1} E|\xi_n|^{1+\theta} < +\infty$.*

PROOF: Since

$$\int_{|\xi_n|\geq a} |\xi_n|\, dP \leq a^{-\theta} E|\xi_n|^{1+\theta}$$

and

$$\sup_{n\geq 1} E|\xi_n|^{1+\theta} = c < +\infty,$$

for every $\varepsilon > 0$, we can choose a to be sufficiently large so that

$$\sup_{n\geq 1} \int_{|\xi_n|\geq a} |\xi_n|\, dP \leq ca^{-\theta} < \varepsilon.$$

By Definition 9, this implies that $\{\xi_n\}$ is uniformly integrable. □

LEMMA 5: *The sequence of random variables $\{\xi_n\}$ converges to the random variable ξ in mean of order r, namely*

$$\lim_{n\to\infty} E[|\xi_n - \xi|^r] = 0, \tag{A11}$$

if and only if $\{\xi_n\}$ *converges to* ξ *in probability and* $\{|\xi_n|^r,\ n \geq 1\}\,(r > 0)$ *is uniformly integrable.*

THEOREM 11: *The necessary and sufficient condition for the position sequence of a Type-1 particle* $\{X_n\}$ *to converge to* p *in mean of order* r $(X_n \xrightarrow{r} p,\ 0 < r < +\infty)$ *is that* $\alpha < e^{\gamma}$.

PROOF: Since the necessary and sufficient condition for $\{X_n\}$ to converge to p in probability is that $\alpha < e^{\gamma}$, from Lemma 5, we only need to prove that $\{|X_n|^r,\ n \geq 1\}\,(r > 0)$ is uniformly integrable.

It is easy to find that whenever $\alpha < e^{\gamma}$, $|X_n - p|$ is bounded, that is to say, there exists a real number $c > 0$ such that $\sup_{n\geq 1} |X_n - p| = c < +\infty$. Thus for every $n \geq 1$, $|X_n - p| < c$, that is, $|X_n| < c + |p|$. Since p is bounded, $|X_n|$ is also bounded. For every $0 < \theta < +\infty$,

$$||X_n|^r|^{1+\theta} = |X_n|^{r+r\theta} < (c + |p|)^{r+r\theta}.$$

Consequently $||X_n|^r|^{1+\theta}$ is bounded, and thus $\sup_{n\geq 1} E||X_n|^r|^{1+\theta} < +\infty$. By Lemma 4 and Lemma 5, $\{|X_n|^r,\ n \geq 1\}$ is uniformly integrable so that $X_n \xrightarrow{r} p$.

This completes the proof of the theorem. □