Stability Analysis of the Particle Swarm Optimization Without Stagnation Assumption

Mohammad Reza Bonyadi and Zbigniew Michalewicz

***Abstract*—In this letter, we study the first- and second-order stabilities of a stochastic recurrence relation that represents a class of particle swarm optimization (PSO) algorithms. We assume that the personal and global best vectors in that relation are random variables (with arbitrary means and variances) that are updated during the run so that our calculations do not require the stagnation assumption. We prove that the convergence of expectation and variance of the positions generated by that relation is independent of the mean and variance of the distribution of the personal and global best vectors. We also provide convergence boundaries for that relation and compare them with those of standard PSO algorithms (as a specific case of the stochastic recurrence relation) provided in earlier studies.**

***Index Terms*—Convergence, particle swarm optimization (PSO), stability analysis, stochastic recurrence.**

I. Introduction

Particle swarm optimization (PSO) is a stochastic population-based optimization algorithm developed by Kennedy and Eberhart [1]. PSO has been applied to many optimization problems such as artificial neural network training and pattern classification [2], [3], to name a few. Since 1995, different theoretical aspects of the original version of PSO have been investigated (local convergence [4], [5], rotation invariance [6], [7], stability [8]–[12], etc.). One of the important aspects related to PSO is to ensure that the generated solutions by the algorithm converge to a point. Convergence to a point for a PSO algorithm is usually analyzed to determine the coefficients' boundaries, for which the generated solutions by the algorithm do not diverge. These boundaries are known as convergence boundaries. For such analyses, it is usually considered that the "memories" of particles (known as personal best and global best) are not updated during the run, referred to as the stagnation assumption [8], [11]. Because such an assumption is not realistic and only covers one state that particles might experience in their lifetime, recently some researchers have investigated the convergence of particles [10], [13] under the weaker stagnation assumption to determine the boundaries under more realistic conditions for the PSO variant introduced in [14].

In this letter, we introduce a stochastic recurrence relation that represents a wide range of PSO variants. Then, we investigate the convergence of expectation and variance of the generated points by this relation to determine its convergence boundaries. We assume that the memories of particles are stochastic variables (updated during the run) and derive the convergence boundaries under such an assumption. This assumption is indeed more generic and realistic than the ones made in previous works that allows us to find convergence boundaries to guarantee convergence, taking into account the update rules for memories.

Without a loss of generality, this letter only considers minimization problems defined as

$$\text{find } \vec{x} \in S \subseteq \mathbb{R}^d \text{ such that } \forall \vec{y} \in S, F(\vec{x}) \leq F(\vec{y}) \tag{1}$$

where S is the search space defined by $\{x | l_i \leq x_i \leq u_i \text{ for all } i\}$, l_i and u_i are the lower and upper bounds of the values of the ith dimension of S, d is the number of dimensions, and $F(.)$ is the objective function.

The rest of this letter is organized as follows. We provide some background information on the original version of PSO in Section II and convergence analysis in Section III. We present our proposed convergence analyses in Section IV and we conclude the letter in Section V.

II. Particle Swarm Optimization

Each particle in the original PSO (OPSO) [1], [15] contains three vectors, as follows.

1) *Position* ($\vec{x}_t^i$)*:* This is the position of the ith particle in the tth iteration that is used to evaluate the particle quality.
2) *Velocity* ($\vec{v}_t^i$)*:* This is the direction and length of movement of the ith particle in the tth iteration.
3) *Personal Best* ($\vec{p}_t^i$)*:* This is the best position (in terms of objective value) that the particle i has visited until iteration t (we refer to this vector as "memory" of particles).

The personal best is updated as: $\vec{p}_{t+1}^i = \vec{x}_t^i$ if $F(\vec{x}_t^i) < F(\vec{p}_t^i) + \epsilon$, otherwise $\vec{p}_{t+1}^i = \vec{p}_t^i$, where $\epsilon > 0$ is an arbitrarily small value that represents the precision of the calculations (see [12] for details).

The value of the jth dimension of each of these vectors is represented by the superscript j throughout the letter. In OPSO, the jth dimension of the velocity of each particle is updated for the next iteration $(t+1)$ by

$$v_{t+1}^{i,j} = v_t^{i,j} + c_1\left(p_t^{i,j} - x_t^{i,j}\right) + c_2\left(g_t^j - x_t^{i,j}\right) \tag{2}$$

where c_1 and c_2 are two random variables uniformly distributed in $[0, \phi_1]$ and $[0, \phi_2]$, respectively; ϕ_1 and ϕ_2 are two real numbers called acceleration coefficients; $p_t^{i,j}$ is the jth dimension of the personal best of the particle i at iteration t; and g_t^j is the jth dimension of the best personal best among all particles in the swarm, known as the global best of the swarm. In the case of local best topologies, g_t^j is selected among the particles in the neighbor (depending on how the topology is defined) of the particle i. The position of the particles is updated by

$$x_{t+1}^{i,j} = x_t^{i,j} + v_{t+1}^{i,j}. \tag{3}$$

Manuscript received June 15, 2015; revised September 22, 2015, November 26, 2015, and December 4, 2015; accepted December 10, 2015. Date of publication December 11, 2015; date of current version September 30, 2016. This work was supported by the Polish Ministry of Science and Higher Education under Grant N N519 5788038.

M. R. Bonyadi is with the Optimisation and Logistics Group, University of Adelaide, Adelaide, SA 5005, Australia, and also with the Data Science Group of Complexica Pty. Ltd., Adelaide, SA 5021, Australia (e-mail: mrbonyadi@cs.adelaide.edu.au).

Z. Michalewicz is with the Optimisation and Logistics Group, The University of Adelaide, Adelaide, SA 5005, Australia, also with the Data Science Group of Complexica Pty. Ltd., Adelaide, SA, Australia; with the Institute of Computer Science, Polish Academy of Sciences, Warsaw 01-237, Poland; and also with the Polish–Japanese Institute of Information Technology, Warsaw 02-008, Poland (e-mail: zbyszek@cs.adelaide.edu.au).

Digital Object Identifier 10.1109/TEVC.2015.2508101

OPSO has been studied by many researchers and many variants have been developed. One of these variants [14], called the inertia PSO (IPSO), proposed to multiply the previous velocity ($v_t^{i,j}$) by an inertia weight (ω) to control the impact of $v_t^{i,j}$ on the movement of particles. The velocity update rule for IPSO was written as

$$v_{t+1}^{i,j} = \omega v_t^{i,j} + c_1\left(p_t^{i,j} - x_t^{i,j}\right) + c_2\left(g_t^{j} - x_t^{i,j}\right) \tag{4}$$

where ω is a constant value called inertia weight. IPSO is usually used by researchers for theoretical analysis. Throughout this letter, $c_1(p_t^{i,j} - x_t^{i,j})$ is called the cognitive component and $c_2(g_t^{j} - x_t^{i,j})$ is called the social component of velocity.

III. Background

In this section, we outline some theorems and studies related to the convergence and stability analysis of PSO.

A. Convergence and Stability of Linear Recurrence Relations

A linear recurrence relation is written as

$$\vec{z}_{t+1} = M\vec{z}_t + \vec{b} \tag{5}$$

where $\vec{z}_t$, $\vec{z}_{t+1}$, and $\vec{b}$ are d dimensional vectors and M is a $d \times d$ matrix. In the case of linear recurrence relations, the following lemma is useful for convergence analysis.

Lemma 1: The sequence $\{\vec{z}_1, \vec{z}_2, \ldots\}$, generated by the recurrence in (5), is convergent if and only if $\rho(M) = \max_{i\in\{1\ldots d\}}(|\gamma_i|) < 1$, where γ_i is the ith eigenvalue of M. $\rho(M)$ is known as the spectral radius of M.

Proof: See [16] for proof. ∎

We define the fixed point of a linear recurrence as follows.

Definition 1: $\hat{z}$, the fixed point of the linear recurrence in (5) is defined by $\lim_{t\to\infty} \vec{z}_t = \hat{z}$.

Lemma 1 introduces a necessary and sufficient condition for convergence. It is, however, not always trivial to calculate and simplify the eigenvalues of the matrix M. The following lemma introduces a necessary condition for convergence of a linear recurrence relation.

Lemma 2: If the sequence $\{\vec{z}_1, \vec{z}_2, \ldots\}$ generated by the recurrence in (5) is convergent, then there exists $\hat{z} \in \mathbb{R}^d$ as a fixed point for the sequence.

Proof: See [16] for proof. ∎

This lemma indicates that the existence[1] of the fixed point $\hat{z}$ is a necessary condition for the convergence of the sequence $\vec{z}_t$. Therefore, if $\hat{z}$ does not exist (it is infinite) for a given M and $\vec{b}$ then $\rho(M) \geq 1$. To calculate this fixed point we can write $\hat{z} = M\hat{z} + \vec{b}$, which results in $\hat{z} = (I - M)^{-1}\vec{b}$, where I is the identity matrix. If M and $\vec{b}$ are given then $\hat{z}$ can be calculated.

The first-order stability for a 1-D stochastic sequence $\{x_1, x_2, \ldots\}$ (x_t is a random variable for all t) is defined as follows.

Definition 2: A sequence $\{x_1, x_2, \ldots\}$ is called "first-order stable" if and only if the sequence $\{E(x_1), E(x_2), \ldots\}$, where $E(.)$ is the expectation operator, is convergent.

The second-order stability for this sequence is defined as follows.

Definition 3: A sequence $\{x_1, x_2, \ldots\}$ is called "second-order stable" if and only if the sequence $\{V(x_1), V(x_2), \ldots\}$, where $V(.)$ is the variance operator, is convergent and V_x, the fixed point of the variance, is zero.

Sometimes a stochastic sequence $\{x_1, x_2, \ldots\}$ is analyzed through a deterministic model, i.e., its stochastic components are replaced by constants. Then, the convergence and stability of the resulted deterministic sequence is investigated. Such analysis is called the "deterministic model stability analysis" throughout this letter.

B. Stability of PSO With Stagnation Assumption

Convergence and stability of IPSO have been investigated through the deterministic model stability analysis [9], [17], first-order stability analysis [10], [11], [18], and second-order stability analysis [8], [13], [19].

One of the earliest attempts to analyze the convergence behavior of IPSO to find convergence boundaries was performed in [9]. In that paper, in order to simplify the formulation of update rules, stochastic components from the system were omitted (deterministic model analysis) and it was assumed that p and g are not updated during the run (stagnation assumption). This analysis led to a relation between acceleration coefficients and inertia weight to guarantee stability of particles.

The first-order stability of particles was analyzed in [11] and [18]. In these studies, the stochastic components (c_1 and c_2) were replaced by their expected values, $\phi_1/2$ and $\phi_2/2$. It was found in [18] that the expected value of the position of particles is convergent if $\omega < 1$, $\phi > 0$, and $4\omega - \phi + 4 > 0$, where $\phi = \phi_1 + \phi_2$. This condition guarantees that the expected value of the positions converge at least at 50% of iterations. However, if convergence of expected position of particles at every iteration is required, then $\omega < 1$, $\phi > 0$, and $2\omega - \phi + 2 > 0$, where $\phi = \phi_1 + \phi_2$ [11]. Also, the positions of particles converge to $((\phi_2 g + \phi_1 p)/(\phi_1 + \phi_2))$.

Particles positions can still move to infinity even though their expected position is convergent. Hence, some researchers [8], [19] studied the convergence of variance[2] of positions of particles during the run. These studies proved that if the variance of positions of a particle converge to fixed point, then $\phi < ((12(\omega^2 - 1))/(5\omega - 7))$, where $\phi = \phi_1 = \phi_2$. Also, the fixed point for the variance of particles is $((\phi(\omega + 1)(g - p)^2)/(4(\phi(5\omega - 7) - 12\omega^2 + 12)))$. This condition guarantees convergence of variance and not necessarily the second-order stability. The particles are second-order stable if this fixed point is zero, which occurs only if $g = p$ [8].

The first- and second-order stabilities were investigated in [20] for IPSO when a generic distribution for the inertia weight and acceleration coefficients was considered, i.e., inertia weight and acceleration coefficients are random variables from an arbitrary probability distribution with expected values μ_ω and μ_c and variances σ_ω^2 and σ_c^2, where $c = c_1 + c_2$. It was proven that IPSO is first-order stable if and only if $-1 < \mu_\omega < 1$ and $0 < \mu_c < 2(\mu_\omega + 1)$. Also, it was found that the upper limit of the values of coefficients to guarantee convergence of variance of the algorithm ($\rho(M) = 1$) is defined by $-a < \mu_\omega < a$ and $0 < \mu_c < b$, where $a = (1/\sqrt{q^2 + 1})$, $b = ((2(1 - (1 + q)\mu_\omega^2))/(1 + d^2 + (d^2 - 1)\mu_\omega))$, $q = \sigma_\omega/\mu_\omega$, and $d = \sigma_c/\mu_c$. This is a necessary condition for convergence of variance. Also, the assumption for the analysis was that the personal and global bests remain constant during the run (stagnation).

C. Weakening the Stagnation Assumption

Recently, some studies have tried to weaken the stagnation assumption to find the convergence regions under more general conditions. For example, it was assumed [10] that the personal best of particles and the global best of the swarm are allowed to move and can occupy an arbitrarily large finite number of unique positions during the run. The main finding of that study was that IPSO is first-order stable if $-1 < \omega < 1$ and $0 < \theta < 2(1 + \omega)$, where $\theta = \theta_1 + \theta_2 < 4$

[1] Note that the existence of this limit entails $\hat{z}$ is finite.

[2] Note that [8] studied the second moment and standard deviation of positions (rather than the variance itself), which are closely related to variance.

and $\theta_i \in [0, \phi_i]$, under this more general assumption. This is indeed the same as what was found by van den Bergh and Engelbrecht [11] when $\theta_i = \phi_i$ and the same as what was found by Trelea [18] when $\theta_i = \phi_i/2$. Hence, one can conclude that the stagnation assumption does not affect the convergence boundaries for first-order stability in IPSO.

The second-order stability of IPSO under a weaker stagnation assumption was investigated by [13]. The assumption in that study was that the personal best is constant from an iteration k to $k + k_0$, where $k_0 \geq 3$. It was proven that, for the global best particle (the particle that its personal best is equal to the global best, $p_t = g_t$ for all t), the convergence boundaries found by Jiang *et al.* [19] as well as by Poli [8] are valid under this assumption, i.e., the variance converges. Also, it was proven that the global best particle in IPSO is second-order stable if and only if its variance is convergent.

IV. Proposed Stability Analysis

The present study extends the findings of previous works from different perspectives. We investigate a stochastic recurrence relation that represents a wide range of PSO variants. We analyze the convergence of expectation and variance for this recurrence relation and find necessary and sufficient conditions to guarantee these convergence properties. We also investigate the first- and second-order stability and find conditions to guarantee these types of stability for that recurrence relation. In all of these calculations, we consider p and g as random variables that are updated during the run with an arbitrary mean and variance. This in fact weakens the stagnation assumption considered by earlier studies.

A. Generalization of PSO Formulation

We assume that the positions of particles are updated for each dimension independently. We also investigate an arbitrary particle in the swarm. In the 1-D space, the position of a particle can be considered as a random variable. In this letter, we study the position update rule of IPSO with generic distribution for its coefficients

$$x_{t+1} = lx_t - \omega x_{t-1} + c_1 p + c_2 g \tag{6}$$

where $l = (1+\omega-c_1-c_2)$, c_1, c_2, p, g, and ω are random variables with expected values (μ) and standard deviations (σ). This formulation allows movement of p and g through a distribution with a mean and variance.

If an arbitrary precision ϵ for the objective value is considered (that is the case for simulations, see Section II), the set of all possible positions for p and g is arbitrarily large but finite (this was also assumed by [10], see Section III-C). Hence, as the search space is bounded, the variance of p and g is finite for any ϵ.

In a simulation, it is usually the case that the locations of p_t (g_t) for each particle show chaotic behavior all over the search space when t is small (exploration). In an extreme case for such behavior, p_t (g_t) samples the search space through a uniform random distribution. However, as t grows, p_t (g_t) becomes more concentrated on a smaller area; hence, the density of the samples picked by p_t (g_t) becomes larger in that specific area. This means that the mean of the distribution of the generated p_t (g_t) converges to a point in that specific area and the variance of the generated p_t (g_t) becomes closer to zero.[3] This can be observed in the convergence curves reported in [6] and [21] where the location of p (g) is updated less frequently when the iteration number grows. Therefore, the distribution of p (g) becomes closer to a gamma distribution with a shrinking variance as t grows.

One should note that, as p and g move through a distribution with a mean and variance, (6) is independent of the topology as it does not matter which other particle updates g. Nevertheless, this formulation does not model fully informed topologies introduced in [22].

B. First Order Stability Analysis

In order to guarantee first-order stability, we need to guarantee convergence of the expectation of the positions (see Definition 2). We calculate the expected position of a particle by applying the expectation operator to both sides of (6)

$$E(x_{t+1}) = E(l)E(x_t) - \mu_\omega E(x_{t-1}) + \mu_{c_1}\mu_p + \mu_{c_2}\mu_g \tag{7}$$

where μ_ω, μ_{c_1}, μ_{c_2}, μ_p, and μ_g, are the expected values of ω, c_1, c_2, p, and g, respectively, and $E(l) = 1 + \mu_\omega - \mu_{c_1} - \mu_{c_2}$. Alternatively, one can rewrite this equation in a matrix form introduced in (5), where $\vec{z}_t = [E(x_t)\ E(x_{t-1})]^T$, $M = \begin{bmatrix} E(l) & -\mu_\omega \\ 1 & 0 \end{bmatrix}$, and $\vec{b} = [\mu_{c_1}\mu_p + \mu_{c_2}\mu_g\ \ 0]^T$. The eigen values of M are $((E(l) \pm \sqrt{E(l)^2 - 4\mu_\omega})/2)$. According to Lemma 1, in order to guarantee convergence of expectation we need to ensure that the magnitude of these eigenvalues are smaller than 1. After simplifications, we found that the expectation of x_t is convergent (spectral radius is smaller than 1) if and only if

$$-1 < \mu_\omega < 1 \text{ and } 0 < \mu_{c_1} + \mu_{c_2} < 2(\mu_\omega + 1). \tag{8}$$

This is called the expectation convergence boundaries. It is clear that the convergence of expectation is independent of the mean and variance of p and g. Therefore, the first-order stability is guaranteed if and only if the conditions in (8) are satisfied, no matter how p and g are updated. Also, this equation suggests that the first-order stability of a PSO variant that follows the recurrence in (6) with any distribution for c_1, c_2, and ω is only dependent on the expectation of these variables.

Let us assume that the expected value of the position of particles is convergent and it converges to a value E_x. If the fixed point of the recurrence $\vec{z}_t = [E(x_t)\ E(x_{t-1})]^T$ is $\hat{z}$ then $E_x = \hat{z}_1$, where $\hat{z}_1$ the first element of $\hat{z}$. After simplifications, E_x is calculated as

$$E_x = \frac{\mu_{c_1}\mu_p + \mu_{c_2}\mu_g}{\mu_{c_1} + \mu_{c_2}}. \tag{9}$$

If particles are first-order stable, then their expected position converges to this point after a long run.

For example, in IPSO, because $\mu_{c_1} = (c_1/2)$, $\mu_{c_2} = (c_2/2)$, and $\mu_\omega = \omega$, the relation to ensure the convergence of expectation of the algorithm is written as

$$-1 < \omega < 1 \text{ and } 0 < c_1 + c_2 < 4(\omega + 1). \tag{10}$$

This equation is aligned with what was found based on the stagnation assumption [9], [11], [18] and the weakened stagnation assumption [10]. Hence, it seems that the stagnation assumption does not affect the convergence boundaries to guarantee first-order stability.

[3]Note that p_t (g_t) cannot be always updated as for some iterations, such as t, when there is no other point in the search space that is better than p_t (g_t) by at least ϵ.

Also, the fixed point of the expectation for IPSO is calculated as $((\phi_1\mu_p + \phi_2\mu_g)/(\phi_1 + \phi_2))$, which is also aligned with previous studies. Note the stagnation assumption results in $\mu_p = p$ and $\mu_g = g$.

C. Second-Order Stability Analysis

In order to guarantee second-order stability of the position of particles, we need to guarantee the convergence of variance and that the fixed point of the variance converges to zero (see Definition 3). We first formulate the variance of positions in the form of (5). Then we study the convergence of variance and the boundaries to guarantee convergence of variance (called the variance convergence boundaries), and, finally, we investigate conditions to guarantee the second-order stability.

1) Matrix Form of Variance: The variance of a random variable α is calculated by

$$V(\alpha) = E\left(\alpha^2\right) - E(\alpha)^2. \tag{11}$$

Thus, in order to calculate $V(x_t)$ we need to calculate $E(x_t^2)$. Let us calculate x_{t+1}^2 as

$$x_{t+1}^2 = \left(lx_t - \omega x_{t-1} + c_1 p + c_2 g\right)^2 = l^2 x_t^2 + \omega^2 x_{t-1}^2 - 2l\omega x_t x_{t-1} + 2lPx_t - 2\omega P x_{t-1} + P^2 \tag{12}$$

where $l = 1 + \omega - c_1 - c_2$ and $P = c_1 p + c_2 g$. Hence, the expected value for x_{t+1}^2 is given by

$$\begin{aligned} E\left(x_{t+1}^2\right) = {} & E\left(l^2\right)E\left(x_t^2\right) + E\left(\omega^2\right)Ex_{t-1}^2 \\ & - 2E(l\omega)E\left(x_t x_{t-1}\right) + 2E(lP)E(x_t) \\ & - 2E(\omega P)E(x_{t-1}) + E\left(P^2\right). \end{aligned} \tag{13}$$

We calculate $E(l^2)$, $E(\omega^2)$, $E(l\omega)$, $E(lP)$, $E(\omega P)$, $E(P^2)$, and $E(x_t x_{t-1})$ in turn. The value for $E(l^2)$ is calculated as

$$\begin{aligned} E\left(l^2\right) = {} & 1 + E\left(\omega^2\right) + E\left(c_1^2\right) + E\left(c_2^2\right) + 2\mu_\omega - \\ & - 2\left(\mu_{c_1} + \mu_{c_2}\right) - 2\mu_\omega(\mu_{c_1} + \mu_{c_2}) + 2\mu_{c_1}\mu_{c_2} \end{aligned} \tag{14}$$

where, according to (11), $E(\omega^2) = \sigma_\omega^2 + \mu_\omega^2$, $E(c_1^2) = \sigma_{c_1}^2 + \mu_{c_1}^2$, and $E(c_2^2) = \sigma_{c_2}^2 + \mu_{c_2}^2$, where σ_ω^2, $\sigma_{c_1}^2$, and $\sigma_{c_2}^2$ are the variance of the random variables ω, c_1, and c_2, respectively.

It is clear that $E(l\omega) = E(\omega + \omega^2 - \omega c_1 - \omega c_2)$ results in $E(l\omega) = \mu_\omega + E(\omega^2) - \mu_\omega(\mu_{c_1} - \mu_{c_2})$. To calculate $E(lP)$ we expand lP and we get

$$\begin{aligned} E(lP) = {} & E\Big(c_1 p + c_2 g + \omega c_1 p + \omega c_2 g - {c_1}^2 p - c_1 c_2 g \\ & - c_2 c_1 p - {c_2}^2 g\Big) = \mu_{c_1}\mu_p + \mu_{c_2}\mu_g + \mu_w \mu_{c_1}\mu_p \\ & + \mu_\omega \mu_{c_2}\mu_g - \mu_p E\left(c_1^2\right) - \mu_{c_1}\mu_{c_2}\left(\mu_p + \mu_g\right) - \mu_g E\left(c_2^2\right). \end{aligned} \tag{15}$$

Because ω and P are two independent random variables, $E(\omega P) = \mu_\omega(\mu_{c_1}\mu_p + \mu_{c_2}\mu_g)$. $E(P^2)$ is calculated as

$$\begin{aligned} E\left(P^2\right) = {} & E\left(c_1^2 p^2 + c_2^2 g^2 + 2c_1 c_2 pg\right) = E\left(p^2\right)E\left(c_1^2\right) \\ & + E\left(g^2\right)E\left(c_2^2\right) + 2\mu_{c_1}\mu_{c_2}\mu_p\mu_g \end{aligned} \tag{16}$$

where $E(p^2) = \sigma_p^2 + \mu_p^2$ and $E(g^2) = \sigma_g^2 + \mu_g^2$—in these, σ_p^2 and σ_g^2 are the variance of movement of p and g. Finally, we calculate $E(x_t x_{t-1})$ by multiplying x_{t+1} by x_t and applying the expectation operator

$$E\left(x_{t+1}x_t\right) = E(l)E\left(x_t^2\right) - E(\omega)E(x_t x_{t-1}) + E(P)E(x_t). \tag{17}$$

Now, we can define a recurrence relation in the form of (5), where

$$M = \begin{bmatrix} E(l) & -E(\omega) & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 2E(lP) & -2E(wP) & E(l^2) & E(\omega^2) & -2E(l\omega) \\ 0 & 0 & 1 & 0 & 0 \\ E(P) & 0 & E(l) & 0 & -E(\omega) \end{bmatrix}$$

$\vec{z}_t = [E(x_t)\ E(x_{t-1})\ E(x_t^2)\ E(x_{t-1}^2)\ E(x_t x_{t-1})]^T$, and $\vec{b} = [E(P)\ 0\ E(P^2)\ 0\ 0]^T$. The matrix M and the vector $\vec{b}$ are generalizations of the matrix form proposed in [8]. Note that this matrix can be used to represent the variance of the positions as $V(x_t) = E(x_t^2) - E(x_t)^2$.

2) Convergence of Variance: We calculated $\gamma_{i=1\ldots5}$ (eigenvalues of the matrix M)[4] and we found that all eigenvalues of M are independent of μ_p, μ_g, σ_p, and σ_g. This means that whether the sequence of $\vec{z}_t$ is convergent is independent of the mean and variance of p and g. As the variance of the sequence is calculated by $E(x_t^2) - E(x_t)^2$, the convergence of variance is also independent of the mean and variance of p and g. This was observed in [12] through an experimental approach for IPSO.

This finding looks counterintuitive. In fact, one may expect that any changes in the values of p and g in (6) cause a change in the stability of the variance of the generated sequence. However, as our theoretical study showed, the changes of p and g through a distribution with some mean and variance does not affect the convergence of variance of the positions generated by (6). Accordingly, findings reported in earlier literature related to the convergence of variance under stagnation assumption can be generalized to the cases where p and g are allowed to move through some arbitrary distribution.

3) Variance Convergence Boundaries: In order to find convergence boundaries to guarantee convergence of variance, we need to find conditions to guarantee $\rho(M) < 1$ (see Lemma 1). Unfortunately, simplification of γ_is is very difficult as they are very complex.[5] Hence, we rather investigate the fixed point of this matrix form to find necessary conditions only for convergent of variance (see Lemma 2). After that, we further investigate this necessary condition to see if it is also sufficient for convergence of variance.

According to Lemma 2, if the relation $\vec{z}_t$ is going to be convergent then the values in M and $\vec{b}$ must ensure the existence of the fixed point $\hat{z}$. One can calculate $\hat{z}$ trivially (see Section IV-C1) and use it to calculate V_x (the fixed point of the variance of positions) as $V_x = \hat{z}_3 - \hat{z}_1^2$ [see (11)], where $\hat{z}_i$ is the ith element of $\hat{z}$. Therefore, we calculate the fixed point of variance as

$$V_x = -\frac{k_3 + k_4}{k_1 k_2} \tag{18}$$

where

$$\begin{aligned} k_1 &= (\mu_{c_1} + \mu_{c_2})^2; \\ k_2 &= k_1(1-\mu_\omega) + 2(\mu_{c_1}+\mu_{c_2})(\mu_\omega^2+\sigma_\omega^2-1) + (\sigma_{c_1}^2+\sigma_{c_2}^2)(\mu_\omega+1); \\ k_3 &= k_1(\mu_\omega+1)(\mu_{c_1}^2\sigma_p^2 + \mu_{c_2}^2\sigma_g^2 + \sigma_{c_1}^2\sigma_p^2 + \sigma_{c_2}^2\sigma_g^2); \\ k_4 &= (\mu_{c_1}^2\sigma_{c_2}^2 + \mu_{c_2}^2\sigma_{c_1}^2)(\mu_\omega+1)(\mu_g - \mu_p)^2. \end{aligned}$$

[4]This calculation involves many simplifications and algebraic operations that have not been included to this letter for the sake of simplicity. We used the Symbolic Math toolbox from MATLAB 2013 together with manual procedures to perform these calculations. Please see the supplementary file for this publication for the MATLAB program.

[5]See the MATLAB function provided as a supplementary file for details of these eigenvalues.

Based on Lemma 2, a necessary condition for the convergence of any linear recurrence relation is the existence of its fixed point. Also, by definition, the variance of any random variable is positive, i.e., $V_x \geq 0$. Moreover, based on the definition of variance (11), expectation of positions needs to be convergent; otherwise, the variance of the position would not be convergent. Hence, in order to guarantee the existence of V_x, we need to guarantee the following.

Criterion 1: Expectation that positions are convergent.

Criterion 2: $V_x \geq 0$ (variance is positive by definition).

Criterion 3: V_x is finite (Lemma 2).

Clearly, Criterion 1 is satisfied by selecting appropriate values for μ_c and μ_ω according to (8), so let us concentrate on conditions to satisfy Criteria 2 and 3.

Lemma 3: Let us assume that the expectation of x_t is convergent (Criterion 1) and σ_p, σ_g, and $(\mu_g - \mu_p)$ are finite. V_x non-negative and finite (Criteria 2 and 3) are guaranteed if and only if $k_2 < 0$.

Proof: Let us start with conditions to guarantee that V_x is non-negative. It is obvious that $k_3 \geq 0$ (all of its components are squared). In addition, because $\mu_\omega + 1 > 0$ (necessary for convergence of expectation), $k_4 \geq 0$ is guaranteed. Thus, $V_x \geq 0$ if and only if k_2 is negative.

A necessary and sufficient condition to guarantee that V_x is finite is that $k_1 k_2 \neq 0$ and k_3 and k_4 are finite. As σ_p, σ_g, and $(\mu_g - \mu_p)$ are finite, k_3 and k_4 are guaranteed to be finite. Also, $k_1 > 0$ because the expectation is convergent [see (8)]; hence, $k_1 \neq 0$. Therefore, a necessary and sufficient condition to guarantee that V_x is finite is that $k_2 \neq 0$. Therefore, $V_x \geq 0$ and finite if and only if $k_2 < 0$. ■

Note that this proof relies on σ_p, σ_g, and $(\mu_g - \mu_p)$ being finite. As discussed in Section IV-A, $(\mu_g - \mu_p)$, σ_p, and σ_g are finite for any arbitrary precision $\epsilon > 0$.

Theorem 1: The following conditions are necessary for the convergence of variance.

Condition 1: $-1 < \mu_\omega < 1$, and $0 < \mu_{c_1} + \mu_{c_2} < 2(\mu_\omega + 1)$.

Condition 2: $k_2 < 0$.

Proof: Criteria 1 to 3 are necessary conditions for the convergence of variance. Criterion 1 is equivalent to Condition 1. Also, according to Lemma 3, Condition 2 is necessary and sufficient for the existence of the variance fixed point. As the existence of the fixed point is a necessary condition for the convergence of variance (Lemma 2), both Conditions 1 and 2 are necessary for the convergence of variance. ■

We simplify conditions in Theorem 1 to

$$- a < \mu_\omega < a \text{ and } 0 < \mu_c < \frac{-2\left(\mu_\omega^2 + \sigma_\omega^2 - 1\right)}{1 - \mu_\omega + q(1 + \mu_\omega)} \qquad (19)$$

where $a = \sqrt{1 - \sigma_\omega^2}$, $q = (\sigma_c^2/\mu_c^2)$, $\sigma_c^2 = \sigma_{c_1}^2 + \sigma_{c_2}^2$, and $\mu_c^2 = (\mu_{c_1} + \mu_{c_2})^2$. Note that μ_ω is real, so $\sigma_\omega^2 \leq 1$, which means $a \leq 1$.

Based on Lemma 2 and Theorem 1, (19) is a necessary condition to ensure $\rho(M) < 1$. However, it is not possible to claim that if the conditions in (19) are satisfied, then $\rho(M) < 1$ [(19) might not be a sufficient condition for $\rho(M) < 1$]. We run an experiment to test whether this necessary condition is also a sufficient condition for the convergence of variance of x_t. The experiment was as follows: μ_ω was changed in $(-1, 1)$ with the step size 0.001 and, for each value of μ_ω, 1000 random combinations for $< \mu_{c_1}, \mu_{c_2}, \sigma_{c_1}, \sigma_{c_2}, \sigma_\omega >$ were generated in a way that the conditions in (19) were satisfied.[6] For each setting, we calculated the value of $\rho(M)$ and tested if $\rho(M) < 1$ is true (1 000 000 combinations of settings in total). Results showed that all tested combinations of parameters that satisfy conditions in (19) also satisfy $\rho(M) < 1$. Although this is an experimental approach, it indicates that the conditions in Theorem 1 are very likely to be sufficient to ensure $\rho(M) < 1$, i.e., (19) is necessary and (most likely) sufficient for convergence of variance.

For IPSO, because $\mu_{c_1} = (\phi_1/2)$, $\mu_{c_2} = (\phi_2/2)$, $\mu_\omega = \omega$, $\sigma_{c_1} = (\phi_1/\sqrt{12})$, $\sigma_{c_2} = (\phi_2/\sqrt{12})$, and $\sigma_\omega = 0$, we have the following.

1) $k_1 = ((\phi_1/2) + (\phi_2/2))^2$.
2) $k_2 = k_1(1-\omega) + (\phi_1 + \phi_2)(\omega^2 - 1) + ((\phi_1^2/12) + (\phi_2^2/12))(\omega + 1)$.
3) $k_3 = k_1(\omega + 1)((\phi_1^2/3)\sigma_p^2 + (\phi_2^2/3)\sigma_g^2)$.
4) $k_4 = ((\phi_1^2 \phi_2^2/24))(\omega + 1)(\mu_g - \mu_p)^2$.

The convergence relation calculated by (19) for the IPSO setting where p and g are allowed to move randomly is then written as

$$- 1 < \omega < 1 \text{ and } 0 < \phi_1 + \phi_2 < 24\frac{\omega^2 - 1}{5\omega - 7} \qquad (20)$$

which is exactly aligned with what was found by Poli [8] and Jiang *et al.* [19] with stagnation assumption [13] and with a weaker stagnation assumption. This means that random movement of p and g does not affect the convergence boundaries to guarantee convergence of variance.

4) Convergence of V_x to Zero: To guarantee second-order stability one needs to guarantee the convergence of variance and that the fixed point of variance (V_x) is zero. Given that Conditions 1 and 2 in Theorem 1 guarantee convergence of variance, we need to guarantee $V_x = 0$ to ensure second-order stability. By using Condition 1 in Theorem 1, it is easy to see that k_1 and k_2 are finite; hence, they do not play a role to ensure $V_x = 0$. Thus, the value of V_x is zero if and only if $k_3 + k_4 = 0$. Because $k_3 \geq 0$ and $k_4 \geq 0$ are always guaranteed (see the proof for Lemma 3), $k_3 + k_4 = 0$ is equivalent to $k_3 = k_4 = 0$. Hence, in order to guarantee second-order stability we need to guarantee satisfaction of Conditions 1 and 2 (Theorem 1) together with $k_3 = k_4 = 0$ (the fixed point of variance is zero). $k_3 = k_4 = 0$ if and only if at least one of the following cases are true:

Case 1: $\sigma_p^2 = 0$, $\sigma_g^2 > 0$, $\mu_{c_2}^2 = \sigma_{c_2}^2 = 0$, and $\mu_{c_1}^2 \neq 0$ or;

Case 2: $\sigma_p^2 > 0$, $\sigma_g^2 = 0$, $\mu_{c_1}^2 = \sigma_{c_1}^2 = 0$, and $\mu_{c_2}^2 \neq 0$ or;

Case 3: $\sigma_p^2 = \sigma_g^2 = 0$ and $\mu_g = \mu_p$ (aka stagnation) or;

Case 4: $\sigma_p^2 = \sigma_g^2 = 0$ and $\mu_g \neq \mu_p$ and

1) $\mu_{c_2}^2 = \sigma_{c_2}^2 = 0$ and $\mu_{c_1}^2 \neq 0$ or;
2) $\mu_{c_1}^2 = \sigma_{c_1}^2 = 0$ and $\mu_{c_2}^2 \neq 0$.

Case 1 refers to when p stops moving while g is always updated,[7] Case 2 refers to when g stops moving and p is always updated, Case 3 refers to when p and g stop moving at the same location, and Case 4 refers to when p and g stop moving but at different locations. Note that, as the fixed point V_x is met in limit when $t \to \infty$, the movement of p and g can also be investigated in limit so that $\sigma_p^2 = 0$ and $\sigma_g^2 = 0$ can be replaced by $\sigma_p^2 \to 0$ and $\sigma_g^2 \to 0$, respectively.

One should note that these cases are very generic and, under some assumptions, they can be simplified. For example, $\mu_g = \mu_p$ and $\sigma_g = \sigma_p$ are always true for the global best particle. Hence, if σ_p converges to zero and its coefficients are in the variance convergence boundaries, then this particle is second-order stable. As discussed in Section IV-A, σ_p converges to zero for any arbitrary precision $\epsilon > 0$ as the number of iterations grows. Hence, for any arbitrary precision $\epsilon > 0$, the convergence of variance is equivalent to the second-order stability for the global best particle.

To summarize, a recurrence relation in the form of (6), under assumptions p and g are random variables, is second-order stable if and only if Conditions 1 and 2 in Theorem 1 are satisfied and one of the four cases mentioned before is true.

[6] The random variables p and g were excluded from this experiment as they do not affect $\rho(M)$ and conditions in (19).

[7] Note that this does not mean that g moves in every iteration, but it means that, for any iteration t, there exists an iteration $t_0 > t$ that g is updated.

For IPSO, we calculate V_x as [see (18)]

$$V_x = \frac{\phi(\omega+1)\left[8\left(\sigma_p^2+\sigma_g^2\right)+(\mu_g-\mu_p)^2\right]}{4\left(\phi(5\omega-7)-12\omega^2+12\right)} \tag{21}$$

where $\phi = \phi_1 = \phi_2$. Under this setting and in stagnation (Case 3), p and g are constants that result in $\sigma_p^2 + \sigma_g^2 = 0$; hence, (21) becomes exactly the same as what was found in [8]. The algorithm becomes second-order stable only if $\mu_p = \mu_g$ (p and g are at the same location) that is the same as what was found by Poli [8].

V. Conclusion

In this letter, we introduced a stochastic recurrence relation that represents a wide class of PSO algorithms. We investigated the convergence boundaries for this relation to guarantee convergence of expectation and variance as well as first-order and second-order stability. Our study assumed that the memories of particles (i.e., personal bests of particles) are random variables that are updated during the run with a mean and a variance. This assumption simulates the movement of those memories and, hence, the convergence boundaries are not specific to stagnation. Also, this assumption makes the calculations independent of the topology as the global best (local best) vector is assumed to be updated randomly no matter which particle updates it.[8] We determined (Section IV-B) necessary and sufficient condition to guarantee convergence of expectation and first-order stability of this recurrence relation. Interestingly, this condition was independent of the mean and variance of memories, meaning that whether the expectation of the position of particles is convergent or the particles are first-order stable is independent of the mean and variance of the movement of memories. We found that the convergence boundaries for the first-order stability of IPSO calculated by our proposed model are exactly the same as the boundaries provided in earlier studies under stagnation and weakened stagnation assumptions. We also proved (Section IV-C2) that the necessary and sufficient condition for convergence of the variance of the recurrence relation is independent of the mean and variance of the movement of the memories (i.e., personal best and global best). We determined (Section IV-C3) convergence boundaries that guarantee convergence of variance for that recurrence relation. We showed that, for a particular case of that recurrence relation that is equivalent to IPSO, these boundaries are the same as what was found in earlier studies for this algorithm. We also found (Section IV-C4) that, to guarantee second-order stability under the assumption of p and g are random variables in a general case, either p or g or both need to stop movement. As a future direction, one can extend our proposed model to more complex topologies (see [22]) and study their effects on the stability and convergence of particles. In addition, as many PSO variants consider an upper limit for the velocity ($v_{\max}$), one can study how such limit can influence convergence of variance and stability of particles.

[8]Note that the formulation models the impact of one other particle only and does not model fully informed topologies.

References

[1] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proc. Int. Conf. Neural Netw.*, vol. 4. Perth, WA, Australia, 1995, pp. 1942–1948.

[2] A. P. Engelbrecht, *Fundamentals of Computational Swarm Intelligence*. Chichester, U.K.: Wiley, 2006.

[3] R. Poli, "Analysis of the publications on the applications of particle swarm optimisation," *J. Artif. Evol. Appl.*, vol. 2008, no. 3, pp. 1–10, 2008. [Online]. Available: http://www.hindawi.com/archive/2008/685175/

[4] M. R. Bonyadi and Z. Michalewicz, "A locally convergent rotationally invariant particle swarm optimization algorithm," *Swarm Intell.*, vol. 8, no. 3, pp. 159–198, 2014.

[5] F. van den Bergh and A. P. Engelbrecht, "A convergence proof for the particle swarm optimiser," *Fund. Informat.*, vol. 105, no. 4, pp. 341–374, 2010.

[6] M. R. Bonyadi, Z. Michalewicz, and X. Li, "An analysis of the velocity updating rule of the particle swarm optimization algorithm," *J. Heuristics*, vol. 20, no. 4, pp. 417–452, 2014.

[7] D. N. Wilke, S. Kok, and A. A. Groenwold, "Comparison of linear and classical velocity update rules in particle swarm optimization: Notes on scale and frame invariance," *Int. J. Numer. Methods Eng.*, vol. 70, no. 8, pp. 985–1008, 2007.

[8] R. Poli, "Mean and variance of the sampling distribution of particle swarm optimizers during stagnation," *IEEE Trans. Evol. Comput.*, vol. 13, no. 4, pp. 712–721, Aug. 2009.

[9] M. Clerc and J. Kennedy, "The particle swarm—Explosion, stability, and convergence in a multidimensional complex space," *IEEE Trans. Evol. Comput.*, vol. 6, no. 1, pp. 58–73, Feb. 2002.

[10] C. W. Cleghorn and A. P. Engelbrecht, "A generalized theoretical deterministic particle swarm model," *Swarm Intell.*, vol. 8, no. 1, pp. 35–59, 2014.

[11] F. van den Bergh and A. P. Engelbrecht, "A study of particle swarm optimization particle trajectories," *Inf. Sci.*, vol. 176, no. 8, pp. 937–971, 2006.

[12] M. Bonyadi and Z. Michalewicz, "Analysis of stability, local convergence, and transformation sensitivity of a variant of particle swarm optimization algorithm," *IEEE Trans. Evol. Comput.*, doi: 10.1109/TEVC.2015.2460753. [Online]. Available: http://ieeexplore.ieee.org/xpl/articleDetails.jsp?arnumber=7166327&queryText=2460753&newsearch=true

[13] Q. Liu, "Order-2 stability analysis of particle swarm optimization," *Evol. Comput.*, vol. 23, no. 2, pp. 187–216, 2014.

[14] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proc. World Congr. Comput. Intell.*, Anchorage, AK, USA, 1998, pp. 69–73.

[15] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. Int. Symp. Micro Mach. Human Sci.*, Nagoya, Japan, 1995, pp. 39–43.

[16] P. B. Lax, *Functional Analysis*. New York, NY, USA: Wiley-Intersci., 2002.

[17] M. Clerc, "The swarm and the queen: Towards a deterministic and adaptive particle swarm optimization," in *Proc. Congr. Evol. Comput.*, vol. 3. Washington, DC, USA, 1999, pp. 1951–1957.

[18] I. C. Trelea, "The particle swarm optimization algorithm: Convergence analysis and parameter selection," *Inf. Process. Lett.*, vol. 85, no. 6, pp. 317–325, 2003.

[19] M. Jiang, Y. P. Luo, and S. Y. Yang, "Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm," *Inf. Process. Lett.*, vol. 102, no. 1, pp. 8–16, 2007.

[20] E. García-Gonzalo and J. L. Fernández-Martínez, "Convergence and stochastic stability analysis of particle swarm optimization variants with generic parameter distributions," *Appl. Math. Comput.*, vol. 249, pp. 286–302, 2014.

[21] P. J. Angeline, "Evolutionary optimization versus particle swarm optimization: Philosophy and performance differences," in *Evolutionary Programming VII*. Heidelberg, Germany: Springer, 1998, pp. 601–610.

[22] R. Mendes, J. Kennedy, and J. Neves, "The fully informed particle swarm: Simpler, maybe better," *IEEE Trans. Evol. Comput.*, vol. 8, no. 3, pp. 204–210, Jun. 2004.