SCIENCE CHINA
Information Sciences

• RESEARCH PAPER •
Special Focus on Pigeon-Inspired Optimization

July 2019, Vol. 62 070207:1–070207:9
https://doi.org/10.1007/s11432-018-9753-5

Theoretical analysis of the convergence property of a basic pigeon-inspired optimizer in a continuous search space

Yushan ZHANG[1,2], Han HUANG[3*], Hongyue WU[3] & Zhifeng HAO[4]

[1]*School of Statistics and Mathematics, Guangdong University of Finance and Economics, Guangzhou* 510320*, China;*
[2]*Big Data and Educational Statistics Application Laboratory, Guangdong University of Finance and Economics, Guangzhou* 510320*, China;*
[3]*School of Software Engineering, South China University of Technology, Guangzhou* 510006*, China;*
[4]*School of Mathematics and Big Data, Foshan University, Foshan* 528000*, China*

Received 20 August 2018/Accepted 30 November 2018/Published online 6 June 2019

Abstract The pigeon-inspired optimization (PIO) algorithm is a newly presented swarm intelligence optimization algorithm inspired by the homing behavior of pigeons. Although PIO has demonstrated effectiveness and superiority in numerous fields, particularly in practical engineering optimization, there have been few results concerning the theoretical foundations of PIO. This paper conducts convergence analysis of basic PIO in a continuous search space in two aspects. First, we analyze the convergence of each pigeon's expected position using a difference equation and prove that the average position of each pigeon in the swarm will converge to the same value. To further study the stochastic global convergence property of the pigeon swarm, we apply the martingale theory to investigate the basic PIO swarm sequence, and achieve a sufficient condition to guarantee global convergence of the basic PIO. Our theoretical analysis shows that this convergence depends upon the accumulation of the minimum probability with which the pigeon swarm jumps to the global-optimal region at each iteration. The mathematical methods proposed in this study, particularly the martingale technique, also provide a new effective approach for the theoretical analysis of bio-inspired algorithms in continuous optimization.

Keywords pigeon-inspired optimization (PIO), convergence analysis, martingale, continuous optimization, swarm intelligence

Citation Zhang Y S, Huang H, Wu H Y, et al. Theoretical analysis of the convergence property of a basic pigeon-inspired optimizer in a continuous search space. Sci China Inf Sci, 2019, 62(7): 070207, https://doi.org/10.1007/s11432-018-9753-5

1 Introduction

Over the last two decades, population-based swarm intelligence (SI) optimization algorithms, such as particle swarm optimization (PSO) [1], ant colony optimization (ACO) [2], and brain storm optimization (BSO) [3], have attracted great interest from researchers and have been successfully applied to many complicated optimization problems. An increasing number of SI-optimization algorithms are appearing as various swarms in nature are simulated. These algorithms offer practical and efficient solutions for various optimization problems.

Recently, a new SI algorithm inspired by the homing behavior of pigeons, known as the pigeon-inspired optimization (PIO) algorithm, was proposed by Duan and Qiao in 2014 [4]. Pigeons can easily find their

* Corresponding author (email: hhan@scut.edu.cn)

homes using three homing tools: magnetic fields, the sun, and landmarks. The magnetic field is used to shape the map and adjust the homing direction according to the altitude of the sun. Landmarks around the pigeons guide them to their destination. To mimic this natural phenomenon, the PIO algorithm utilizes two operators to describe the flocking behavior of homing pigeons: the map and compass operator represents the effects of the magnetic field and the sun, and the landmark operator describes the effects of landmarks [5, 6].

Since this SI-optimization algorithm was first proposed, many research results have proposed improvements and applications for it. A series of comparative experiments using benchmarks and practical optimization problems have shown that the PIO offers better performance than other bio-inspired algorithms on efficiency and stability. Duan and Qiao [4] applied PIO to solve air-robot-path planning problems and found that PIO outperformed the standard differential evolution (DE) algorithm in terms of convergence speed and stability. To improve the performance of the power system of an unmanned aerial vehicle (UAV), a modified PIO algorithm called the adjacent-disturbances and integrated-dispatching PIO (ADID-PIO) algorithm was presented to optimize the design parameters of a dc brushless motor [6]; the comparative experimental results indicated that the convergence rate, efficiency, and stability of ADID-PIO were better than those of PIO, BSO, or predator-prey BSO (PPBSO) [7] in the design process of a DC brushless motor. A variant of PIO called predator-prey PIO (PPPIO) was proposed to solve unmanned combat aerial vehicle three-dimensional (3D) path-planning problems in a dynamic environment; the comparative simulation results showed that the PPPIO algorithm was more efficient than basic PIO, PSO, and DE [5]. For more applications of PIO and a comparison of its performance with other bio-inspired algorithms, please refer to [8–12].

Although PIO has demonstrated its effectiveness and superiority in numerous fields (particularly in practical engineering optimization) and has received extensive attention from researchers, the theoretical foundations of PIO—including convergence analysis and parameter-setting principles—still remain weak [13]. Currently, theoretical studies of PIO are mainly based on empirical and intuitive statistical results, and rigorous mathematical arguments are lacking [13]. Among theoretical studies of PIO, convergence analysis is a key problem of great significance that concerns the effect of essential factors on pigeon-swarm dynamics and the conditions under which pigeon swarms converge to certain constant positions [14–16]. Zhang and Duan [5] conducted preliminary convergence analysis of PIO by treating the state of the PIO algorithm's population sequence as a finite Markov chain. To the best of our knowledge, this has been the only study concerning convergence analysis of PIO thus far. However, PIO is mainly used to solve continuous optimization problems; i.e., PIO is a continuous bio-inspired algorithm. Therefore, the analysis in [5] can be regarded as a special case and must be generalized for a continuous situation.

The contributions of the present study can be divided into two parts. First, we use a difference equation to calculate the expected value of each pigeon's position at each iteration. By computing the limits, we obtain the statistically average convergent position of a pigeon swarm. However, a sequence of stochastic variables may not converge even if the corresponding expectation sequence converges [17]. Second, to further study the global convergence property of the pigeon swarm, we employ the martingale theory, which investigates basic PIO's evolutionary process and provides a sufficient condition to guarantee the algorithm's global convergence. The martingale technique adopted in the stochastic global convergence analysis of basic PIO does not require additional assumptions such as the Markov property, and can thus be extended to the theoretical analysis of bio-inspired algorithms in continuous optimization.

2 Basic PIO algorithm and its stochastic process model

2.1 Introduction to a basic PIO algorithm

According to [4], basic PIO employs two operators to mimic the behavior of homing pigeons: the map-and-compass, and landmark operators.

(1) Map and compass operator. Pigeons are randomly initialized in a D-dimension search space,

$\mathbb{R}^D$. The total number of pigeons is N_p, $(D, N_p \in \mathbb{Z}^+)$. The position and velocity of the k-th pigeon at the t-th iteration are denoted as $\boldsymbol{x}_k(t) = (x_{k1}(t), x_{k2}(t), \ldots, x_{kD}(t))$, and $\boldsymbol{v}_k(t) = (v_{k1}(t), v_{k2}(t), \ldots, v_{kD}(t))$, respectively, where $k = 1, \ldots, N_p$, $t = 0, 1, \ldots$.

The new velocity $\boldsymbol{v}_k(t)$, and position $\boldsymbol{x}_k(t)$, at the t-th iteration are updated as follows:

$$\boldsymbol{v}_k(t) = \boldsymbol{v}_k(t-1) \cdot \mathrm{e}^{-Rt} + r \cdot (\boldsymbol{P}_g(t-1) - \boldsymbol{x}_k(t-1)), \tag{1}$$

$$\boldsymbol{x}_k(t) = \boldsymbol{x}_k(t-1) + \boldsymbol{v}_k(t), \tag{2}$$

where $R \in (0, 1)$ is the map and compass factor; $r \sim U(0, 1)$ is a uniform random variable; and $\boldsymbol{P}_g(t-1)$ is the best position found up to the $(t-1)$-th iteration for the entire swarm (i.e., the global best position).

When the number of loops reaches the required number of iterations, the map-and-compass operator stops executing, and the landmark operator proceeds to work.

(2) Landmark operator. We consider the maximization problem as an example. We sort all pigeons in descending order of size according to their fitness values. Thus, the total number of pigeons in every generation is halved, and the pigeons in the less-fit half are abandoned. Let $\boldsymbol{X}_C(t)$ be the center of the pigeons' positions at the t-th generation; thus the position updating rule for each pigeon (k) at iteration t can be given as follows:

$$N_p(t) = \text{ceil}\left(\frac{N_p(t-1)}{2}\right), \tag{3}$$

$$\boldsymbol{X}_C(t) = \frac{\sum_{j=1}^{N_p(t)} \boldsymbol{x}_j(t) \cdot \text{fitness}\,(\boldsymbol{x}_j(t))}{N_p(t) \sum_{j=1}^{N_p(t)} \text{fitness}\,(\boldsymbol{x}_j(t))}, \tag{4}$$

$$\boldsymbol{x}_k(t) = \boldsymbol{x}_k(t-1) + r \cdot (\boldsymbol{X}_C(t) - \boldsymbol{x}_k(t-1)), \quad k = 1, 2, \ldots, N_p. \tag{5}$$

In the abovementioned equations, $N_p(t)$ represents the number of pigeons in the t-th generation, $N_p(0) = N_p$. The function ceil$(\cdot)$ returns the smallest integer value that is greater than or equal to each input value, after a certain number of iterations, $N_p(t) = \text{ceil}(\frac{Np}{2^t})$, as $\frac{Np}{2^t} \in (0, 1)$; thus, $N_p(t)$ will remain 1. Therefore, $\boldsymbol{X}_C(t)$ will be the current best position (note: not the global best position) in the t-th generation. $r \sim U(0, 1)$ is a uniform random variable independent of both $\boldsymbol{x}_k(t)$ and $\boldsymbol{X}_C(t)$.

The implementation procedure for basic PIO is described below [4], as shown in Algorithm 1.

The basic PIO algorithm can be regarded as a new version of the evolutionary algorithm (EA), with the map-and-compass operator corresponding to the exploration phase and the landmark operator corresponding to the exploitation phase of classical EAs. As $N_p(t)$ will converge to 1 in finite time, PIO is similar to an elitist $(1 + N_p)$ EA acting on a continuous real space.

2.2 Description of the optimization problem

Without loss of generality, we assume that the PIO algorithm analyzed herein is used to tackle maximization problems in a continuous search space.

Definition 1 (Maximization problem)**.** Let $S = \prod_{i=1}^{D} [-a_i, a_i] \subset \mathbb{R}^D$, $a_i > 0$ be a D-dimensional continuous search space, and let $f : S \to \mathbb{R}$ be a D-dimensional function. For maximization problems, we find a global optimum $\boldsymbol{x}^* \in S$, such that $f^* \triangleq f(\boldsymbol{x}^*) = \max_{\boldsymbol{x} \in S} f(\boldsymbol{x})$.

The function $f : S \to \mathbb{R}$ is called the objective function of the maximization problem. We do not require f to be continuous; however, it must be bounded. Furthermore, in this study, we only consider unconstrained optimization.

In addition, the following properties are assumed.

(1) The subset containing the global optimal solutions in S is non-empty.

(2) Let $S^*(\varepsilon) = \{\boldsymbol{x} \in S | f(\boldsymbol{x}) > f^* - \varepsilon\}$ be the global optimum ε-neighborhood. Each element of $S^*(\varepsilon)$ may be considered as a maximum.

(3) For any $\varepsilon > 0$, the Lebesgue measure of $S^*(\varepsilon)$, denoted as $m(S^*(\varepsilon)) > 0$.

The first assumption describes the existence of global optima for the problem. The second assumption presents a rigorous definition of the global optimum for continuous maximization problems. The third

Algorithm 1 Basic pigeon-inspired optimization (PIO) algorithm

Input: N_p: number of individuals in a pigeon swarm.
D: dimension of the search space.
R: the map and compass factor.
$Nc_{1\max}$: maximum number of generations for which the map-and-compass operation is performed.
$Nc_{2\max}$: maximum number of generations for which the landmark operation is performed.
Output: $\boldsymbol{P}_g$: the global best position.
1. Initialization
Set initial values for $Nc_{1\max}$, $Nc_{2\max}$, N_p, D, and R.
Set initial position $\boldsymbol{x}_k$ and velocity $\boldsymbol{v}_k$ for each pigeon, $k = 1, \ldots, N_p$.
2. Map and compass operations
for $t = 1$ to $Nc_{1\max}$ **do**
  for $k = 1$ to N_p **do**
    Calculate $\boldsymbol{v}_k(t)$ and $\boldsymbol{x}_k(t)$ according to Eqs. (1) and (2);
  end for
  Evaluate $\boldsymbol{x}_k(t)$, $k = 1, \ldots, N_p$ and update $\boldsymbol{P}_g(t)$;
end for
3. Landmark operations
for $t = Nc_{1\max} + 1$ to $Nc_{2\max}$ **do**
  Rank all pigeon individuals according to their fitness values;
  $N_p(t) = \mathrm{ceil}(\frac{N_p(t-1)}{2})$;
  Keep $N_p(t)$ individuals with better fitness values and abandon the others;
  Calculate $\boldsymbol{X}_C(t)$ and update $\boldsymbol{x}_k(t)$, $k = 1, \ldots, N_p$ according to Eqs. (4) and (5);
  Evaluate $\boldsymbol{x}_k(t)$, $k = 1, \ldots, N_p$ and update $\boldsymbol{P}_g(t)$;
end for
4. Output
$\boldsymbol{P}_g(Nc_{2\max})$ is output as the global optimum.

assumption shows that there always exist solutions whose objective values are continuously distributed and are arbitrarily close to the global optimum, thus making the maximization problem solvable.

2.3 Stochastic process model of the basic PIO algorithm

Our convergence analyses represent the basic PIO algorithm as a stochastic process. In this subsection, we will explain the notations and terminologies used in this study.

Definition 2 (State of the pigeon swarm)**.** The state of the pigeon swarm at iteration t ($t = 0, 1, \ldots$) is defined as $\boldsymbol{\eta}(t) = (\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_{N_p}(t), \boldsymbol{P}_g(t))$, where $\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_{N_p}(t), \boldsymbol{P}_g(t) \in S$.

Definition 3 (State space of the pigeon swarm)**.** The set of all possible pigeon swarm states is called the state space of the pigeon swarm, denoted as $\Omega = S^{N_p+1} = \{\boldsymbol{\eta} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N_p}, \boldsymbol{\xi}) | \boldsymbol{x}_k \in S, k = 1, \ldots, N_p; \boldsymbol{\xi} \in S\}$.

Definition 4 (ε-global optimum state space of the pigeon swarm)**.** The ε-global-optimum state space of the pigeon swarm is defined as $\Omega^*(\varepsilon) = \{\boldsymbol{\eta} = (\boldsymbol{x}_1, \ldots, \boldsymbol{x}_{N_p}, \boldsymbol{\xi}) | \exists \boldsymbol{x}_k \in S^*(\varepsilon), k = 1, \ldots, N_p; \boldsymbol{\xi} \in S\}$.

In Subsection 3.3, we will discuss the stochastic convergence of the pigeon swarm sequence to $\Omega^*(\varepsilon)$.

Definition 5 (Discrete time stochastic process of PIO)**.** The discrete time stochastic process associated with the PIO algorithm is denoted as $\{\boldsymbol{\eta}(t) = (\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_{N_p}(t), \boldsymbol{P}_g(t))\}_{t=0}^{+\infty}$, whose state space is Ω.

3 Convergence analysis of basic PIO: from individual to swarm

Intuitively, the convergence analysis of PIO investigates the algorithm's characteristics when the number of iterations approaches infinity. According to the description of Algorithm 1, we can see that the basic PIO procedure can be divided into two main stages: map-and-compass operations and landmark operations; the maximum numbers of iterations of both stages are predefined. Therefore, during convergence analysis, the first stage can be ignored because previous finite iterations do not affect the convergence property. In the following study, our convergence analysis begins from the second stage: landmark operations, reconsidering $t = 0, 1, \ldots$ instead of $t = Nc_{1\max}+1$ to $Nc_{2\max}$.

3.1 Convergence analysis of the expectation value of a pigeon's position

Let $x_{ki}(t), X_{Ci}(t)$ be the i-th components of random vectors $\boldsymbol{x}_k(t), \boldsymbol{X}_C(t)$, respectively; $i = 1, 2, \ldots, D$, $k = 1, 2, \ldots, N_p$, $t = 0, 1, \ldots$. According to (5), we obtain the following difference equation with regard to the expectation sequence $\{E(x_{ki}(t))\}_{t=1}^{+\infty}$:

$$E(x_{ki}(t)) = \frac{1}{2}E(x_{ki}(t-1)) + \frac{1}{2}E(X_{Ci}(t)). \tag{6}$$

Eq. (6) can be considered as a linear non-homogeneous first-order difference equation, whose general solution is provided in [18]:

$$E(x_{ki}(t)) = \frac{\frac{1}{2}E(X_{Ci}(t))}{1-\frac{1}{2}} + A \cdot \left(\frac{1}{2}\right)^t = E(X_{Ci}(t)) + A \cdot \left(\frac{1}{2}\right)^t, \tag{7}$$

where A is a real constant.

Considering the limit of (7) as t approaches infinity, we obtain

$$\lim_{t\to+\infty} E(x_{ki}(t)) = \lim_{t\to+\infty} E(X_{Ci}(t)). \tag{8}$$

Furthermore, we derive the convergence result of each pigeon's statistically average position amidst the entire swarm:

$$\begin{aligned}
\lim_{t\to+\infty} E(\boldsymbol{x}_k(t)) &= \lim_{t\to+\infty} (E(x_{k1}(t)), E(x_{k2}(t)), \ldots, E(x_{kD}(t))) \\
&= \left(\lim_{t\to+\infty} E(x_{k1}(t)), \lim_{t\to+\infty} E(x_{k2}(t)), \ldots, \lim_{t\to+\infty} E(x_{kD}(t))\right) \\
&= \left(\lim_{t\to+\infty} E(X_{C1}(t)), \lim_{t\to+\infty} E(X_{C2}(t)), \ldots, \lim_{t\to+\infty} E(X_{CD}(t))\right) \\
&= \lim_{t\to+\infty} E(\boldsymbol{X}_C(t)), \quad k \in \{1, 2, \ldots, N_p\}.
\end{aligned} \tag{9}$$

Notably, $N_p(t)$ in (3) is equal to 1 post finite t iterations; then, based on the procedure of Algorithm 1 and (4), we obtain

$$\boldsymbol{X}_C(t) = \arg\max\left\{f(\boldsymbol{x}_1(t)), f(\boldsymbol{x}_2(t)), \ldots, f\left(\boldsymbol{x}_{N_p}(t)\right)\right\}. \tag{10}$$

Thus, $\boldsymbol{X}_C(t)$ is the current best position at the t-th generation of the entire pigeon swarm.

Based on (9) and (10), we finally conclude that the average position of each pigeon in the swarm will converge to the same value, i.e., $\lim_{t\to+\infty} E(\boldsymbol{X}_C(t))$, provided that this limit exists.

3.2 A brief introduction to martingale theory

In Subsection 3.1, we only indicate that the statistically average position of each pigeon $E(\boldsymbol{x}_k(t))$ has the same limit as that of the current best solution (i.e., pigeon) $E(\boldsymbol{X}_C(t))$ when t approaches infinity. This does not mean that the convergent position is a global or even a local optimum. In Subsection 3.4, we will employ martingale theory to investigate the global convergence of the stochastic PIO process, as defined in Definition 5.

In probability theory, a martingale is a stochastic process in which the conditional expectation of the next value, given the current and preceding ones, is equal to the current value. The submartingale is a special case of the martingale, where the conditional expectation of the next value, given the current and preceding values, is not less than the current value. The formal definition of the submartingale is as follows.

Definition 6 (Submartingale [17])**.** For two stochastic processes $\{Y_j\}_{j=0}^{+\infty}$ and $\{Z_j\}_{j=0}^{+\infty}$, $\{Y_j\}_{j=0}^{+\infty}$ is called a submartingale with respect to $\{Z_j\}_{j=0}^{+\infty}$ if the following conditions hold for $\forall j \geqslant 0$:

(a) $E(|Y_j|) < +\infty$;

(b) $E(Y_{j+1}|Z_0, Z_1, \ldots, Z_j) \geqslant Y_j$;

(c) Y_j is a function of $Z_0, Z_1, \ldots, Z_j$.

The submartingale convergence theorem is now presented as follows.

Theorem 1 (Submartingale convergence theorem [17])**.** Suppose $\{Y_j\}_{j=0}^{+\infty}$ is a submartingale with respect to $\{Z_j\}_{j=0}^{+\infty}$ and $\sup_{j\geqslant 0} E\left(|Y_j|\right) < +\infty$. Then, there exists a random variable Y_∞, such that $\{Y_j\}_{j=0}^{+\infty}$ converges to Y_∞ with probability 1; i.e., $P(\lim_{j\to+\infty} Y_j = Y_\infty) = 1$ and $E\left(|Y_\infty|\right) < +\infty$.

3.3 Some related convergence analyses on continuous EAs

As mentioned in Subsection 2.1, PIO is similar to an elitist $(1+N_p)$EA acting on a continuous real space; therefore, related theoretical results for the convergence analyses of continuous EAs are helpful for the convergence analysis of PIO.

Rudolph [19, 20] introduced the formalization of stochastic process models of EAs and presented some of the most frequently used stochastic convergence concepts regarding EAs. He also brought to attention that modeling the transition probability function [21] is the primary task when establishing a link between an EA and a stochastic process. Using the martingale theory and a transition probability function, Rudolph presented two sufficient conditions for the convergence of an EA acting on a continuous space: (i) elitism, i.e., the best state found thus far cannot be lost from one iteration to another, and (ii) a positively bounded probability for reaching the target zone in one step from any point of the space.

Agapie et al. [21, 22] deeply analyzed the transition probability function of the stochastic process associated with continuous EAs, indicating that the associated one-step kernel can be described as a sum of two measures, one singular (Dirac) and one continuous. This result shows us how to accurately calculate the conditional mathematical expectation in Theorem 2 of this study. By modeling the continuous EA as a renewal process, Agapie et al. [22, 23] analyzed the computation time of continuous EAs, such as $(1+\lambda)$ ES and $(\mu+\lambda)$ ES [24].

3.4 Stochastic convergence analysis of a basic PIO based on the martingale theory

For any pigeon swarm $\boldsymbol{\eta}\,(t) = (\boldsymbol{x}_1(t), \ldots, \boldsymbol{x}_{N_p}(t), \boldsymbol{P}_g\,(t))$, $t = 0, 1, \ldots$, the fitness function is denoted as $F\,(\boldsymbol{\eta}\,(t))$. Assuming $f\,(\cdot)$ to be the objective function of the considered maximization problem, similar to [20], we define

$$F\left(\boldsymbol{\eta}\,(t)\right) \triangleq \max\left\{f\left(\boldsymbol{x}_1\,(t)\right), \ldots, f\left(\boldsymbol{x}_{N_p}\,(t)\right), f\left(\boldsymbol{P}_g\,(t)\right)\right\}.$$

As $\boldsymbol{P}_g\,(t)$ is the best position found until the t-th iteration by the entire swarm (i.e., the global best position), we can let $F\left(\boldsymbol{\eta}\,(t)\right) = f(\boldsymbol{P}_g\,(t))$; therefore, $\{F\left(\boldsymbol{\eta}\,(t)\right)\}_{t=0}^{+\infty}$ is a monotonic non-decreasing sequence.

Lemma 1. The stochastic process $\{F\left(\boldsymbol{\eta}\,(t)\right)\}_{t=0}^{+\infty}$ is a submartingale with respect to $\{\boldsymbol{\eta}\,(t)\}_{t=0}^{+\infty}$.

Proof. We verify the three conditions in Definition 6.

(a) As mentioned in Subsection 2.2, $f\,(\cdot)$ is a bounded function on the search space S, hence for $\forall t = 0, 1, \ldots$, $E\left(|F\left(\boldsymbol{\eta}\,(t)\right)|\right) = E(|f(\boldsymbol{P}_g\,(t))|) < +\infty$.

(b) As $\{F\left(\boldsymbol{\eta}\,(t)\right)\}_{t=0}^{+\infty}$ is a monotonic non-decreasing sequence, we obtain

$$\begin{aligned}&F\left(\boldsymbol{\eta}\,(t+1)\right) \geqslant F\left(\boldsymbol{\eta}\,(t)\right)\\ &\quad\Rightarrow E\left(F\left(\boldsymbol{\eta}\,(t+1)\right)|\boldsymbol{\eta}\,(0)\,, \boldsymbol{\eta}\,(1)\,, \ldots, \boldsymbol{\eta}\,(t)\right) \geqslant E\left(F\left(\boldsymbol{\eta}\,(t)\right)|\boldsymbol{\eta}\,(0)\,, \boldsymbol{\eta}\,(1)\,, \ldots, \boldsymbol{\eta}\,(t)\right) = F\left(\boldsymbol{\eta}\,(t)\right).\end{aligned}$$

(c) Furthermore, $F\left(\boldsymbol{\eta}\,(t)\right)$ is a function of $\boldsymbol{\eta}\,(0)\,, \boldsymbol{\eta}\,(1)\,, \ldots, \boldsymbol{\eta}\,(t)$.

In conclusion, this completed the proof.

Therefore, according to Theorem 1, there exists a random variable $F\left(\boldsymbol{\eta}\,(\infty)\right)$ to which the stochastic process $\{F\left(\boldsymbol{\eta}\,(t)\right)\}_{t=0}^{+\infty}$ converges with probability 1 as $t \to +\infty$. Similarly, the stochastic process $\{\boldsymbol{\eta}\,(t)\}_{t=0}^{+\infty}$ converges to $\boldsymbol{\eta}\,(\infty)$ with probability 1, i.e., $P(\lim_{t\to+\infty} \boldsymbol{\eta}\,(t) = \boldsymbol{\eta}\,(\infty)) = 1$.

Definition 7 (Global convergence of basic PIO)**.** Suppose the discrete time stochastic process associated with PIO is $\{\boldsymbol{\eta}\,(t)\}_{t=0}^{+\infty}$. If $\lim_{t\to+\infty} P\left(\boldsymbol{\eta}\,(t) \in \Omega^*\,(\varepsilon)\right) = 1$ holds for any $0 < \varepsilon \in \mathbb{R}$, we call this the global convergence of basic PIO.

In the following analysis, before a sufficient condition for the guaranteed global convergence of basic PIO is provided, the following lemma is presented.

Lemma 2. When $\sum_{k=0}^{\infty} a_k b_k < +\infty$ $(0 \leqslant a_k, b_k \leqslant 1, k = 0, 1, \ldots)$, the following two cases hold:

(a) If $\sum_{k=0}^{\infty} a_k = +\infty$, then $\lim_{k\to\infty} b_k = 0$.

(b) If $\sum_{k=0}^{\infty} a_k < +\infty$, then $\lim_{k\to\infty} b_k = \delta \in [0, 1]$ or it does not exist.

Proof. (a) Assume $\lim_{k\to\infty} b_k = \beta > 0$; then, for $\frac{\beta}{2}$, $\exists K \in \mathbb{N}^+$, such that $\forall k > K$, we have $\frac{\beta}{2} < b_k$. Therefore, $\sum_{k=0}^{\infty} a_k b_k = \sum_{k=0}^{K} a_k b_k + \sum_{k=K+1}^{\infty} a_k b_k > \sum_{k=0}^{K} a_k b_k + \frac{\beta}{2}\sum_{k=K+1}^{\infty} a_k = +\infty$, which contradicts $\sum_{k=0}^{\infty} a_k b_k < +\infty$, i.e., $\lim_{k\to\infty} b_k = 0$ holds.

(b) Note that $a_k b_k \leqslant a_k$, $k = 0, 1, \ldots$; hence, $\sum_{k=0}^{\infty} a_k < +\infty$ implies $\sum_{k=0}^{\infty} a_k b_k < +\infty$, which means that b_k can take any value in [0,1]. Specifically, let $b_k = \delta \in [0, 1]$, and then, $\lim_{k\to\infty} b_k = \delta \in [0, 1]$; let $b_k = |\sin k|$, and then, $\lim_{k\to\infty} b_k$ does not exist.

Theorem 2. Let $q_t^* = \min_{\boldsymbol{y}\in\Omega\backslash\Omega^*(\varepsilon)} P\left(\boldsymbol{\eta}(t+1) \in \Omega^*(\varepsilon) \,|\boldsymbol{\eta}(t) = \boldsymbol{y}\right)$, $t = 0, 1, \ldots$; consequently, we obtain the following inferences:

(a) If $\sum_{t=0}^{\infty} q_t^* = +\infty$, then the global convergence of basic PIO can be guaranteed.

(b) If $\sum_{t=0}^{\infty} q_t^* < +\infty$, then the global convergence of basic PIO cannot be guaranteed.

Proof. According to the properties of conditional expectation, for $t = 0, 1, \ldots$, we have

$$E\left(F\left(\boldsymbol{\eta}(t+1)\right)\right) - E\left(F\left(\boldsymbol{\eta}(t)\right)\right) = E\left[E\left(F\left(\boldsymbol{\eta}(t+1)\right)|\boldsymbol{\eta}(t)\right)\right] - E\left(F\left(\boldsymbol{\eta}(t)\right)\right). \tag{11}$$

Suppose the probability distribution function of $\boldsymbol{\eta}(t)$ to be $P_t(\boldsymbol{y})$ and the conditional-probability-distribution function of $\boldsymbol{\eta}(t+1)$ given $\boldsymbol{\eta}(t) = \boldsymbol{y}$ to be $P_t(\boldsymbol{z}|\boldsymbol{y})$. Therefore, based on (11), we obtain

$$\begin{aligned}
&E\left[E\left(F\left(\boldsymbol{\eta}(t+1)\right)|\boldsymbol{\eta}(t)\right)\right] - E\left(F\left(\boldsymbol{\eta}(t)\right)\right)\\
&= \int_\Omega E\left[F\left(\boldsymbol{\eta}(t+1)\right)|\boldsymbol{\eta}(t) = \boldsymbol{y}\right] \mathrm{d}P_t(\boldsymbol{y}) - \int_\Omega F(\boldsymbol{y}) \mathrm{d}P_t(\boldsymbol{y})\\
&= \int_\Omega \left[\int_\Omega F(\boldsymbol{z}) \mathrm{d}P_t(\boldsymbol{z}|\boldsymbol{y})\right] \mathrm{d}P_t(\boldsymbol{y}) - \int_\Omega F(\boldsymbol{y}) \mathrm{d}P_t(\boldsymbol{y})\\
&= \int_\Omega \left[\int_\Omega F(\boldsymbol{z}) \mathrm{d}P_t(\boldsymbol{z}|\boldsymbol{y}) - F(\boldsymbol{y})\right] \mathrm{d}P_t(\boldsymbol{y})\\
&= \int_\Omega \left[\int_\Omega (F(\boldsymbol{z}) - F(\boldsymbol{y})) \mathrm{d}P_t(\boldsymbol{z}|\boldsymbol{y})\right] \mathrm{d}P_t(\boldsymbol{y}).
\end{aligned}$$

Noting that $\int_\Omega F(\boldsymbol{y}) \mathrm{d}P_t(\boldsymbol{z}|\boldsymbol{y}) = F(\boldsymbol{y}) \int_\Omega \mathrm{d}P_t(\boldsymbol{z}|\boldsymbol{y}) = F(\boldsymbol{y})$, the above last equation holds, so we get

$$\int_\Omega \left[\int_\Omega (F(\boldsymbol{z}) - F(\boldsymbol{y})) \mathrm{d}P_t(\boldsymbol{z}|\boldsymbol{y})\right] \mathrm{d}P_t(\boldsymbol{y}) \geqslant \int_{\Omega\backslash\Omega^*(\varepsilon)} \left[\int_{\Omega^*(\varepsilon)} (F(\boldsymbol{z}) - F(\boldsymbol{y})) \mathrm{d}P_t(\boldsymbol{z}|\boldsymbol{y})\right] \mathrm{d}P_t(\boldsymbol{y}).$$

Let $\alpha = \min\left\{F(\boldsymbol{z}) - F(\boldsymbol{y}) \,|\boldsymbol{z} \in \Omega^*(\varepsilon), \boldsymbol{y} \in \Omega\backslash\Omega^*(\varepsilon)\right\}$, as $F(\boldsymbol{z}) > f^* - \varepsilon, F(\boldsymbol{y}) \leqslant f^* - \varepsilon$; thus, we obtain $\alpha > 0$ provided that α exists, it holds that

$$\begin{aligned}
\int_{\Omega\backslash\Omega^*(\varepsilon)} \left[\int_{\Omega^*(\varepsilon)} (F(\boldsymbol{z}) - F(\boldsymbol{y})) \mathrm{d}P_t(\boldsymbol{z}|\boldsymbol{y})\right] \mathrm{d}P_t(\boldsymbol{y}) &\geqslant \alpha \int_{\Omega\backslash\Omega^*(\varepsilon)} \left[\int_{\Omega^*(\varepsilon)} \mathrm{d}P_t(\boldsymbol{z}|\boldsymbol{y})\right] \mathrm{d}P_t(\boldsymbol{y})\\
&= \alpha \int_{\Omega\backslash\Omega^*(\varepsilon)} P\left(\boldsymbol{\eta}(t+1) \in \Omega^*(\varepsilon) \,|\boldsymbol{\eta}(t) = \boldsymbol{y}\right) \mathrm{d}P_t(\boldsymbol{y}).
\end{aligned}$$

Let $q_t^* = \min_{\boldsymbol{y}\in\Omega\backslash\Omega^*(\varepsilon)} P\left(\boldsymbol{\eta}(t+1) \in \Omega^*(\varepsilon) \,|\boldsymbol{\eta}(t) = \boldsymbol{y}\right)$; this results in

$$\alpha \int_{\Omega\backslash\Omega^*(\varepsilon)} P\left(\boldsymbol{\eta}(t+1) \in \Omega^*(\varepsilon) \,|\boldsymbol{\eta}(t) = \boldsymbol{y}\right) \mathrm{d}P_t(\boldsymbol{y}) \geqslant \alpha q_t^* \int_{\Omega\backslash\Omega^*(\varepsilon)} \mathrm{d}P_t(\boldsymbol{y}) = \alpha q_t^* P\left(\boldsymbol{\eta}(t) \notin \Omega^*(\varepsilon)\right). \tag{12}$$

Combining (11) and (12), we get

$$\sum_{t=0}^{M} \left[E\left(F\left(\boldsymbol{\eta}(t+1)\right)\right) - E\left(F\left(\boldsymbol{\eta}(t)\right)\right)\right] = E\left(F\left(\boldsymbol{\eta}(M+1)\right)\right) - E\left(F\left(\boldsymbol{\eta}(0)\right)\right)$$

$$\geqslant \alpha \sum_{t=0}^{M} [q_t^* P(\boldsymbol{\eta}(t) \notin \Omega^*(\varepsilon))]. \tag{13}$$

As $\{F(\boldsymbol{\eta}(t))\}_{t=0}^{+\infty}$ is a bounded submartingale, let $M \to +\infty$ in (13); thus, we obtain

$$\sum_{t=0}^{\infty} [q_t^* P(\boldsymbol{\eta}(t) \notin \Omega^*(\varepsilon))] < +\infty. \tag{14}$$

Now, we discuss two situations related to the global convergence of basic PIO using Lemma 2:

(a) If $\sum_{t=0}^{\infty} q_t^* = +\infty$, it holds that $\lim_{t\to+\infty} P(\boldsymbol{\eta}(t) \notin \Omega^*(\varepsilon)) = 0$, i.e., $\lim_{t\to+\infty} P(\boldsymbol{\eta}(t) \in \Omega^*(\varepsilon)) = 1$.

As mentioned above, the stochastic process $\{\boldsymbol{\eta}(t)\}_{t=0}^{+\infty}$ converges to $\boldsymbol{\eta}(\infty)$ with probability 1, i.e., $P(\lim_{t\to+\infty} \boldsymbol{\eta}(t) = \boldsymbol{\eta}(\infty)) = 1$; this means that $\boldsymbol{\eta}(\infty)$ almost surely takes values in $\Omega^*(\varepsilon)$; as a result, the global convergence of basic PIO is guaranteed.

(b) If $\sum_{t=0}^{\infty} q_t^* < +\infty$, it holds that $\lim_{t\to+\infty} P(\boldsymbol{\eta}(t) \notin \Omega^*(\varepsilon)) = \delta \in [0,1]$ or $\lim_{t\to+\infty} P(\boldsymbol{\eta}(t) \notin \Omega^*(\varepsilon))$ does not exist.

Similar to situation (a), basic PIO possesses global convergence if $\lim_{t\to+\infty} P(\boldsymbol{\eta}(t) \notin \Omega^*(\varepsilon)) = 0$.

If $\lim_{t\to+\infty} P(\boldsymbol{\eta}(t) \notin \Omega^*(\varepsilon)) = \delta \in (0,1]$, $\boldsymbol{\eta}(\infty)$ has a positive probability of adopting values outside $\Omega^*(\varepsilon)$, it implies that the basic PIO is not globally convergent.

If $\lim_{t\to+\infty} P(\boldsymbol{\eta}(t) \notin \Omega^*(\varepsilon))$ does not exist, then the basic PIO is obviously not globally convergent.

In summary, under situation (b), the global convergence of the basic PIO cannot be guaranteed.

Intuitively, q_t^* represents the minimum probability with which the pigeon swarm escapes to the ε-global optimum region from the outside at time $t = 0, 1, \ldots$. It reflects the ability with which the pigeon swarm approaches the global optimum: with the growth of q_t^*, the pigeon swarm can approach the global optimum more easily. Although each q_t^* may be small, basic PIO still can attain a global optimum provided that the accumulation of q_t^* becomes sufficiently large.

4 Conclusion

We studied the convergence property of basic PIO from two perspectives. First, we analyzed the convergence of the pigeon's expected position using a difference equation and proved that the average position of each pigeon in the swarm will converge to the same value; however, this does not mean that the convergent position is a global or even a local optimum. Next, to further study the global convergence property of the pigeon swarm, we introduced the martingale theory and investigated the PIO's swarm sequence; we proved that the global convergence of basic PIO depends on $\sum_{t=0}^{\infty} q_t^*$. When $\sum_{t=0}^{\infty} q_t^* = +\infty$, the basic PIO is guaranteed to globally converge. When $\sum_{t=0}^{\infty} q_t^* < +\infty$, global convergence cannot be guaranteed. Based on the theoretical analysis provided in Subsection 3.4, we can see that the adopted martingale technique does not require additional assumptions such as a Markov property, and is suitable for the theoretical analysis of bio-inspired algorithms in continuous optimization. In future work, we will attempt to improve PIO using the theoretical results of the PIO's convergence analysis, and perform PIO-runtime analysis.

Acknowledgements This work was supported by Natural Science Foundation of China-Guangdong Joint Fund (Grant No. U1501254), National Natural Science Foundation of China (Grant Nos. 61772225, 61876207), Guangdong Natural Science Funds for Distinguished Young Scholar (Grant No. 2014A030306050), the Ministry of Education - China Mobile Research Funds (Grant No. MCM20160206), Guangdong High-level Personnel of Special Support Program (Grant No. 2014TQ01X664), International Cooperation Project of Guangzhou (Grant No. 201807010047), and Science and Technology Program of Guangzhou (Grant Nos. 201804010276, 201707010227, 201707010228).

References

1 Bonyadi M R, Michalewicz Z. Particle swarm optimization for single objective continuous space problems: a review. Evolary Comput, 2017, 25: 1–54

2 Dorigo M, Blum C. Ant colony optimization theory: a survey. Theor Comput Sci, 2005, 344: 243–278
3 Shi Y H. An optimization algorithm based on brainstorming process. Int J Swarm Intell Res, 2011, 2: 35–62
4 Duan H B, Qiao P X. Pigeon-inspired optimization: a new swarm intelligence optimizer for air robot path planning. Int J Intell Comput Cyber, 2014, 7: 24–37
5 Zhang B, Duan H B. Three-dimensional path planning for uninhabited combat aerial vehicle based on predator-prey pigeon-inspired optimization in dynamic environment. IEEE/ACM Trans Comput Biol Bioinf, 2017, 14: 97–107
6 Xu X B, Deng Y M. UAV power component-DC brushless motor design with merging adjacent-disturbances and integrated-dispatching pigeon-inspired optimization. IEEE Trans Magn, 2018, 54: 1–7
7 Duan H B, Li S T, Shi Y H. Predator-prey brain storm optimization for DC brushless motor. IEEE Trans Magn, 2013, 49: 5336–5340
8 Li T, Zhou C J, Wang B, et al. A hybrid algorithm based on artificial bee colony and pigeon inspired optimization for 3D protein structure prediction. J Bionanosci, 2018, 12: 100–108
9 Sushnigdha G, Joshi A. Trajectory design of re-entry vehicles using combined pigeon inspired optimization and orthogonal collocation method. IFAC-PapersOnLine, 2018, 51: 656–662
10 Xin L, Xian N. Biological object recognition approach using space variant resolution and pigeon-inspired optimization for UAV. Sci China Technol Sci, 2017, 60: 1577–1584
11 Rajendran S, Sankareswaran U M. A novel pigeon inspired optimization in ovarian cyst detection. Curr Med Imaging Rev, 2016, 12: 43–49
12 Lei X, Ding Y, Wu F X. Detecting protein complexes from DPINs by density based clustering with Pigeon-Inspired Optimization Algorithm. Sci China Inf Sci, 2016, 59: 070103
13 Duan H B, Ye F. Progresses in pigeon-inspired optimization algorithms. J Beijing Univ Technol (Nat Sci Ed), 2017, 43: 1–7
14 Xu G, Yu G S. On convergence analysis of particle swarm optimization algorithm. J Comput Appl Math, 2018, 333: 65–73
15 Liu Q F. Order-2 stability analysis of particle swarm optimization. Evolary Comput, 2015, 23: 187–216
16 Bonyadi M R, Michalewicz Z. Analysis of stability, local convergence, and transformation sensitivity of a variant of the particle swarm optimization algorithm. IEEE Trans Evol Comput, 2016, 20: 370–385
17 Nguyen H T, Wang T H. A Graduate Course in Probability and Statistics, Volume I, Essentials of Probability for Statistics. Beijing: Tsinghua University Press, 2008
18 Elaydi S N. An Introduction to Difference Equations. 3rd ed. New York: Springer, 2005
19 Rudolph G. Convergence properties of evolutionary algorithms. Dissertation for Ph.D. Degree. Hamburg: University of Hamburg, 1997
20 Rudolph G. Stochastic convergence. In: Handbook of Natural Computing. Berlin: Springer, 2010
21 Agapie A, Agapie M. Transition functions for evolutionary algorithms on continuous state-space. J Math Model Algor, 2007, 6: 297–315
22 Agapie A, Agapie M, Zbaganu G. Evolutionary algorithms for continuous-space optimisation. Int J Syst Sci, 2013, 44: 502–512
23 Agapie A, Agapie M, Rudolph G, et al. Convergence of evolutionary algorithms on the n-dimensional continuous space. IEEE Trans Cybern, 2013, 43: 1462–1472
24 Beyer H G. The Theory of Evolution Strategies. New York: Springer, 2001