Information Sciences 279 (2014) 18–36

Contents lists available at ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Behavioral analysis of the leader particle during stagnation in a particle swarm optimization algorithm

Sarthak Chatterjee [a], Debdipta Goswami [a], Sudipto Mukherjee [a], Swagatam Das [b,*]

[a] Department of Electronics and Telecommunication Engineering, Jadavpur University, Kolkata, West Bengal, India
[b] Electronics and Communication Sciences Unit, Indian Statistical Institute, Kolkata, India

ARTICLE INFO

Article history:
Received 29 October 2013
Received in revised form 12 February 2014
Accepted 22 March 2014
Available online 25 April 2014

Keywords:
Particle swarm optimization
Leader particle
Stagnation
Stable region
Gradient of the objective function

ABSTRACT

Concept of the particle swarms emerged from a simulation of the collective behavior of social creatures. It gradually evolved into a powerful derivative-free optimization techniques, now known as Particle Swarm Optimization (PSO) for solving multi-dimensional, multi-modal, and non-convex optimization problems. The dynamics governing the movement of the particles in PSO has invoked a great deal of research interest over the last decade. Theoretical investigations of PSO has mostly focused on particle trajectories in the search space and the parameter-selection. This work looks into the PSO algorithm from the perspective of the leader particle and takes into account stagnation, a situation where particles are trapped at less coveted local optima, thus preventing them from reaching more coveted global optima. We show that the points sampled by the leader particle satisfy a simple mathematical relation which demonstrates that they lie on a specific line. We demonstrate the condition under which for certain values of the parameters, particles stick to exploring one side of the stagnation point only and ignore the other side, and also the case where both sides are explored. We also obtain information about the gradient of the objective function during stagnation in PSO. We provide a large number of machine simulations which support our claims over several ranges of the control parameters. This sheds light on possible modifications to the basic PSO algorithm which would help future researchers to work with even more efficient and state-of-the-art PSO variants.

1. Introduction

Kennedy and Eberhart [29,18] introduced the concept of function optimization by means of a particle swarm in 1995 [10,18]. In the basic PSO scheme, a "swarm" of particles move around in the search space influenced by continually better and improved positions discovered by other particles. PSO does not require any derivative information of the function to be optimized, uses only rudimentary mathematical operators, and is conceptually very simple.

Since its inception in 1995, PSO has attracted a great deal of attention of the researchers all over the globe resulting into nearly uncountable number of variants of the basic algorithm, theoretical and empirical investigations of the dynamics of the particles, parameter selection and control, and applications of the algorithm to a wide spectrum of real world problems from diverse fields of science and engineering [8,17,19,20,23,24,30–32,37].

* Corresponding author.
E-mail addresses: sarthak.chatterjee92@gmail.com (S. Chatterjee), eigenvalue_debdipta@yahoo.in (D. Goswami), sudipto.dip15@gmail.com (S. Mukherjee), swagatam.das@isical.ac.in (S. Das).

http://dx.doi.org/10.1016/j.ins.2014.03.098
0020-0255/

Being a stochastic search process, PSO is not free from false and/or premature convergence, especially over multi-modal fitness landscapes. Quite often, PSO does not work very well, and may require considerable tuning of it's parameters to specifically adapt to deceptive or intensely multi-modal optimization problems. For detailed analysis of parameter-tuning in PSO, see van den Bergh [1], Trelea [34], Shi and Eberhart [33], Carlisle and Dozier [7], and Clerc [9]. A few more important works in the field of PSO parameter-selection are [4,11,12,33].

Mathematical analysis of the dynamics of PSO has attracted a good deal of research interest over the last decade. Most of such analytical studies that have so far been undertaken focuses on the trajectories of the particles and the choices of parameters that will guarantee the convergence and stability of the trajectories. These issues have been addressed by [2,3,5,6,8,16,25,26,34,35]. The sampling distributions of the PSO were investigated in [9,21,27,28]. A stagnation-state analysis of the particle dynamics in PSO is an extremely important work which has not been comprehensively studied earlier. [21] reported an analysis of the dynamic equation of the leading (globally best) particle. However, the authors only provided a sufficient stable region for the parametric space of the PSO algorithm. The present work attempts to introduce some degree of rigor on the dynamics of the so-called "fittest" particle in the stagnation state of a PSO algorithm with deterministic control parameters. In doing so, we arrive at some interesting conclusions. The points sampled by the leader particle satisfy a mathematical relation which shows that they lie on a line. Moreover, the dynamics of the leader particle is wholly governed by a parameter-dependent function which may take three different forms. If we assume that the swarm stays in the stagnation state forever, then this function may be seen to converge to 0. Some simple mathematics leads us to the conclusion that this function, which we call $g(t)$, is never negative for certain choices of parameters. This case is significant in pinpointing the fact that if this happens, PSO loses it's exploratory nature and hence, that these parameters may not be good choices for PSO.

A dynamic, stagnation-state analysis of the leader particle also gives us some information about the relation between the line on which the points sampled by the leader particle lie and the gradient of the objective function. Our work leads us to the conclusion that the aforementioned line is either orthogonal to the direction of the gradient of the objective function or has a descent direction. As will be shown, these two cases are intricately linked to the choice of the parameters and also to the sign of the function $g(t)$.

The organization of the paper is as follows. Section 2 reviews the classical PSO algorithm. Section 3 is devoted to the dynamic analysis of the leader particle during stagnation. Section 4 presents the analysis according to which we conclude that the points sampled by the leader particle lie on a line. We also discuss the dynamic and limitative behaviors of the dominant particle here. Section 5 gives the relationship between the dynamic behavior discussed and the choice of parameters. We also see how restricting parameters to lie in certain sets (or their unions) imposes strict rules on the sign that $g(t)$ possesses. This is carried forward in Section 6 to obtain valuable information about the relationship between the line on which the points sampled by the leader particle lie and the gradient of the objective function. Section 7 provides some numerical verification of the theory that we have put forth. The work ends with a short discussion about prospective future developments which seamlessly follow from the conclusions garnered forthwith.

2. The PSO algorithm

This section provides a brief introduction to the basic PSO Algorithm. PSO maintains a swarm containing m particles where $m \in \mathbb{N}$ is a constant. Each particle is characterized by a position, a velocity and a knowledge of it's own neighborhood, utilizing which it can share information about the hitherto best position it has attained with the other particles traversing the experimental search space. This so-called "best" position is governed by the fitness value or simply fitness, which determines a particular particle's progress towards coveted local or global minima of the objective function under consideration. The particles traverse the search space dynamically and their movement is governed by the following fundamental equations:

$$v_{ij}(k+1) = \omega v_{ij}(k) + C_1(p_{ij}(k) - x_{ij}(k)) + C_2(g_j(k) - x_{ij}(k)) \tag{1}$$

and

$$x_{ij}(k+1) = x_{ij}(k) + v_{ij}(k+1) \tag{2}$$

where

(1) $\boldsymbol{x}_i(k) = (x_{i1}(k), x_{i2}(k), \ldots, x_{in}(k))^T$ is the position of the i-th particle at iteration k, and $\boldsymbol{v}_i(k) = (v_{i1}(k), v_{i2}(k), \ldots, v_{in}(k))^T, i = (1, 2, \ldots, m)$ is the velocity of the i-th at iteration k.

(2) $\boldsymbol{p}_i(k) = (p_{i1}(k), p_{i2}(k), \ldots, p_{in}(k))^T$ and $\boldsymbol{g}_i(k) = (g_{i1}(k), g_{i2}(k), \ldots, g_{in}(k))^T$ are the personal best position and the neighborhood best position of particle i at iteration k respectively. Their values are defined as follows:

$$\boldsymbol{p}_i(k) = \underset{0 \leqslant t \leqslant k}{\operatorname{argmin}} f(\boldsymbol{x}_i(t)) \tag{3}$$

and

$$\boldsymbol{g}_i(k) = \underset{i \in N_i}{\operatorname{argmin}} f(\boldsymbol{p}_i(k)) \tag{4}$$

where N_i is the neighborhood of particle i. It is important at this juncture to note that $\boldsymbol{x}_i(k), \boldsymbol{v}_i(k), \boldsymbol{g}_i(k)$ and $\boldsymbol{p}_i(k)$ are all vectors of size $n \times 1$.

(3) The parameters ω, C_1 and C_2 satisfy the following for the canonical PSO algorithm:

$$\omega \in (0,1), \quad C_1 \sim U(0,\phi_1), \quad C_2 \sim U(0,\phi_2) \tag{5}$$

where ω, ϕ_1 and ϕ_2 are often termed as the accelerating coefficients or inertia weights.

In practice, for any particular PSO application, the choice of N_i is extremely crucial and often bears deep linkages to the purpose of the application in which the PSO algorithm is being used [22,29]. It is the population topology which ultimately determines the choice of N_i. Relevant in this regard are works like [14,15], in which this problem of choosing N_i suitably is studied empirically. Popular examples to be found in the literature are the gbest, lbest, and random topologies. In all our subsequent treatments, we have used the gbest topology.

In the canonical PSO algorithm, the parameters ω, C_1 and C_2 are randomly generated for each iteration, particle and dimension. Our work analyses the dynamics of the leading particle in PSO with constant parameters. This considerably reduces the mathematical complexity but, as we shall see, also gives us deep enough insights into the problem at hand to strike a delicate compromise between the ease of calculation and the conclusions obtained. The parameters are seen to satisfy:

$$\omega \in (-1,1), \quad C_1 > 0, \quad C_2 > 0, \quad 0 < C_1 + C_2 < 2(1+\omega) \tag{6}$$

thus providing us with a popular stable region of the PSO algorithm [2,8,13,34].

3. Dynamics of the leader particle

In this section, we provide a detailed analysis of the leader particle's dynamics and solve it in the special case of stagnation. At any instant of time t, the position and velocity updation equations are given by (1) and (2) as

$$v_{ij}(t+1) = \omega v_{ij}(t) + C_1(p_{ij}(t) - x_{ij}(t)) + C_2(g_{ij}(t) - x_{ij}(t)) \tag{7}$$

and

$$x_{ij}(t+1) = x_{ij}(t) + v_{ij}(t+1). \tag{8}$$

Substituting (1) in (2), we obtain

$$x_{ij}(t+1) = x_{ij}(t) + \omega v_{ij}(t) + C_1(p_{ij}(t) - x_{ij}(t)) + C_2(g_{ij}(t) - x_{ij}(t)). \tag{9}$$

Simplifying the above, we get

$$x_{ij}(t+1) = (1 - C_1 - C_2)x_{ij}(t) + \omega v_{ij}(t) + C_1 p_{ij}(t) + C_2 g_{ij}(t). \tag{10}$$

Replacing t by $(t-1)$ in (8) and substituting in (10), we get

$$x_{ij}(t+1) = (1 - C_1 - C_2)x_{ij}(t) + \omega(x_{ij}(t) - x_{ij}(t-1)) + C_1 p_{ij}(t) + C_2 g_{ij}(t), \tag{11}$$

which immediately reduces to the difference equation model of the PSO algorithm, given by

$$x_{ij}(t+1) = (1 + \omega - C_1 - C_2)x_{ij}(t) - \omega x_{ij}(t-1) + C_1 p_{ij}(t) + C_2 g_{ij}(t) \tag{12}$$

where $\omega \in (-1,1)$ and $C_1, C_2 > 0$. Now let us define the stagnation of the swarm as follows: If the whole swarm cannot improve itself from the K-th iteration up to the $(K+M)$-th iteration ($M \geqslant 3$), i.e. the global best position remains the same for this period, then the swarm is said to be in stagnation during the iterations $(K, K+M)$, and $\boldsymbol{x_i}(K)$ is the stagnation point where i is the leader particle.

Now, if the particle is in the stagnation state, then the global best position as well as the local best position of the leader particle will be the stagnation point, i.e. $\boldsymbol{p}_i(t) = \boldsymbol{g}_i(t) = \boldsymbol{x}_i(K), t \in [K, K+M)$. Thus, for the leader particle, we have the difference equation as

$$x_{ij}(t+1) = a_1 x_{ij}(t) + a_2 x_{ij}(t-1) + \phi x_{ij}(K). \tag{13}$$

for $t \in [K, K+M)$, where $\phi = C_1 + C_2, a_1 = 1 + \omega - \phi$ and $a_2 = -\omega$

To analyze and solve this difference equation, let us construct a column vector

$$\boldsymbol{Y}(t+1) = \begin{bmatrix} x(t+1) - x(K) \\ x(t) - x(K) \end{bmatrix}. \tag{14}$$

Here, we have dropped the subscripts i and j for the sake of convenience, since we are interested in the analysis of the leader particle only. Now the Eq. (13) can be written in an iterative functional form like the following,

$$\boldsymbol{Y}(t+1) = M\boldsymbol{Y}(t), \tag{15}$$

where

$$M = \begin{bmatrix} a_1 & a_2 \\ 1 & 0 \end{bmatrix}.$$

We can also write

$$\boldsymbol{Y}(K+t+1) = [M]^t \boldsymbol{Y}(K+1) \tag{16}$$

Now the eigenvalues of the matrix M are given by

$$\lambda_{1,2} = \frac{a_1}{2} \mp \frac{\sqrt{\Delta}}{2} \tag{17}$$

where $\Delta = a_1^2 + 4a_2 \Rightarrow a_1^2 - 4\omega$.

Case A: $\Delta > 0, \lambda_1, \lambda_2$ are real, $\lambda_1 \neq \lambda_2$:

In this case, since the two eigenvalues are distinct and $|\lambda_{1,2}| < 1$, the matrix M will be diagonalizable and we can write $M = P\Lambda P^{-1}$. So from Eq. (16) we can conclude

$$\boldsymbol{Y}(K+t+1) = P\Lambda^t P^{-1}\boldsymbol{Y}(K+1) \tag{18}$$

where

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix} \tag{19}$$

Now, upon finding the change-of-basis matrix (that which maps from the standard basis to that of the eigenvectors) we can write

$$P = \begin{bmatrix} \lambda_1 & \lambda_2 \\ 1 & 1 \end{bmatrix} \tag{20}$$

and thus, $P\Lambda^t P^{-1} = \frac{1}{\sqrt{\Delta}} \begin{bmatrix} -\lambda_1^{t+1} + \lambda_2^{t+1} & \lambda_1\lambda_2(\lambda_1^t - \lambda_2^t) \\ -\lambda_1^t + \lambda_2^t & \lambda_1\lambda_2(\lambda_1^{t-1} - \lambda_2^{t+1}) \end{bmatrix}$. Hence from the above expression and the Eq. (18) we can conclude that

$$x(K+t) = x(K) + \frac{1}{\sqrt{\Delta}}(\lambda_2^t - \lambda_1^t)[x(K+1) - x(K)] \tag{21}$$

Case B: $\Delta < 0, \lambda_1, \lambda_2$ are complex conjugates:

In this case, too, the eigenvalues being different, the matrix M is diagonalizable. Here $\lambda_{1,2} = |\lambda_{1,2}|e^{\pm i\theta}$, where $|\lambda_{1,2}| = \sqrt{\omega}$ and $\theta = \tan^{-1}\frac{\sqrt{4\omega - a_1^2}}{a_1}$. Here also Eq. (18) will hold and

$$P\Lambda^t P^{-1} = \frac{1}{\sqrt{\Delta}} \begin{bmatrix} \omega^{(t+1)/2}(e^{i(t+1)\theta} - e^{-i(t+1)\theta}) & \omega^{(t+2)/2}(e^{-it\theta} - e^{it\theta}) \\ \omega^{t/2}(e^{it\theta} - e^{-it\theta}) & \omega^{(t+1)/2}(e^{-i(t-1)\theta} - e^{i(t-1)\theta}) \end{bmatrix}$$

Since $\sin\theta = \frac{-\sqrt{\Delta}}{2\sqrt{\omega}}$, the above matrix relation along with the Eq. (18) boils down to

$$x(K+t) = x(K) + \frac{\sin(t\theta)}{\sin\theta}\omega^{(t-1)/2}[x(K+1) - x(K)] \tag{22}$$

Here, since $\sin\theta \neq 0$ and $|\omega| < 1$

$$\frac{\sin(t\theta)}{\sin\theta}\omega^{(t-1)/2} \to 0 \tag{23}$$

as $t \to \infty$.

Case C: $\Delta = 0, \lambda_1$ and λ_2 are real and equal:

Here, it can be easily shown that the algebraic multiplicity of the matrix M is less than its geometric multiplicity, the former being two and the latter one. So the matrix cannot be diagonalized. But, since the characteristic polynomial of M splits, we can conclude that M is similar to a Jordan canonical matrix given by

$$J = \begin{bmatrix} \frac{a_1}{2} & 1 \\ 0 & \frac{a_1}{2} \end{bmatrix} \tag{24}$$

i.e. $M = PJP^{-1}$ for some P. One eigenvector of M is given by $v_1 = \begin{bmatrix} \frac{a_1}{2} \\ 1 \end{bmatrix}$ and another generalized eigenvector v_2 can be found out using the equation $(M - \frac{a_1}{2}I)v_2 = v_1$. One particular solution of v_2 (chosen arbitrarily) is $v_2 = \begin{bmatrix} 0 \\ -\frac{2}{a_1} \end{bmatrix}$. So the change-of-basis matrix is $P = \begin{bmatrix} \frac{a_1}{2} & 0 \\ 1 & -\frac{2}{a_1} \end{bmatrix}$. Thus from Eq. (16) we can conclude

$$\boldsymbol{Y}(K+t+1) = PJ^tP^{-1}\boldsymbol{Y}(K+1), \tag{25}$$

where J is given by Eq. (24). Now,

$$PJ^tP^{-1} = \begin{bmatrix} \left(\frac{a_1}{2}\right)^t + t\left(\frac{a_1}{2}\right)^t & -t\left(\frac{a_1}{2}\right)^{t+1} \\ \left(\frac{a_1}{2}\right)^{t-1} + (t-1)\left(\frac{a_1}{2}\right)^{t-1} & -(t-1)\left(\frac{a_1}{2}\right)^t \end{bmatrix}.$$

So, from Eq. (25) we can conclude that

$$x(K+t) = x(K) + t\left(\frac{a_1}{2}\right)^{t-1}[x(K+1) - x(K)]. \tag{26}$$

Here $a_1^2 = 4\omega$ and since

$$|\omega| < 1, \quad |a_1| < 2. \tag{27}$$

Therefore, if the whole swarm is in the stagnation state, the positions sampled by the leader particle i during the stagnation state follow the rule given below,

$$\boldsymbol{x_i}(K+t) = \boldsymbol{x_i}(K) + g(t)(\boldsymbol{x_i}(K+1) - \boldsymbol{x_i}(K)), \quad t \in [0, M), \tag{28}$$

where $x_i(K)$ is the stagnation point, and

$$g(t) = \begin{cases} \frac{(\lambda_2^t - \lambda_1^t)}{\sqrt{\Delta}}, & \Delta > 0 \\ t\left(\frac{a_1}{2}\right)^{t-1}, & \Delta = 0 \\ \frac{\sin(t\theta)}{\sin\theta}(\sqrt{\omega})^{t-1}, & \Delta < 0 \end{cases}. \tag{29}$$

Here, $\Delta = a_1^2 - 4\omega$, and λ_1, λ_2 are defined by Eq. (17).

4. Dynamic and limitative behavior of the leader particle

The conclusions of the previous section lead us to investigate in detail about the dynamic and limitative behaviors of the leader particle.

4.1. Dynamic behavior of the leader

From Eq. (28), we can infer that the positions sampled by the leader particle i lie on a line identified by $\boldsymbol{x_i}(K+1)$ and $\boldsymbol{x_i}(K)$. If $x_{ij}(K+1) = x_{ij}(K)$,

$$x_{ij}(K+t) = x_{ij}(K), \quad t \in [0, M) \tag{30}$$

This implies that the jth components of the positions of the leader particle remain constant during the iterations $[K, K+M)$, i.e., the trajectory of the particle i lies on an $(n-1)$ dimensional subspace of Ω.

If $x_{ij}(K+1) \neq x_{ij}(K)$, then the positions sampled by the leader particle will oscillate along the line identified by $\boldsymbol{x_i}(K+1)$ and $\boldsymbol{x_i}(K)$, the amplitudes being smaller with increasing t.

4.2. Limiting behavior of the leader

Now we assume that $M \to \infty$, i.e., the swarm stays in the stagnation state after the Kth iteration. Since $|\omega| < 1$, from Eq. (29), (23) and (27) and keeping in mind that $|\lambda_{1,2}| < 1$ for $\Delta > 0$, we can conclude (see Appendix A)

$$\lim_{t\to\infty} g(t) = 0, \tag{31}$$

and therefrom

$$\lim_{t\to\infty} \boldsymbol{x}(K+t) = \boldsymbol{x}(K). \tag{32}$$

Thus, if the stagnation continues, the positions sampled by the leader particle will finally converge to the stagnation point.

5. Dynamic behavior and parameter choice

The derivation presented in the previous section leads us to the conclusion that the points sampled by the leader particle during stagnation lie on a line identified by $\boldsymbol{x_i}(K+1)$ and $\boldsymbol{x_i}(K)$. The purpose of the section in question is to show that the entire stable region

$$\omega \in (-1,1), \quad C_1 > 0, \quad C_2 > 0, \quad 0 < C_1 + C_2 < 2(1+\omega) \tag{33}$$

can be divided into two sub-regions wherein for the first, $g(t) \geqslant 0$ for any t and for the second $g(t) \geqslant 0$ for some t and $g(t) < 0$ for other t. In other words, the stagnation-time sample points of the leader particle are heavily dependent on the choice of parameters ω, C_1 and C_2. If the aforementioned parameters lie inside the first sub-region, then the points sampled by the leader particle during stagnation always lie on one side of the stagnation point. However, restricting them to lie in the second sub-region ensures that the sampled points corresponding to the leader particle lie on either side of the stagnation point.

We now investigate the behavior of the function $g(t)$ corresponding to the values of ω, C_1 and C_2.

Case 1: $\omega = 0$, $C_1 + C_2 = 1$

In this case, $\omega = 0$ and $C_1 + C_2 = 1$. Consequently, $a_1 = 0$, which means that $\Delta = 0$. Hence, $g(t) = 0 \forall t \in \mathbb{N}$.

Case 2: $\Delta > 0$

At the outset, we notice that $\Delta > 0$ gives $a_1^2 > 4\omega$, and hence either $a_1 > 2\sqrt{\omega}$ or $\omega < 0$.

a. $0 < \lambda_1 < \lambda_2$

We first take up the case where $0 < \lambda_1 < \lambda_2$. This shows that either $a_1 > -\sqrt{\Delta}$ or $a_1 > \sqrt{\Delta}$. Combining the two, we have $a_1^2 > \Delta$. Since $\Delta = a_1^2 - 4\omega$, this gives $\omega > 0$. Hence, the only possibility is $a_1 > 2\sqrt{\omega}$. Since $\lambda_2 > \lambda_1$, $\frac{a_1}{2} + \frac{\sqrt{\Delta}}{2} > \frac{a_1}{2} - \frac{\sqrt{\Delta}}{2}$ and hence $\Delta > 0$. Since $a_1 > 2\sqrt{\omega}$, and $a_1 = 1 + \omega - C_1 - C_2$, we have $1 + \omega - C_1 - C_2 > 2\sqrt{\omega}$ and therefore $C_1 + C_2 < (1 - \sqrt{\omega})^2$. Hence, from the information that $|\omega| < 1$ and $C_1 + C_2 > 0$, we conclude that.

- $0 < \omega < 1$
- $0 < C_1 + C_2 < (1 - \sqrt{\omega})^2$

b. $\lambda_1 < 0 < \lambda_2$, $|\lambda_1| < |\lambda_2|$

We can write $\frac{a_1}{2} + \frac{\sqrt{\Delta}}{2} > 0$ and $\frac{a_1}{2} - \frac{\sqrt{\Delta}}{2} < 0$. This leads to $a_1 < \sqrt{\Delta}$ and $a_1 > -\sqrt{\Delta}$. Writing in a single statement, this gives

$$-\sqrt{\Delta} < a_1 < \sqrt{\Delta}$$

In other words, $a_1^2 < \Delta$. Putting $\Delta = a_1^2 - 4\omega$, we get $\omega < 0$, which immediately permits us to write down

$$-1 < \omega < 0$$

since the theoretical lower limit of ω is -1. Since $|\lambda_1| < |\lambda_2|$,

$$\left|\frac{a_1}{2} - \frac{\sqrt{\Delta}}{2}\right| < \left|\frac{a_1}{2} + \frac{\sqrt{\Delta}}{2}\right|.$$

Since $\lambda_1 < 0$ and $\lambda_2 > 0$, we can write

$$\frac{\sqrt{\Delta}}{2} - \frac{a_1}{2} < \frac{a_1}{2} + \frac{\sqrt{\Delta}}{2}$$

which means that $a_1 > 0$. Hence $1 + \omega - C_1 - C_2 > 0$ and hence $0 < C_1 + C_2 < 1 + \omega, 0$ being the lower bound of $C_1 + C_2$. For this case therefore, the conclusions that can be drawn are

- $-1 < \omega < 0$
- $0 < C_1 + C_2 < 1 + \omega$

Case 3: $\Delta = 0$ *and* $a_1 > 0$

Since $\Delta = 0, a_1^2 = 4\omega$ and hence $a_1 = 2\sqrt{\omega}$ since a_1 is positive. Therefore, $1 + \omega - C_1 - C_2 = 2\sqrt{\omega}$ and hence

$$C_1 + C_2 = (1 - \sqrt{\omega})^2.$$

Since $a_1 > 0$ and $a_1^2 = 4\omega$, we have $\omega > 0$. We are therefore in a position to modify the result of case 2a. to include the equality sign and immediately write down

- $0 < \omega < 1$ (as has been already obtained)
- $0 < C_1 + C_2 \leqslant (1 - \sqrt{\omega})^2$

The above discussion on the three cases and their sub-cases points to those values of the parameters ω, C_1, C_2 for which $g(t) > 0$ for any $t \in \mathbb{N}$.

Case 4: $\Delta > 0$, $\lambda_1 < 0 < \lambda_2$, $|\lambda_1| = |\lambda_2|$

The first part of the calculation proceeds in exactly the same way as that of case 2b. Writing $|\lambda_1| = |\lambda_2|$ gives

$$\left|\frac{a_1}{2} - \frac{\sqrt{\Delta}}{2}\right| = \left|\frac{a_1}{2} + \frac{\sqrt{\Delta}}{2}\right|.$$

Therefore this gives

$$\frac{\sqrt{\Delta}}{2} - \frac{a_1}{2} = \frac{a_1}{2} + \frac{\sqrt{\Delta}}{2}$$

and hence $a_1 = 0$. In other words, $1 + \omega - C_1 - C_2 = 0$ and consequently,

$$1 + \omega = C_1 + C_2.$$

This allows us to modify 2b. to include the equality sign and hence conclude that

- $-1 < \omega < 0$ (as has been already obtained)
- $0 < C_1 + C_2 \leqslant 1 + \omega$

This case points to the condition where λ_1 and λ_2 are equispaced about 0, λ_1 being negative and λ_2 positive. As is evident from Eq. (29), $g(t) > 0$ for any odd t and $g(t) = 0$ for any even t when the parameters satisfy the above.

Case 5: $\Delta > 0$, $\lambda_1 < \lambda_2 < 0$

We readily obtain

$$\frac{a_1}{2} - \frac{\sqrt{\Delta}}{2} < 0$$

and

$$\frac{a_1}{2} + \frac{\sqrt{\Delta}}{2} < 0.$$

This means that $a_1 < \sqrt{\Delta}$ and $a_1 < -\sqrt{\Delta}$ or that $a_1^2 < \Delta$. This means that $a_1^2 < a_1^2 - 4\omega$ or that $\omega < 0$ and hence $-1 < \omega < 0$. We also see that $\lambda_1 < \lambda_2$. Therefore,

$$\frac{a_1}{2} - \frac{\sqrt{\Delta}}{2} < \frac{a_1}{2} + \frac{\sqrt{\Delta}}{2}.$$

or $\sqrt{\Delta} > 0$. Also, since $a_1 < -\sqrt{\Delta}$ and $\sqrt{\Delta}$ is positive, we have $a_1 < 0$. In other words, $1 + \omega - C_1 - C_2 < 0$ or

$$C_1 + C_2 > 1 + \omega.$$

Since the canonical PSO predicts that for stability $C_1 + C_2 < 2(1 + \omega)$, we can summarize that

- $-1 < \omega < 0$
- $1 + \omega < C_1 + C_2 < 2(1 + \omega)$

Case 6: $\Delta > 0$, $\lambda_1 < 0 < \lambda_2$, $|\lambda_1| > |\lambda_2|$

Here

$$\lambda_1 = \frac{a_1}{2} - \frac{\sqrt{\Delta}}{2} < 0$$

and

$$\lambda_2 = \frac{a_1}{2} + \frac{\sqrt{\Delta}}{2} > 0.$$

Hence, $a_1 < \sqrt{\Delta}$ and $a_1 > -\sqrt{\Delta}$ and consequently, $a_1^2 < \Delta$. Therefore, $a_1^2 < a_1^2 - 4\omega$ and hence $\omega < 0$. In other words, $-1 < \omega < 0$.

Again, $|\lambda_1| > |\lambda_2|$ or

$$\left|\frac{a_1}{2} - \frac{\sqrt{\Delta}}{2}\right| > \left|\frac{a_1}{2} + \frac{\sqrt{\Delta}}{2}\right|.$$

So,

$$\frac{\sqrt{\Delta}}{2}-\frac{a_1}{2}>\frac{a_1}{2}+\frac{\sqrt{\Delta}}{2}$$

and therefore $a_1<0$. This gives $1+\omega-C_1-C_2<0$ or that

$$C_1+C_2>1+\omega.$$

Thus, the conclusions drawn are

- $-1<\omega<0$
- $1+\omega<C_1+C_2<2(1+\omega)$

Which is the same as case 5.

Case 7: $\Delta=0$, $a_1<0$

Here, $\Delta=0$ which means that $a_1^2=4\omega$. Since $a_1<0$, we take the negative square root and write $a_1=-2\sqrt{\omega}$. This means that $1+\omega-C_1-C_2=-2\sqrt{\omega}$ and therefore

$$C_1+C_2=(1+\sqrt{\omega})^2.$$

Since, the right-hand-side of the above can be expanded to write $(1+\omega+2\sqrt{\omega})$, we can show by applying the arithmetic mean-geometric mean inequality on 1 and ω (which we can because $\omega>0$) that

$$1+\omega>2\sqrt{\omega}.$$

Hence $(1+\sqrt{\omega})^2<2(1+\omega)$ and therefore

$$C_1+C_2<2(1+\omega).$$

Further, since $\sqrt{\omega}$ is positive, it is almost trivial to show that $1+\sqrt{\omega}>1-\sqrt{\omega}$ and hence that

$$C_1+C_2>(1-\sqrt{\omega})^2.$$

We can hence write

- $0<\omega<1$
- $(1-\sqrt{\omega})^2<C_1+C_2<2(1+\omega)$

For the cases 5, 6 and 7 we find that $g(t)>0$ for any $t\in T_1=\{t\in\mathbb{N}|t\ is\ odd\}$ and $g(t)<0$ for any $t\in T_2=\{t\in\mathbb{N}|t\ is\ even\}$.

Case 8: $\Delta<0$

In this case, $a_1^2-4\omega<0$ and hence $a_1^2<4\omega$. We conclude that $\omega>0$ and hence $0<\omega<1$. Again $a_1<2\sqrt{\omega}$ which implies that $1+\omega-C_1-C_2<2\sqrt{\omega}$ and therefore

$$C_1+C_2>(1-\sqrt{\omega})^2.$$

We therefore can write

- $0<\omega<1$
- $(1-\sqrt{\omega})^2<C_1+C_2<2(1+\omega)$

The case 8 leads us to the conclusion that $g(t)>0$ for any $t\in T_1=\{t\in\mathbb{N}|\sin(t\theta)>0\}$ and that $g(t)<0$ for any $t\in T_2=\{t\in\mathbb{N}|\sin(t\theta)<0\}$.

5.1. Subregions of the parameter space

We now proceed to define four sets S_i where $i=1,2,3,4$ as follows:

$$S_1=\{(\omega,C_1,C_2)|-1<\omega<0,0<C_1+C_2\leqslant 1+\omega\} \tag{34a}$$

$$S_2=\left\{(\omega,C_1,C_2)|0\leqslant\omega<1,0<C_1+C_2\leqslant(1-\sqrt{\omega})^2\right\}-\{(\omega,C_1,C_2)|\omega=0,C_1+C_2=1\} \tag{34b}$$

$$S_3=\{(\omega,C_1,C_2)|-1<\omega<0,1+\omega<C_1+C_2<2(1+\omega)\} \tag{34c}$$

$$S_4=\left\{(\omega,C_1,C_2)|0\leqslant\omega<1,(1-\sqrt{\omega})^2<C_1+C_2<2(1+\omega)\right\} \tag{34d}$$

We therefore find that if $M \to \infty$, or in other words, if the swarm stays in stagnation forever, we have the following results:

- If $\omega = 0, C_1 + C_2 = 1$, then $g(t) = 0$ for any $t \in \mathbb{N}$.
- If the parameters $(\omega, C_1, C_2) \in S_1 \cup S_2$, then $g(t) \geqslant 0$ for all $t \in \mathbb{N}$. Furthermore, there exits an infinite set $T' \subset \mathbb{N}$ such that $g(t) > 0$ for all $t \in T'$.
- If the parameters $(\omega, C_1, C_2) \in S_3 \cup S_4$, then there exist two infinite sets T_1 and T_2, such that $g(t) > 0$ for all $t \in T_1$ while $g(t) < 0$ for all $t \in T_2$.

The above discussion serves to illuminate one extremely vital ingredient when it comes to stagnation in PSO. Since PSO is prima facie an exploratory algorithm, the leader particle needs to search on either side of the stagnation point. Since $(\omega, C_1, C_2) \in S_1 \cup S_2$ ensures that $g(t)$ is never less than 0 for any $t \in \mathbb{N}$, we can conclude that the leader particle explores the landscape on only one side of the stagnation point $\boldsymbol{x_i}(K)$ and completely ignores the other side. This highlights that the parameter subspace $S_1 \cup S_2$ may not be a good choice when it comes to PSO.

6. Objective function gradient and the parameter choice

In this section, we will look into the surface of the objective function and determine some relation between its gradient and the values of the parameter.

Let us assume that the whole swarm is in the stagnation state permanently i.e. $M \to \infty$ and the objective function $f(x)$ is continuously differentiable at least in a small neighborhood of the stagnation point.

During stagnation for the leader particle i, we can write,

$$f(\boldsymbol{x_i}(K+t)) \geqslant f(\boldsymbol{x_i}(K)), \quad t = 1, 2, \ldots$$

This relation, along with Eq. (28), gives us

$$f(\boldsymbol{x_i}(K) + g(t)\boldsymbol{d}) - f(\boldsymbol{x_i}(K)) \geqslant 0, \quad t = 1, 2, \ldots, \tag{35}$$

where $\boldsymbol{d} = \boldsymbol{x_i}(K+1) - \boldsymbol{x_i}(K)$.

A: $(\omega, C_1, C_2) \in S_1 \cup S_2$.

From the results of the previous section, in this region $g(t) \geqslant 0, \forall t \in \mathbb{N}$; and $\exists T' \subset \mathbb{N}$ such that $g(t) > 0, \forall t \in T'$. Combining with Eq. (35), we have

$$\frac{f(\boldsymbol{x_i}(K) + g(t)\boldsymbol{d}) - f(\boldsymbol{x_i}(K))}{g(t)} \geqslant 0, t \in T'.$$

If $t \in T', t \to \infty$, along with (31) we have,

$$\nabla f(\boldsymbol{x_i}(K))^T \boldsymbol{d} \geqslant 0. \tag{36}$$

Therefore $-\boldsymbol{d} = \boldsymbol{x_i}(K) - \boldsymbol{x_i}(K+1)$ is along the descent direction of the objective function at the stagnation point $\boldsymbol{x_i}(K)$.

B: $(\omega, C_1, C_2) \in S_3 \cup S_4$.

Again from the results of the previous section, we know in this region there exists infinite sets $T_1 \in \mathbb{N}$ and $T_2 \in \mathbb{N}$, such that $g(t) > 0$ for $t \in T_1$ and $g(t) < 0$ for $t \in T_2$. Combining with Eq. (35), we have

$$\frac{f(\boldsymbol{x_i}(K) + g(t)\boldsymbol{d}) - f(\boldsymbol{x_i}(K))}{g(t)} \geqslant 0, \quad t \in T_1.$$

and

$$\frac{f(\boldsymbol{x_i}(K) + g(t)\boldsymbol{d}) - f(\boldsymbol{x_i}(K))}{g(t)} \leqslant 0, \quad t \in T_2.$$

Again letting $t \to \infty$, with the help of Eq. (31), we have both

$$\nabla f(\boldsymbol{x_i}(K))^T \boldsymbol{d} \geqslant 0.$$

and

$$\nabla f(\boldsymbol{x_i}(K))^T \boldsymbol{d} \leqslant 0.$$

Therefore we have

$$\nabla f(\boldsymbol{x_i}(K))^T \boldsymbol{d} = 0. \tag{37}$$

So, $\boldsymbol{d} = \boldsymbol{x_i}(K+1) - \boldsymbol{x_i}(K)$ is orthogonal to the gradient of the objective function at the stagnation point.

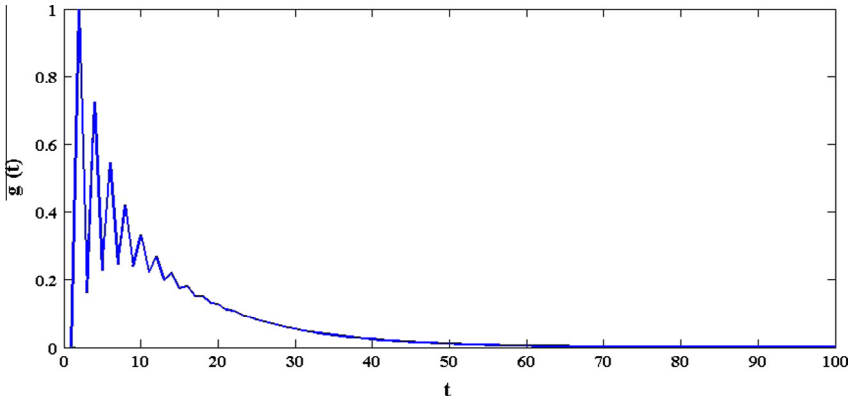

Fig. 1. Plot of $g(t)$ vs t with $C_1 = C_2 = 0.07$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_1$.

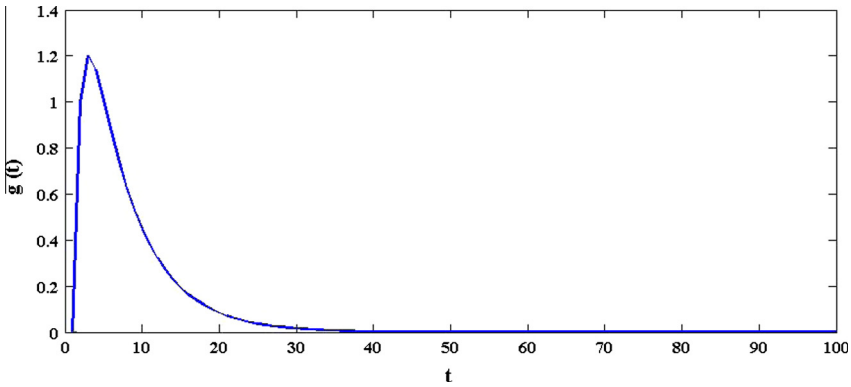

Fig. 2. Plot of $g(t)$ vs t with $C_1 = C_2 = 0.05$ and $\omega = 0.3$, i.e. $(\omega, C_1, C_2) \in S_2$.

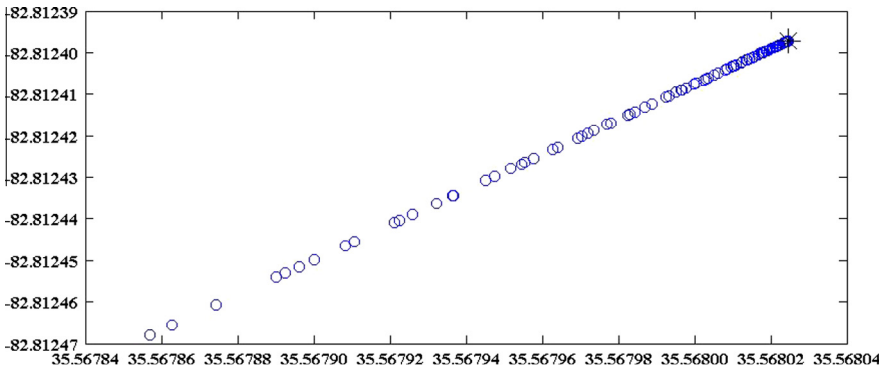

Fig. 3. Positions sampled by the leader particle with $C_1 = C_2 = 0.07$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_1$ on the f_2 i.e., shifted Schwefel's function.

7. Simulation results

In this section we provide some numerical results generated through computer simulations to back the theoretical results obtained in the previous sections. We choose several parameter values in $S_1 \cup S_2$ and $S_3 \cup S_4$ as well. For each case we provide the time-variation of $g(t)$ and the sample history of the leader particle.

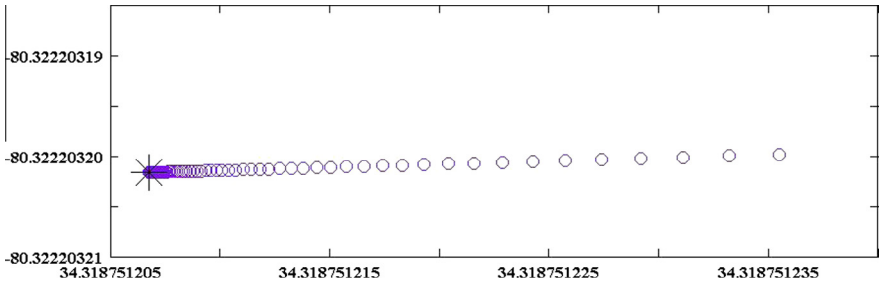

Fig. 4. Positions sampled by the leader particle with $C_1 = C_2 = 0.07$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_1$ on the f_4 i.e., shifted Schwefel's function with noise in fitness.

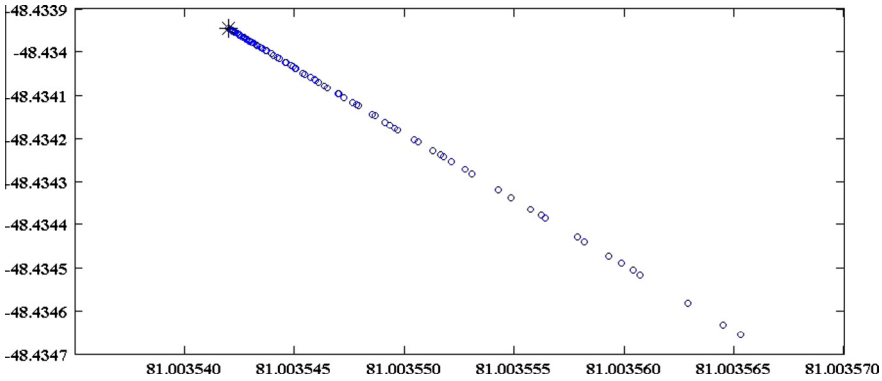

Fig. 5. Positions sampled by the leader particle with $C_1 = C_2 = 0.07$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_1$ on the f_6 i.e., shifted Rosenbrock's function.

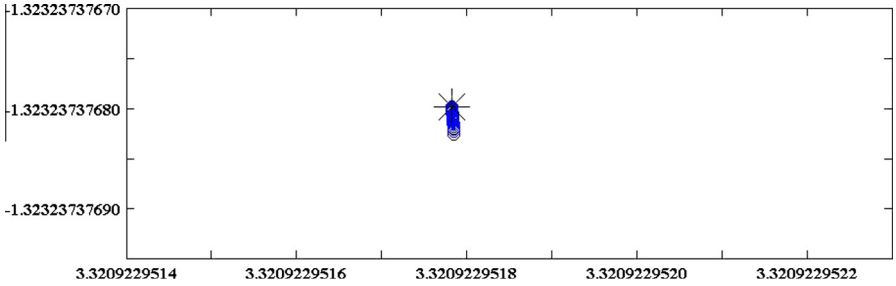

Fig. 6. Positions sampled by the leader particle with $C_1 = C_2 = 0.07$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_1$ on the f_{15} i.e., Hybrid Composition Function.

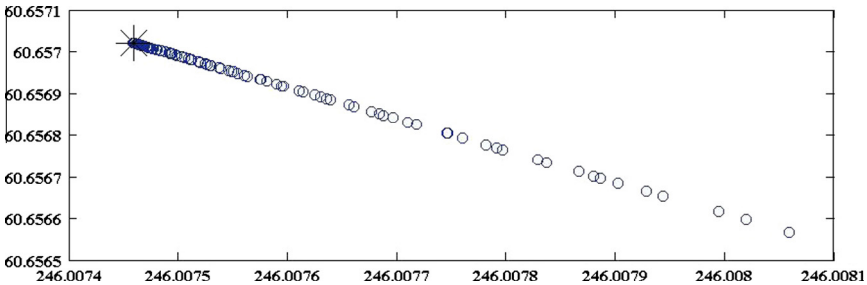

Fig. 7. Positions sampled by the leader particle with $C_1 = C_2 = 0.07$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_1$ on the f_7 i.e., shifted rotated Griewank's function without bounds.

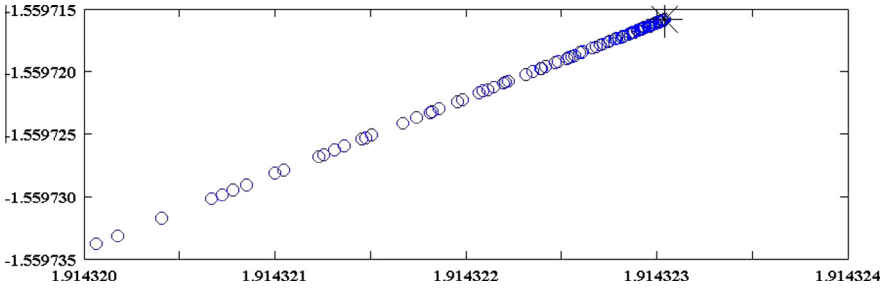

Fig. 8. Positions sampled by the leader particle with $C_1 = C_2 = 0.07$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_1$ on the f_{10} i.e., Shifted Rotated Rastrigin's Function.

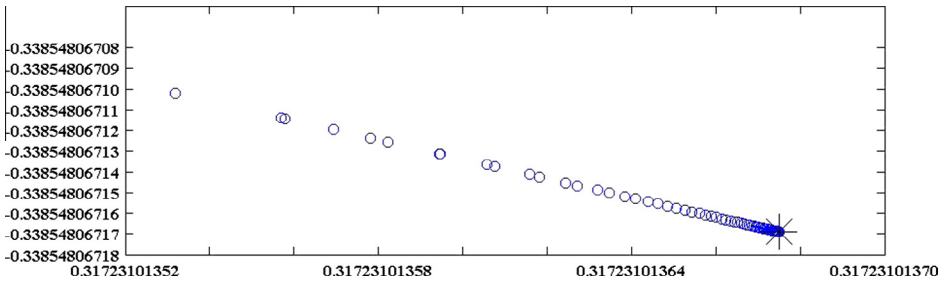

Fig. 9. Positions sampled by the leader particle with $C_1 = C_2 = 0.07$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_1$ on the f_{11} i.e., Shifted Rotated Weierstrass Function.

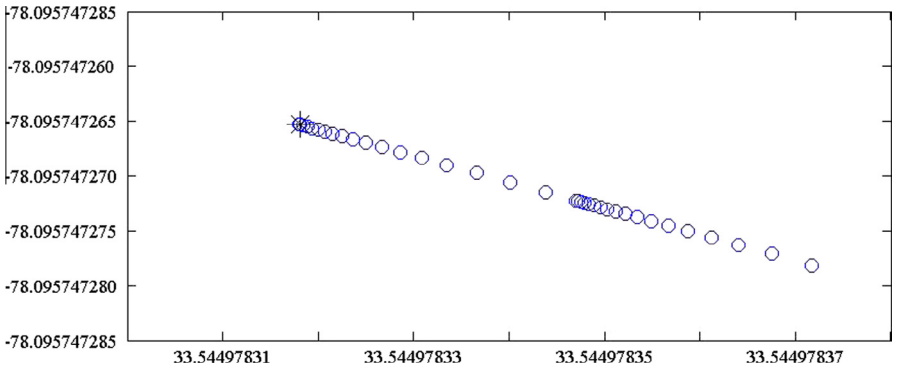

Fig. 10. Positions sampled by the leader particle with $C_1 = C_2 = 0.05$ and $\omega = 0.3$, i.e. $(\omega, C_1, C_2) \in S_2$ on the f_2 i.e., Shifted Schwefel's Function.

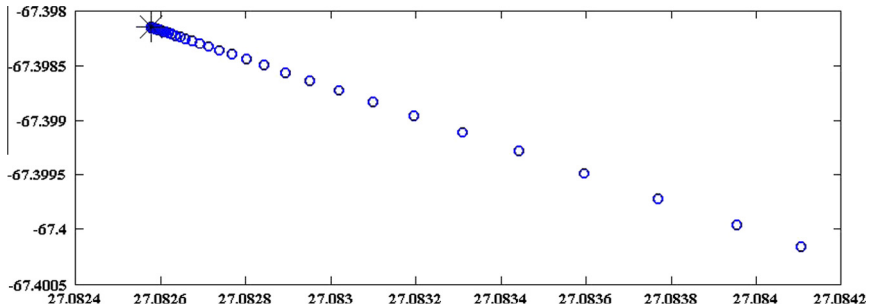

Fig. 11. Positions sampled by the leader particle with $C_1 = C_2 = 0.05$ and $\omega = 0.3$, i.e. $(\omega, C_1, C_2) \in S_2$ on the f_4 i.e., Shifted Schwefel's Function with noise in fitness.

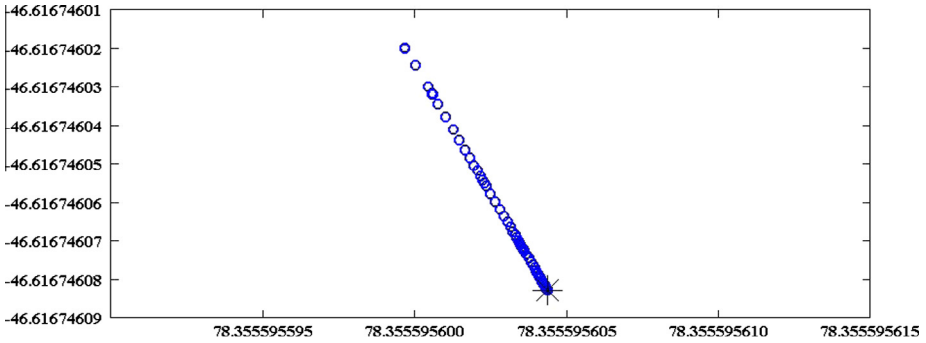

Fig. 12. Positions sampled by the leader particle with $C_1 = C_2 = 0.05$ and $\omega = 0.3$, i.e. $(\omega, C_1, C_2) \in S_2$ on the f_6 i.e., Shifted Rosenbrock's Function.

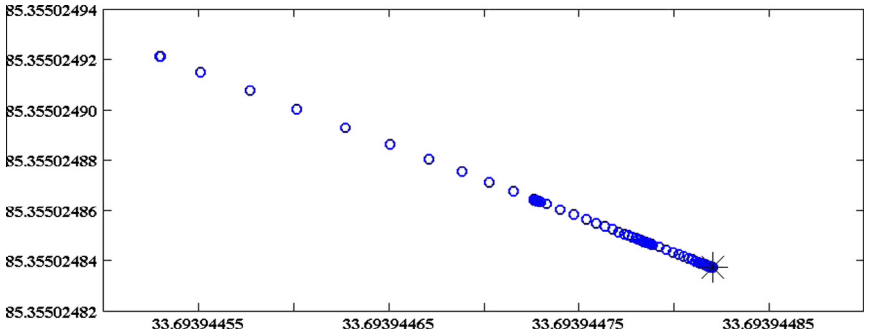

Fig. 13. Positions sampled by the leader particle with $C_1 = C_2 = 0.05$ and $\omega = 0.3$, i.e. $(\omega, C_1, C_2) \in S_2$ on the f_7 i.e., Shifted Rotated Griewank's Function without Bounds.

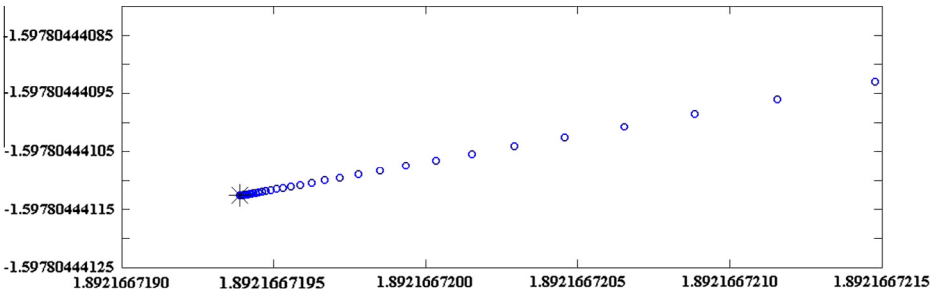

Fig. 14. Positions sampled by the leader particle with $C_1 = C_2 = 0.05$ and $\omega = 0.3$, i.e. $(\omega, C_1, C_2) \in S_2$ on the f_{10} i.e., Shifted Rotated Rastrigin's Function.

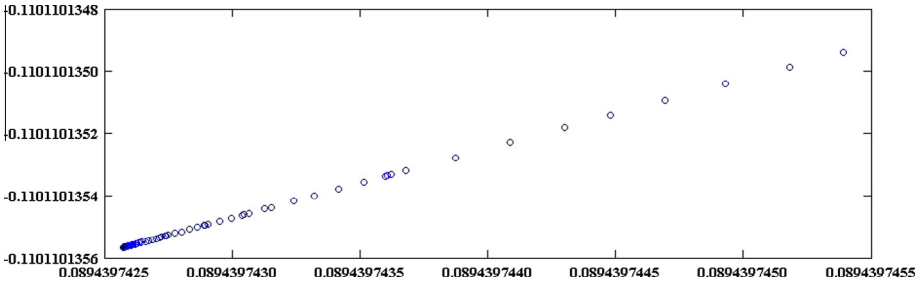

Fig. 15. Positions sampled by the leader particle with $C_1 = C_2 = 0.05$ and $\omega = 0.3$, i.e. $(\omega, C_1, C_2) \in S_2$ on the f_{11} i.e., Shifted Rotated Weierstrass Function.

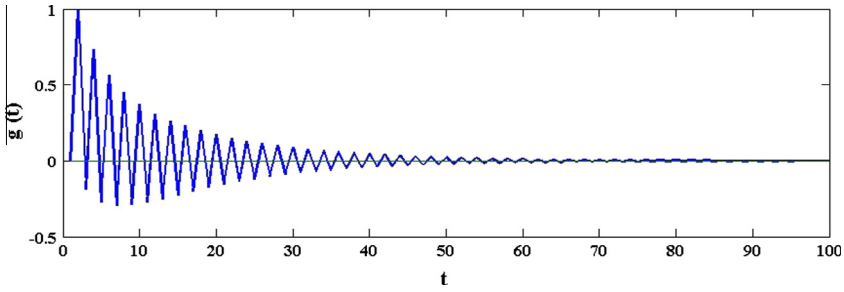

Fig. 16. Plot of $g(t)$ vs t with $C_1 = C_2 = 0.244$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_3$.

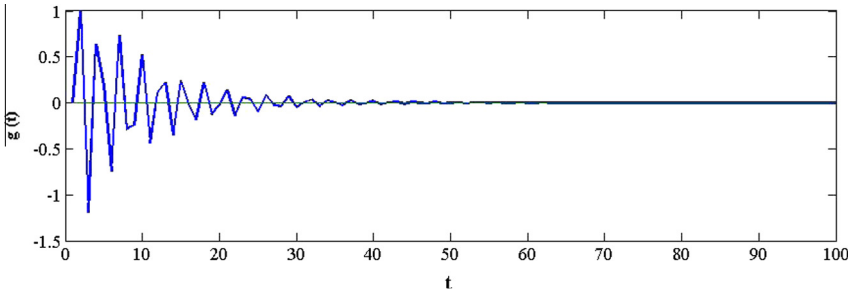

Fig. 17. Plot of $g(t)$ vs t with $C_1 = C_2 = 1.5$ and $\omega = 0.8$, i.e. $(\omega, C_1, C_2) \in S_4$.

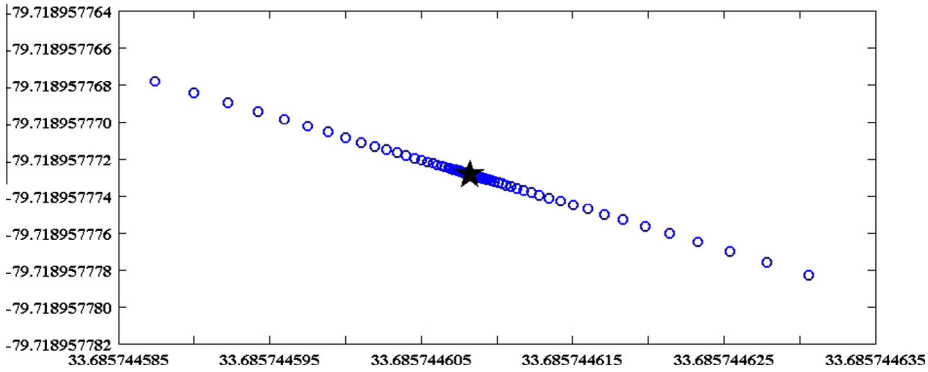

Fig. 18. Positions sampled by the leader particle with $C_1 = C_2 = 0.244$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_3$ on the f_4 i.e., Shifted Schwefel's Function with noise in fitness.

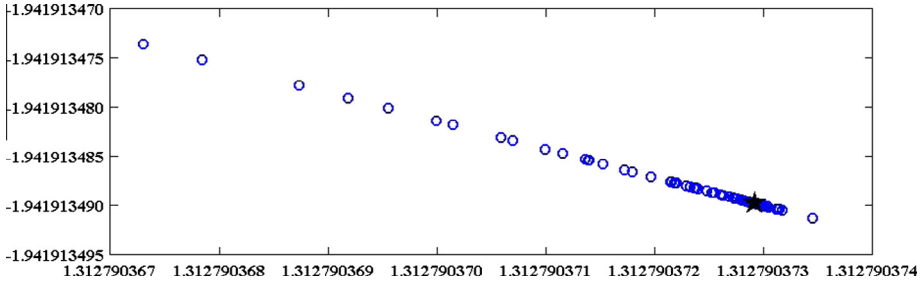

Fig. 19. Positions sampled by the leader particle with $C_1 = C_2 = 0.244$ and $\omega = -0.7$, i.e. $(\omega, C_1, C_2) \in S_3$ on the f_{10} i.e., Shifted Rotated Rastrigin's Function.

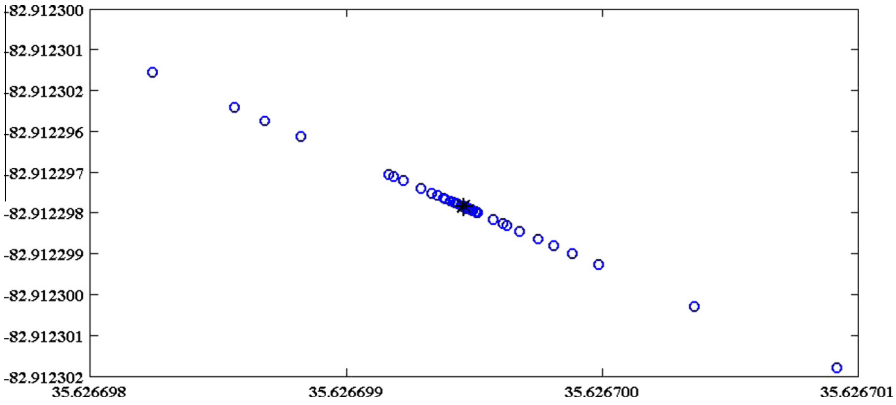

Fig. 20. Positions sampled by the leader particle with $C_1 = C_2 = 1.5$ and $\omega = 0.8$, i.e. $(\omega, C_1, C_2) \in S_4$ on the f_2 i.e., Shifted Schwefel's Function.

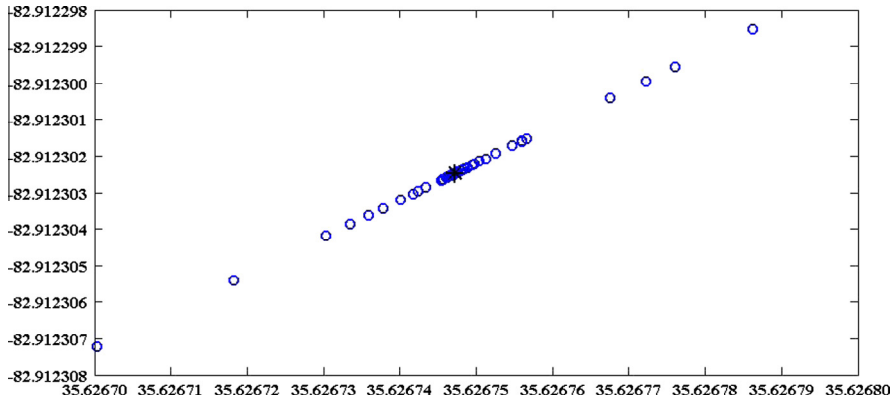

Fig. 21. Positions sampled by the leader particle with $C_1 = C_2 = 1.5$ and $\omega = 0.8$, i.e. $(\omega, C_1, C_2) \in S_4$ on the f_4 i.e., Shifted Schwefel's Function with noise in fitness.

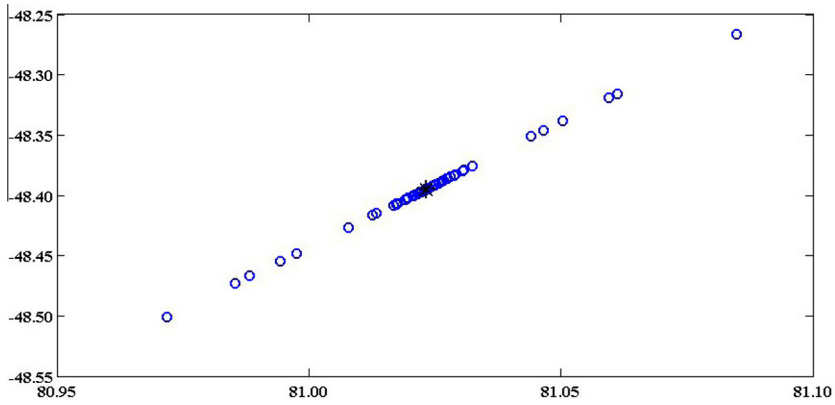

Fig. 22. Positions sampled by the leader particle with $C_1 = C_2 = 1.5$ and $\omega = 0.8$, i.e. $(\omega, C_1, C_2) \in S_4$ on the f_6 i.e., Shifted Rosenbrock's Function.

7.1. Parameter choices in subregion $S_1 \cup S_2$

From the conclusion of Section 5 we know that when $(\omega, C_1, C_2) \in S_1 \cup S_2, g(t) \geqslant 0$, the leader particle searches only one side of the stagnation point along the line $\boldsymbol{x_i}(K) - \boldsymbol{x_i}(K+1)$.

Here, two parameter combinations are chosen within the subregion $S_1 \cup S_2 : (C_1 = C_2 = 0.07, \omega = -0.7), (C_1 = C_2 = 0.05, \omega = 0.3)$; the first lies in S_1 and the second in S_2. For each combination we have provided the plot of $g(t)$ with time.

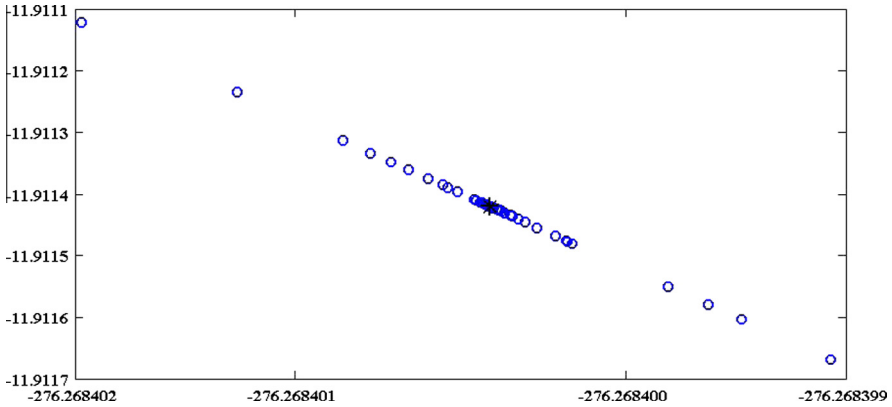

Fig. 23. Positions sampled by the leader particle with $C_1 = C_2 = 1.5$ and $\omega = 0.8$, i.e. $(\omega, C_1, C_2) \in S_4$ on the f_7 i.e., Shifted Rotated Griewank's Function without Bounds.

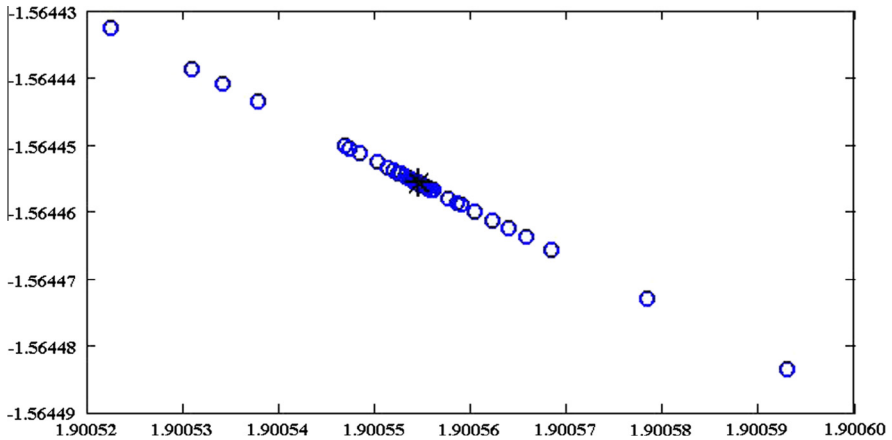

Fig. 24. Positions sampled by the leader particle with $C_1 = C_2 = 1.5$ and $\omega = 0.8$, i.e. $(\omega, C_1, C_2) \in S_4$ on the f_{10} i.e., Shifted Rotated Rastrigin's Function.

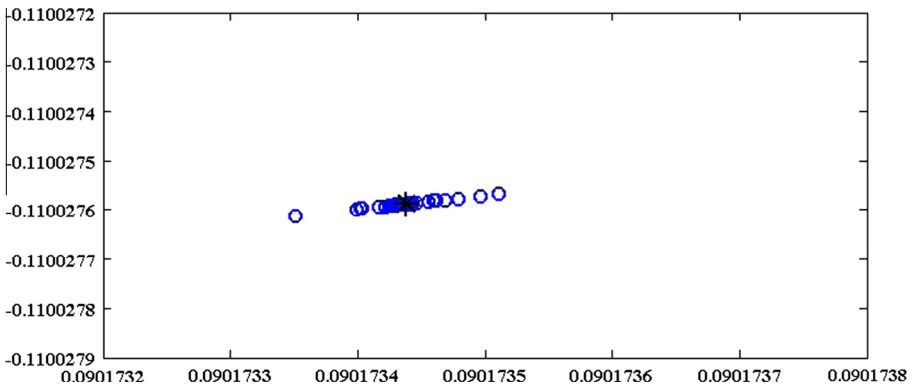

Fig. 25. Positions sampled by the leader particle with $C_1 = C_2 = 1.5$ and $\omega = 0.8$, i.e. $(\omega, C_1, C_2) \in S_4$ on the f_{11} i.e., Shifted Rotated Weierstrass Function.

To observe the dynamic behavior of the leader particle we choose to use the benchmark functions provided in the IEEE Congress on Evolutionary Computation (CEC) 2005 competition and special session on real-parameter optimization [36] with constant C_1, C_2 and with random C_1, C_2 as well.

All the functions are 2D. Initial positions and velocities are randomized and stagnation is detected before starting to plot the positions sampled by the leader particle. The stagnation point in each relevant diagram is marked with an asterisk.

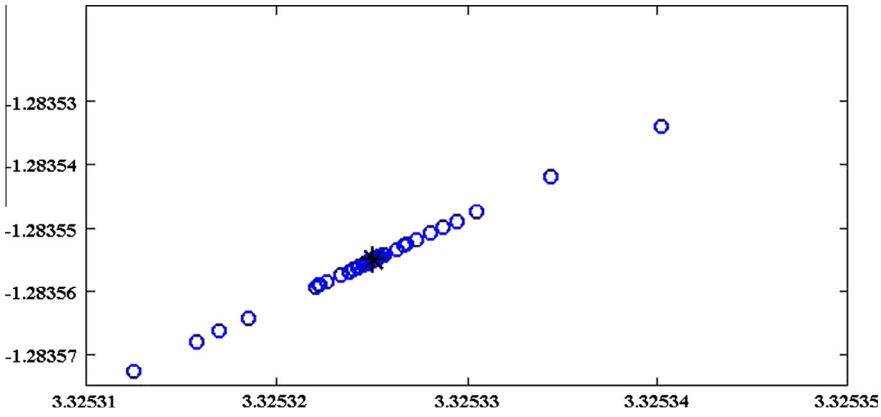

Fig. 26. Positions sampled by the leader particle with $C_1 = C_2 = 1.5$ and $\omega = 0.8$, i.e. $(\omega, C_1, C_2) \in S_4$ on the f_{15} i.e., Hybrid Composition Function.

Figs. 1–7 show the plot of $g(t)$, Figs. 3–14 show the sample histories of the leader particle when C_1, C_2 are constant. We can see from Figs. 1 and 2, that $g(t) \geqslant 0$ always holds for this set of parameter values. Again Figs. 3–15 show that the positions sampled by the leader particle lie on only one side of the stagnation point. Hence the theoretical results proposed in Section 5 are supported by the numerical experiments. The side of the stagnation point on which the sampled points will lie are determined by $x(K+1)$.

7.2. Parameter choices in subregion $S_3 \cup S_4$

Now, the parameters are chosen within the subregion $S_3 \cup S_4$ where, from the mathematical analysis, we know that the leader particle samples the positions on both the sides of the stagnation point.

From Figs. 16 and 17 we can conclude that $g(t) < 0$ for some t and $g(t) > 0$ for other t. which reinforces the results obtained in Section 5. Here the leader particle explores both sides of the stagnation point. which can be seen from the Figs. 17–25. Here also the positions sampled by the leader particle lie on a line determined by $x_i(K)$ and $x_i(K+1)$.

The two sets of parameters are $(C_1 = C_2 = 0.244, \omega = -0.7)$ and $(C_1 = C_2 = 1.5, \omega = 0.8)$. The former combination lies within the subregion S_3 with $\Delta > 0$ and the latter lies in S_4 with $\Delta < 0$. Figs. 16 and 17 show the plot of $g(t)$ vs t. Figs. 18–26 show the sample histories of the leader particle where C_1 and C_2 are constant.

8. Conclusion

The behavioral dynamics of the leader particle in a particle swarm optimization algorithm have been investigated here considering an iterative functional approach.

Taking the characteristic parameters of the PSO, namely ω, C_1 and C_2 as constants, we have proved that the points sampled by the leader particle satisfy a simple relation. The points sampled by the leader particle lie on a straight line identified by $\boldsymbol{x_i}(K)$ and $\boldsymbol{x_i}(K+1)$. It has also been demonstrated that the function $g(t)$ regulates the behavior of the leader particle. The popular convergence region of the PSO algorithm can be divided into two sub-regions designated $S_1 \cup S_2$ and $S_3 \cup S_4$. In the former $g(t) \geqslant 0$ is always true. This immediately points to the fact that the leader particle searches only a single side of the stagnation point. The other side is never explored. Parameter choices made in this region are bad choices for the PSO as it destroys the exploratory nature of the algorithm.

When stagnation remains, i.e. when $M \to \infty$ we also obtain crucial information about the gradient of the objective function. We have successfully shown that when the parameters lie in $S_1 \cup S_2$, then the direction $\boldsymbol{x_i}(K) - \boldsymbol{x_i}(K+1)$ is a descent direction of the objective function at the stagnation point $\boldsymbol{x_i}(K)$. On the other hand, if the parameters lie in $S_3 \cup S_4$, then the direction $\boldsymbol{x_i}(K) - \boldsymbol{x_i}(K+1)$ is orthogonal to the gradient of the objective function at the stagnation point $\boldsymbol{x_i}(K)$. Such a gradient-oriented analysis of the PSO algorithm is a unique area of work that is not found in the current literature on PSO.

The analytical results point to an improvement in the parameter choice so that there is an improvement in the performance of the PSO algorithm. It seeks to obtain this objective by shedding light on the dynamics of the leader particle in stagnation. The leader particle being the one that reaches stagnation point first during stagnation, the study of its dynamics holds importance. Although existing literature has tried to analyze the convergence, particle trajectories and stability of PSO, the leader particle dynamics presented in this paper provides a unique approach to the study. A clear indication of the theoretical results of our paper is towards the parameter choice $(\omega, C_1, C_2) \in S_3 \cup S_4$. This will help to design PSO models that are both effective and efficient. There is a twofold support for this parameter space. On the first place, for $(\omega, C_1, C_2) \in S_3 \cup S_4$, the leader particle is able to search both sides of the stagnation point. In such a situation, there is a greater probability of obtaining better result. For instance, if the optimum happens to be on the side the leader particle never searches, then

the optimization using such a PSO would not yield desirable results. The goal of any optimization algorithm should lay emphasis on sufficient exploration of the search space. Even if the algorithm eventually reaches stagnation, it would be more effective if it can search a greater space before there is no further improvement in position of the particles. On the second hand, the objective function gradient and the points sampled by the leader particle corroborate the parameter choice in $S_3 \cup S_4$. If $\boldsymbol{x_i}(K) - \boldsymbol{x_i}(K+1)$ is orthogonal to the gradient of the objective function, then there is more exploration. Both the sides of the stagnation point are covered as all points sampled in successive iterations lie on the search space. If, however, $\boldsymbol{x_i}(K) - \boldsymbol{x_i}(K+1)$ is along the descent direction of the objective function, or $\boldsymbol{x_i}(K+1) - \boldsymbol{x_i}(K)$ is along the ascent direction of the objective function, successive iterations search only the descent direction, which lies on the search space. There is a loss in explorative nature. The theoretical analysis presented in this paper can lead to similar analysis for other optimization algorithms where the fittest particle in the process of optimization may be analyzed. While that might help to improve the parameter choice, it would also provide useful insight into the subtle changes leading to stagnation.

Further extensions of the work include a statistical analysis of the dynamics of the leader particle considering the parameters ω, C_1 and C_2 to be completely random in nature. Possible improvements to the PSO algorithm can possibly be observed then. Besides it can focus on the damped oscillation that occurs about the stagnation point during stagnation. The cases of under-damped, over-damped and critically damped oscillations depending on parameter choice can be taken up and their implications on PSO performance dealt with.

Appendix A. Evaluation of the limit of $g(t)$ as $t \to \infty$

From Eq. (29),

$$g(t) = \frac{(\lambda_2^t - \lambda_1^t)}{\sqrt{\Delta}}, \quad \Delta > 0$$

Now since

$$\lambda_{1,2} = \frac{a_1}{2} \pm \frac{\sqrt{\Delta}}{2}$$

where $a_1 = 1 + \omega - \phi, \phi = C_1 + C_2$ and $a_2 = -\omega$; and for stable region $0 < (C_1 + C_2) < 2(1+\omega)$, we can conclude $|a_1| < 1 + \omega$. Again $\Delta = a_1^2 + 4a_2 = a_1^2 - 4\omega$. So,

$$|\lambda_{1,2}| = \left|\frac{a_1}{2} \pm \frac{\sqrt{\Delta}}{2}\right| < \frac{1}{2}\left[(1+\omega) \pm \sqrt{(1+\omega)^2 - 4\omega}\right] = \frac{1}{2}[(1+\omega) \pm (1-\omega)]$$

$= 1$ or ω and $|\omega| < 1$. Thus $|\lambda_{1,2}| < 1$ and $g(t) \to 0$ as $t \to \infty$ for $\Delta > 0$.

Again we know form Eq. (29),

$$g(t) = t\left(\frac{a_1}{2}\right)^{t-1}, \quad \Delta = 0.$$

Since $|a_1| < 1 + \omega$ and $-1 < \omega < 1, |a_1| < 2 \Rightarrow \left|\frac{a_1}{2}\right| < 1$,

$$\left(\frac{a_1}{2}\right)^{t-1} \to 0$$

as $t \to 0$. Therefore $g(t) \to 0$ as $t \to 0$.

Lastly from Eq. (29),

$$g(t) = \frac{\sin(t\theta)}{\sin\theta}\left(\sqrt{\omega}\right)^{t-1}, \quad \Delta < 0.$$

Now $|\sin(t\theta)| < 1 \forall t$ and $\sin\theta \neq 0$; therefore since $|\omega| < 1$,

$$g(t) \to 0$$

as $t \to 0$.

So we can conclude

$$\lim_{t\to\infty} g(t) = 0,$$

References

[1] F. van den Bergh, An Analysis of Particle Swarm Optimizers (Ph.D. thesis). University of Pretoria, Faculty of Natural and Agricultural Science.
[2] F. van den Bergh, A.P. Engelbrecht, A study of particle swarm optimization particle trajectories, Information Science 176 (8) (2006) 937–971.
[3] T.M. Blackwell, Particle swarms and population diversity, Soft Comput. 9 (2005) 793–802.
[4] D. Bratton, T. Blackwell, A simplified recombinant PSO, J. Artif. Evol. Appl. 2008 (2008).

[5] E.F. Campana, G. Fasano, A. Pinto, Dynamic system analysis and initial particles position in particle swarm optimization, in: Proceedings of the IEEE Swarm Intelligence Symposium (SIS), IEEE, Indianapolis, Piscataway, 2006.
[6] E.F. Campana, G. Fasano, D. Peri, A. Pinto, Particle swarm optimization: efficient globally convergent modifications, in: C.A. Mota Soares, et al. (Eds.), Proceedings of the III European Conference on Computational Mechanics, Solids, Structures and Coupled Problems in Engineering, Lisbon, Portugal, 2006.
[7] A. Carlisle, G. Dozier, An off-the-shelf PSO, in: Proceedings of the Particle Swarm Optimization Workshop, pp. 1–6.
[8] M. Clerc, J. Kennedy, The particle swarm – explosion, stability and convergence in a multidimensional complex space, IEEE Trans. Evol. Comput. 6 (1) (2002) 58–73.
[9] M. Clerc, Stagnation Analysis in Particle Swarm Optimization or What Happens When Nothing Happens, Technical Report CSM-460, Department of Computer Science, University of Essex, August 2006.
[10] R.C. Eberhart, J. Kennedy, A new optimizer using particle swarm theory, in: Proceedings of the Sixth International Symposium on Micro Machine and Human Science, IEEE, Nagoya, Japan, Piscataway, 1995, pp. 39–43.
[11] R.C. Eberhart, Y. Shi, Comparing inertia weights and constriction factors in particle swarm optimization, in: Proceedings of the Congress on Evolutionary Computation, vol. 1, pp. 84–88.
[12] G. Evers, An Automatic Regrouping Mechanism to Deal with Stagnation in Particle Swarm Optimization (Master's thesis), The University of Texas – Pan American, Department of Electrical Engineering.
[13] J.L. Fernández Martínez, E. García Gonzalo, J.P. Fernández Alvarez, Theoretical analysis of particle swarm trajectories through a mechanical analogy, Int. J. Comput. Intell. Res. 4 (2) (2008).
[14] J. García-Nieto, E. Alba, Empirical computation of the quasi-optimal number of informants in particle swarm optimization, in: ACM Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'11), Dublin, Ireland, July 2011, pp. 147–154.
[15] J. García-Nieto, E. Alba, Why six informants is optimal in PSO, in: ACM Proceedings of the Genetic and Evolutionary Computation Conference (GECCO'12), Philadelphia, USA, July 2012, pp. 25–32.
[16] J. Kennedy, The behavior of particles, in: V.W. Porto, N. Saravanan, D. Waagen, A.E. Eiben (Eds.), Evolutionary Programming VII: Proceedings of the 7th Annual Conference on Evolutionary Programming Conference, Springer-Verlag, San Diego, CA, Berlin, 1998, pp. 581–589.
[17] J. Kennedy, Bare bones particle swarms, in: Proceedings of the IEEE Swarm Intelligence Symposium, Indianapolis, IN, 2003, pp. 80–87.
[18] J. Kennedy, R.C. Eberhart, Particle swarm optimization, in: proceeding of the IEEE International Conference on Neural Networks IV, IEEE, Piscataway, 1995, pp. 1942–1948.
[19] J. Kennedy, R.C. Eberhart, A discrete binary version of the particle swarm algorithm, in: Proceedings of the Conference on Systems, Man, and Cybernetics, IEEE, Piscataway, 1997, pp. 4104–4109.
[20] J. Kennedy, R. Mendes, Population structure and particle swarm performance, in: Proceedings of the IEEE Congress on Evolutionary Computation (CEC), IEEE, Honolulu, HI, Piscataway, 2002, pp. 1671–1676.
[21] V. Kadirkamanathan, K. Selvarajah, P.J. Fleming, Stability analysis of the particle dynamics in particle swarm optimizer, IEEE Trans. Evolut. Comput. 10 (3) (2006) 245–255.
[22] R. Mendes, Population Topologies and Their Influence in Particle Swarm Performance, Ph.D. thesis, Departamento de Informatica, Escola de Engenharia, Universidade do Minho, 2004.
[23] R. Mendes, P. Cortes, M. Rocha, J. Neves, Particle swarms for feedforward neural net training, in: Proceedings of the International Joint Conference on Neural Networks, IEEE, Honolulu, HI, Piscataway, 2002, pp. 1895–1899.
[24] R. Mendes, J. Kennedy, J. Neves, Watch thy neighbor or how the swarm can learn from its environment, in: Proceedings of the IEEE Swarm Intelligence Symposium (SIS), IEEE, Piscataway, 2003, pp. 88–94.
[25] E. Ozcan, C.K. Mohan, Analysis of a simple particle swarm optimization system, Intell. Eng. Syst. Thro. Artif. Neural Netw. 8 (1998) 253–258.
[26] E. Ozcan, C.K. Mohan, Particle swarm optimization: surfing the waves, in: Proceeding of Congress on Evolutionary Computation, IEEE Service Center, Piscataway, NJ, 1999, pp. 1939–1944.
[27] R. Poli, On the moments of the sampling distribution of particle swarm optimisers, in: Proceedings of the Workshop on Particle Swarm Optimization: The Second Decade of the Genetic and Evolutionary Computation Conference (GECCO), ACM, London, New York, 2007.
[28] R. Poli, D. Broomhead, Exact analysis of the sampling distribution for the canonical particle swarm optimiser and its convergence during stagnation, in: Genetic and Evolutionary Computation Conference (GECCO), ACM, London, New York, 2007.
[29] R. Poli, J. Kennedy, T. Blackwell, Particle swarm optimization: an overview, Swarm Intell. 1 (2007) 33–57.
[30] R. Poli, Analysis of the publications on the applications of particle swarm optimisation, J. Artif. Evolut. Appl. 2008 (2008) 1–10.
[31] P. Rocca, M. Benedetti, M. Donelli, D. Franceschini, A. Massa, Evolutionary optimization as applied to inverse scattering problems, Inverse Prob. 25 (2009) 1–41.
[32] Y. Shi, R.C. Eberhart, A modified particle swarm optimizer, in: Proceedings of the IEEE International Conference on Evolutionary Computation, IEEE, Piscataway, 1998, pp. 69–73.
[33] Y. Shi, R.C. Eberhart, Parameter selection in particle swarm optimization, in: Proceedings of Evolutionary Programming VII (EP98), pp. 591–600.
[34] I.C. Trelea, The particle swarm optimization algorithm: convergence analysis and parameter selection, Inform. Process. Lett. 85 (6) (2003) 317–325.
[35] K. Yasuda, A. Ide, N. Iwasaki, Adaptive particle swarm optimization, in: Proceedings of the IEEE International Conference on Systems Man, and Cybernetics, IEEE, Piscataway, 2003, pp. 155–159.
[36] P.N. Suganthan, N. Hansen, J.J. Liang, K. Deb, Y.-P. Chen, A. Auger, S. Tiwari, 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.
[37] S. Ghosh, S. Das, D. Kundu, K. Suresh, A. Abraham, Inter-particle communication and search-dynamics of lbest particle swarm optimizers: an analysis, Inform. Sci. 182 (1) (2012) 156–168. ISSN 0020-0255.