Particle Convergence Expected Time in The PSO Model with Inertia Weight

Krzysztof Trojanowski and Tomasz Kulpa
Cardinal Stefan Wyszyński University, Faculty of Mathematics and Natural Sciences, Warsaw, Poland

Keywords: Particle Swarm with Inertia Weight, Particle Location Convergence Expected Time, Location Variance Convergence Expected Time.

Abstract: Theoretical properties of particle swarm optimization approach with inertia weight are investigated. Particularly, we focus on the convergence analysis of the expected value of the particle location and the variance of the location. Four new measures of the expected particle convergence time are defined: (1) convergence of the expected location of the particle, (2) the particle location variance convergence and (3-4) their respective weak versions. For the first measure an explicit formula of its upper bound is also given. For the weak versions of the measures graphs of recorded values are presented.

1 INTRODUCTION

Particle swarm optimization (PSO) is a stochastic population-based search algorithm successfully applied in numerous real-world problems (Poli, 2008a; Bonyadi and Michalewicz, 2016b). Usually, when PSO is implemented some drawbacks or limitations can be observed. They can be divided into two main groups: related to transformation invariance and to convergence (Bonyadi and Michalewicz, 2016b). The latter group concerns problems with stability and local convergence of a swarm, patterns of particle movements, and first hitting time. All of them were a subject of theoretical analysis.

Phenomenon of uncontrolled growth of particle velocities for some values of velocity equation coefficients was one of the first identified limitation of PSO. Obtaining a non divergent behavior of a swarm needed to identify boundaries for a so called convergence region of safe coefficients values. Even for the PSO configuration from this region there appeared a problem of swarm stagnation. This is a case when swarm obtains its equilibrium state and converges to a point which is, however, not a local optimum.

Another issue concerning effectiveness of the search process are the patterns of particle movements. For velocity equation coefficients from the convergence region one can observe different patterns of particles paths. Depending on the optimized function different configurations prove to be the most efficient. However, there exist coefficients settings commonly regarded as a "good starting point" of PSO configuration tuning for selected classes of problems.

In the case of the PSO first hitting time issue, the subject of interest is the time (precisely, a number of evaluation function calls) necessary to obtain satisfactory solution. Due to stochastic nature of PSO an expected runtime of the algorithm is rather investigated. In the presented research we focus on this very aspect of the theoretical analysis. New definitions of particle convergence in the stochastic model of the particle movement are proposed and estimations of the number of steps necessary for the particle to obtain the stability state are presented.

The paper consists of six sections. In Section 2 a brief review of selected areas of PSO theoretical analysis can be found, that is, analysis concerning (1) stability and region of stable particle parameter configurations and (2) runtime analysis, particularly, estimation of times necessary to hit a staisfying solution. In Section 3 the stochastic model of the particle movement is presented. Section 4 introduces definitions of particle convergence expected time (*pcet*) and particle weak convergence expected time (*pwcet*). Section 5 focuses on the convergence of particle location variance and introduces next two definitions of the particle location variance convergence time $pvct(\delta)$ and its weak version. Section 6 concludes the paper.

Trojanowski, K. and Kulpa, T.
Particle Convergence Expected Time in The PSO Model with Inertia Weight.
DOI: 10.5220/0006048700690077
In *Proceedings of the 8th International Joint Conference on Computational Intelligence (IJCCI 2016) - Volume 1: ECTA*, pages 69-77
ISBN: 978-989-758-201-1

2 RELATED WORK

The PSO model with inertia weight implements following velocity and position equations:

$$\begin{cases} \mathbf{v}_{t+1} = w \cdot \mathbf{v}_t + \varphi_{t,1} \otimes (\mathbf{y}_t - \mathbf{x}_t) + \varphi_{t,2} \otimes (\mathbf{y}^*_t - \mathbf{x}_t), \\ \mathbf{x}_{t+1} = \mathbf{x}_t + \mathbf{v}_{t+1} \end{cases} \quad (1)$$

where $\mathbf{v}_t$ is a particle's velocity, $\mathbf{x}_t$ — particle's location, $\mathbf{y}_t$ — the best location the particle has found so far, $\mathbf{y}^*_t$ — the best location found by particles in its neighborhood, w – inertia coefficient, $\varphi_{t,1}$ and $\varphi_{t,2}$ control influence of the attractors on the velocity, $\varphi_{t,1} = R_{t,1}c_1$, $\varphi_{t,2} = R_{t,2}c_2$, and c_1,c_2 represent acceleration coefficients, $R_{t,1}, R_{t,2}$ are two vectors of random values uniformly generated in range $[0,1]$ and $\otimes$ denotes pointwise vector product. Values of coefficients w, c_1 and c_2 define convergence properties of the particle.

2.1 Stability and Stable Regions

In (Cleghorn and Engelbrecht, 2015) assumptions accompanying theoretical PSO research can be classified into the following four: (1) deterministic assumption, where $\varphi_1 = \varphi_{t,1}$ and $\varphi_2 = \varphi_{t,2}$, for all t, (2) stagnation assumption, where $\mathbf{y}_t = \mathbf{y}$ and $\mathbf{y}^*_t = \mathbf{y}^*$, for all t sufficiently large, (3) weak chaotic assumption, where both $\mathbf{y}_t$ and $\mathbf{y}^*_t$ will occupy an arbitrarily large but finite number of unique position, and (4) weak stagnation assumption, where the global attractor of the particle that has obtained the best objective function evaluation remains constant for all t sufficiently large. Under the deterministic assumption the following region of particle convergence was derived ((Trelea, 2003; van den Bergh and Engelbrecht, 2006)):

$$\begin{cases} \mathbf{0} < \varphi_1 + \varphi_2 < \mathbf{2}(1+w), \\ 0 < w < 1, \quad \varphi_1 > \mathbf{0} \wedge \varphi_2 > \mathbf{0} \end{cases} \quad (2)$$

and the stability is defined as $\lim_{t\to\infty} \mathbf{x}_t = \mathbf{y}$.

To deal with randomness of $\varphi_{t,1}$ and $\varphi_{t,2}$ they are replaced with their expectations $c_1/2$ and $c_2/2$ respectively. In this case stability is defined as $\lim_{t\to\infty} E|\mathbf{x}_t| = \mathbf{y}$ ((Poli, 2009)) and is called the order-1 stability. The region defined with Ineq. (2) satisfies this stability, thus, it is also called the order-1 stable region. In later publications (e.g. (Cleghorn and Engelbrecht, 2014; Bonyadi and Michalewicz, 2016a; Liu, 2015)) the region is extended to $|w| < 1$ and $\mathbf{0} < \varphi_1 + \varphi_2 < \mathbf{2}(1+w)$.

Unfortunately, the order-1 stability is not enough to ensure convergence, simply the particle may oscillate or even diverge and the expectation converges to a point. The convergence of the variance (or standard deviation) is also necessary, which is called the order-2 stability condition ((Jiang et al., 2007; Poli, 2009)). In (Jiang et al., 2007) the stability is defined as $\lim_{t\to\infty} E[\mathbf{x}_t - \mathbf{y}]^2 = 0$ where $\mathbf{y} = \lim_{t\to\infty} E[\mathbf{x}_t]$. In (Poli, 2009) the stability is defined as $\lim_{t\to\infty} E[\mathbf{x}^2_t] = \beta_0$ and $\lim_{t\to\infty} E[\mathbf{x}_t\mathbf{x}_{t-1}] = \beta_1$ where β_0 and β_1 are constant. Eventually, both authors obtain the same set of inequalities which define the so called order-2 stable region:

$$\varphi < \frac{12(1-w^2)}{7-5w} \text{ where } \varphi_1 = \varphi_2 = \varphi. \quad (3)$$

2.2 Runtime Analysis

For applications of PSO for real-world problems it is important to estimate when a swarm or a particle reaches close vicinity of the optimum. Need for analysis of this problem appeared in (Witt, 2009) and in (Lehre and Witt, 2013) authors introduced formal definition of the first hitting time (FHT) and expected FHT (EFHT). Both concepts refer to an entire swarm, precisely, FHT represents the number of times the evaluation function f_{eval} is called until the swarm for the first time contains a particle $\mathbf{x}$ for which $|f_{\text{eval}}(\mathbf{x}) - f_{\text{eval}}(\mathbf{y}^*)| < \delta$.

Another approach can be found in (Trojanowski and Kulpa, 2015), where subsequent locations of particles are a subject of analysis. Authors proposed a concept of particle convergence time (*pct*) as a measure of speed at which the equilibrium state is reached. In this case the "equilibrium state" is the state when the distance between current and the next location of the particle is never greater than the given threshold value δ. Authors assumed that the global attractor remains unchanged (the so-called stagnation assumption), that is, the value of global attractor is never worse than the value of any location visited during the convergence process. This means that the shape of evaluation function f_{eval} is negligible as far as this condition is satisfied.

Definition 2.1 (The particle convergence time). *Let δ be a given positive number and $S(\delta)$ be a set of natural numbers such that:*

$$s \in S(\delta) \iff ||\mathbf{x}_{t+1} - \mathbf{x}_t|| < \delta \text{ for all } t \geq s. \quad (4)$$

The particle convergence time ($pct(\delta)$) is the minimal number in the set $S(\delta)$, that is

$$pct(\delta) = \min\{s \in S(\delta)\}. \quad (5)$$

Under the deterministic and stagnation assumptions, and also the best particle stagnation assumption (that is, $\mathbf{y}_t = \mathbf{y}^*_t = \mathbf{y}$), the explicit version of an upper bound formula of (pct), that is, $pctb(\delta)$ is given ((Trojanowski and Kulpa, 2015)).

3 THE STOCHASTIC MODEL

Under the best particle stagnation assumption the update equation of the particle location in one-dimensional search space can be reformulated as follows:

$$x_{t+1} = (1+w-\phi_t)x_t - wx_{t-1} + \phi_t y, \quad (6)$$

where w is a constant parameter of inertia and ϕ_t is the sum of two independent random variates, $\phi_t = \varphi_{t,1} + \varphi_{t,2}$, $\varphi_{t,i} \sim U(0, c_i)$, $i = 1,2$. It is also assumed that ϕ_t, $t = 1,2,3\ldots$ are independent and identically distributed.

Thus, in the further evaluations $\mathrm{E}[\phi_t]$ and $\mathrm{E}[\phi_t^2]$ equal

$$\begin{aligned} \mathrm{E}[\phi_t] =& \mathrm{E}[\phi_{t,1}] + \mathrm{E}[\phi_{t,2}] = \frac{c_1+c_2}{2} \\ \mathrm{E}[\phi_t^2] =& \mathrm{Var}[\phi_t] + (\mathrm{E}[\phi])^2 = \\ & \frac{c_1^2}{12} + \frac{c_2^2}{12} + \left(\frac{c_1+c_2}{2}\right)^2 \end{aligned}$$

Set $e_t = E[x_t]$, $m_t = E[x_t^2]$, $h_t = E[x_t x_{t-1}]$, $f = E[\phi_t]$ and g=$E[\phi_t^2]$.

The proposed model is a simplified version of the model presented in (Poli and Broomhead, 2007; Poli, 2008b; Poli, 2009), particularly, we apply the same analysis of dynamics of first and second moments of the PSO sampling distribution.

We apply the expectation operator to both sides of Eq. (6). Because of the statistical independence between ϕ_t and x_t we obtain

$$e_{t+1} = (1+w-f)e_t - we_{t-1} + fy. \quad (7)$$

Eq. (7) gives us the same model as the model described by Eq. (6), however, instead of the acceleration coefficient ϕ_t we have its expected value f and instead of the particle location x_t we have particle expected location e_t. We can say that the update of expected position of a particle follows in the same way as the particle trajectory in the deterministic model described by Eq. (6).

We raise both sides of Eq. (6) to the second power and obtain

$$\begin{aligned} x_{t+1}^2 =& (1+w-\phi_t)^2 x_t^2 + w^2 x_{t-1}^2 + \phi_t^2 y^2 \\ & - 2(1+w-\phi_t)wx_t x_{t-1} - 2wy\phi_t x_{t-1} \\ & + 2y\phi_t(1+w-\phi_t)x_t \end{aligned} \quad (8)$$

Applying the expectation operator to both sides of Eq. (8) and again because of the statistical independence between ϕ_t, x_t and x_{t-1} we obtain

$$\begin{aligned} m_{t+1} =& m_t((1+w)^2 - 2(1+w)f + g) \\ & + m_{t-1}w^2 - h_t 2w(1+w-f) \\ & + e_t 2y(f(1+w) - g) \\ & - e_{t-1} 2wyf + y^2 g \end{aligned} \quad (9)$$

Multiplying both sides of Eq. (6) by x_t we get

$$x_{t+1}x_t = (1+w-\phi_t)x_t^2 - wx_t x_{t-1} + \phi_t y x_t \quad (10)$$

Again, we apply the expectation operator to (10) and obtain

$$h_{t+1} = (1+w-f)m_t - wh_t + fye_t \quad (11)$$

Now, a vector $\mathbf{z}_t = (e_t, e_{t-1}, m_t, m_{t-1}, h_t)^T$ can be introduced. Equations (7), (9), and (11) can be rewritten as a matrix equation

$$\mathbf{z}_{t+1} = M_t \mathbf{z}_t + \mathbf{b} \quad (12)$$

where

$$M_t = \begin{bmatrix} m_{1,1} & -w & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ m_{3,1} & m_{3,2} & m_{3,3} & w^2 & m_{3,5} \\ 0 & 0 & 1 & 0 & 0 \\ fy & 0 & m_{5,3} & 0 & -w \end{bmatrix} \quad (13)$$

where the matrix components are

$$\begin{aligned} m_{1,1} &= 1+w-f, \\ m_{3,1} &= 2y(f(1+w)-g), \\ m_{3,2} &= -2wyf, \\ m_{3,3} &= (1+w)^2 - 2(1+w)f + g, \\ m_{3,5} &= -2w(1+w-f), \\ m_{5,3} &= 1+w-f. \end{aligned}$$

and

$$\mathbf{b} = (fy, 0, y^2 g, 0, 0)^T \quad (14)$$

The particle is order-2 stable if e_t, m_t, and h_t converge to to stable fixed points. This happens when all absolute values of eigenvalues of M are less than 1.

In that case, there exist a fixed point of the system described by equation

$$\mathbf{z}^* = (I-M)^{-1}\mathbf{b}. \quad (15)$$

When the system is order-2 stable, by the change of variables $\mathbf{u}_t = \mathbf{z}_t - \mathbf{z}^*$, we can rewrite Eq. (12)

$$\mathbf{u}_{t+1} = M\mathbf{u}_t, \quad (16)$$

which can be integrated to obtain the explicit formula

$$\mathbf{u}_t = M^t \mathbf{u}_0. \quad (17)$$

The order-2 analysis of the system described by Eq. (17) is not easy because of complicated formulas for eigenvalues of M. However, the order-1 analysis can be done, because two of them are known as

$$\begin{aligned} \lambda_1 &= \frac{1+w-f+\gamma}{2}, \\ \lambda_2 &= \frac{1+w-f-\gamma}{2}, \end{aligned} \quad (18)$$

where

$$\gamma = \sqrt{(1+w-f)^2 - 4w}. \quad (19)$$

For fixed initial values of e_0 and e_1, the explicit formula for e_t, first time obtained by (van den Bergh and Engelbrecht, 2006), is given by equation

$$e_t = k_1 + k_2\lambda_1^t + k_3\lambda_2^t, \tag{20}$$

where

$$\begin{aligned} k_1 &= y, \\ k_2 &= \frac{\lambda_2(e_0 - e_1) - e_1 + e_2}{\gamma(\lambda_1 - 1)}, \\ k_3 &= \frac{\lambda_1(e_1 - e_0) + e_1 - e_2}{\gamma(\lambda_2 - 1)}, \\ e_2 &= (1 + w - f)e_1 - we_0 + fy. \end{aligned} \tag{21}$$

4 PARTICLE CONVERGENCE EXPECTED TIME

Due to the analogy between the deterministic model based on the update equation of the particle location (6) and the studied order-1 stochastic model of PSO described by Eq. (7) we can define a measure of particle convergence expected time (*pcet*) respectively to the idea given in Def. (2.1),

Definition 4.1 (The particle convergence expected time). *Let δ be a given positive number and $S(\delta)$ be a set of natural numbers such that:*

$$s \in S(\delta) \iff |e_{t+1} - e_t| < \delta \text{ for all } t \geq s. \tag{22}$$

The particle convergence expected time ($pcet(\delta)$) is the minimal number in the set $S(\delta)$, that is

$$pcet(\delta) = \min\{s \in S(\delta)\}. \tag{23}$$

Briefly, the particle convergence expected time *pcet* is the minimal number of steps necessary for the expected particle location to obtain its stable state as defined above.

The explicit formula for solutions of the recurrence Eq. (6) is given in (van den Bergh and Engelbrecht, 2006). This formula was used in (Trojanowski and Kulpa, 2015) to find an upper bound formula of *pct*, that is, $pctb(\delta)$. Because of the analogy between the models described by Eq. (6) and Eq. (7) we obtain the following upper bound for *pcet*, namely *pcetb*

$$pcetb(\delta) = \max\left(\frac{\ln\delta - \ln(2|k_2||\lambda_1 - 1|)}{\ln|\lambda_1|}, \frac{\ln\delta - \ln(2|k_3||\lambda_2 - 1|)}{\ln|\lambda_2|}\right) \tag{24}$$

for real value of γ given by (19) and

$$pcetb(\delta) = \frac{\ln\delta - \ln(|\lambda_1 - 1|(|k_2| + |k_3|))}{\ln|\lambda_1|} \tag{25}$$

for imaginary value of γ, where λ_1 and λ_2 are given by Eq. (18) and k_1, k_2 and k_3 are given by Eq. (21).

Obviously, characteristics of $pcetb(\delta)$ depicted in Fig. 1 (generated for $\delta = 0.0001$) looks the same as the characteristics of *pctb* (see (Trojanowski and Kulpa, 2015) for comparisons) and have the same distinctive shape of a funnel. Thus, as in the case of *pctb*, they can also be classified into four main types.

Empirical evaluation of *pcet* is difficult, so, we introduce the less restrictive measure, that is, a particle weak convergence time.

Definition 4.2 (The particle weak convergence expected time). *Let δ be a given positive number. The particle weak convergence expected time $pwcet(\delta)$ is the minimal number of steps necessary to get the expected value of difference between subsequent particle locations lower than δ, that is*

$$pwcet(\delta) = min\{t : |e_t - e_{t+1}| < \delta\}. \tag{26}$$

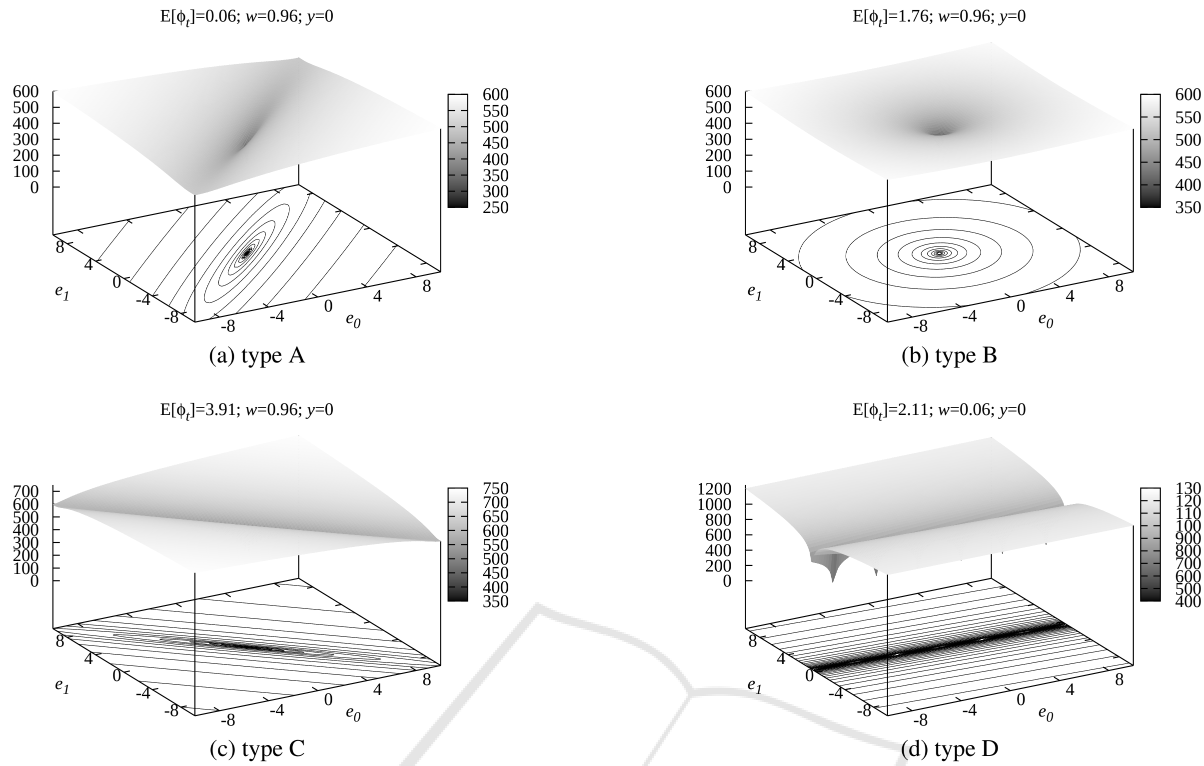
It is obvious that $pwcet(\delta) \leq pcet(\delta)$ and equality generally does not hold. Empirical characteristics of *pwcet* are depicted in Fig. 2 and Fig. 3. The characteristics were obtained with Algorithm 1.

Algorithm 1: Particle weak convergence expected time evaluation procedure.

```
Initialize: T_max = 1e+5, two successive expected
    locations e_0 and e_1, and an attractor of a particle,
    for example, y = 0.
s_1 = e_1 - e_0
f = (c_1 + c_2)/2
t = 1
repeat
    e_{t+1} = (1 + w - f)e_t - we_{t-1} + fy
    s_{t+1} = e_{t+1} - e_t
    t = t + 1
until (s_t > δ) ∧ (s_t < 1e+10) ∧ (t < T_max)
if s_t < 1e+10 then
    return t
else
    return T_max
end if
```

Fig. 2 depicts the values of *pwcet* generated for $\delta = 0.0001$ as a function of initial location and velocity represented by expected locations e_0 and e_1 where $E[\phi_t]$ and w are fixed. A grid of pairs $[e_0, e_1]$ consists of 40000 points (200×200) varying from -10 to 10 for both e_0 and e_1.

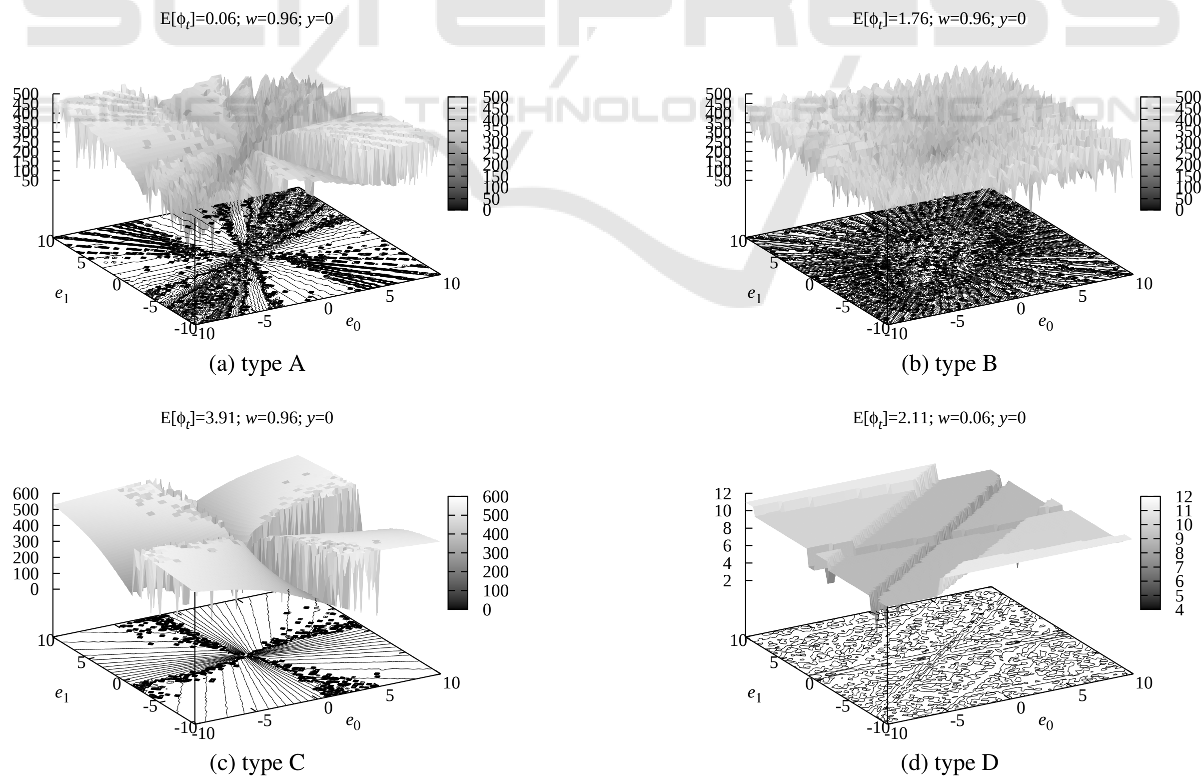
Fig. 3 shows the values of *pwcet* also for $\delta = 0.0001$ obtained for a grid of configurations $(\phi_{\max}, w)$ starting from $[\phi_{\max} = 0.0,\ w = -1.0]$ and changing with step 0.02 for w and step 0.04 for $\phi_{\max}$ (which gave 200×100 points).



(a) type A (b) type B

(c) type C (d) type D

Figure 1: Graphs of $pcetb(e_0, e_1)$ for selected configurations $(E[\phi_t], w)$.



(a) type A (b) type B

(c) type C (d) type D

Figure 2: Graphs of recorded values of $pwcet(e_0, e_1)$ for selected configurations $(E[\phi_t], w)$.

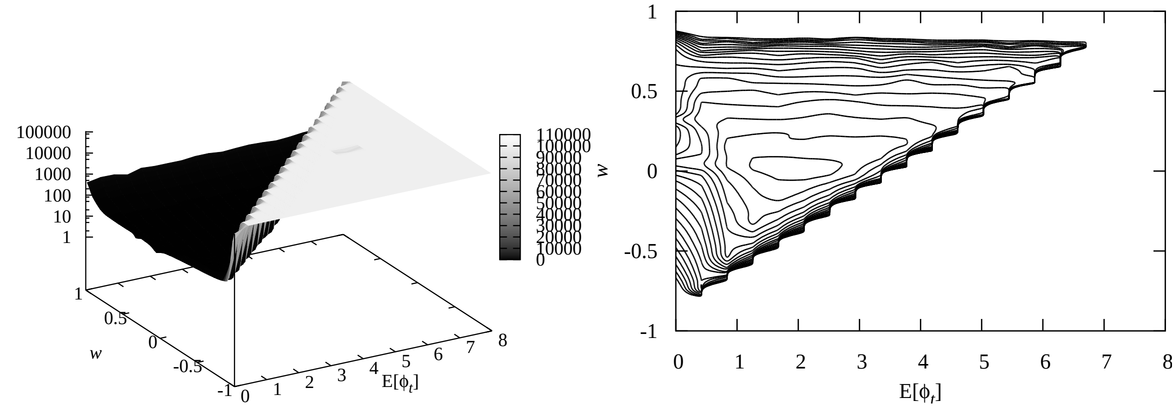


Figure 3: Recorded convergence times of the particle location $pwcet(E[\phi_t],w)$ for example starting conditions: $e_0 = -9$ and $e_1 = -5$; 3D shape with logarithmic scale for $pwcet(E[\phi_t],w)$ (left graph), and isolines from 0 to 100 with step 5 (right graph).

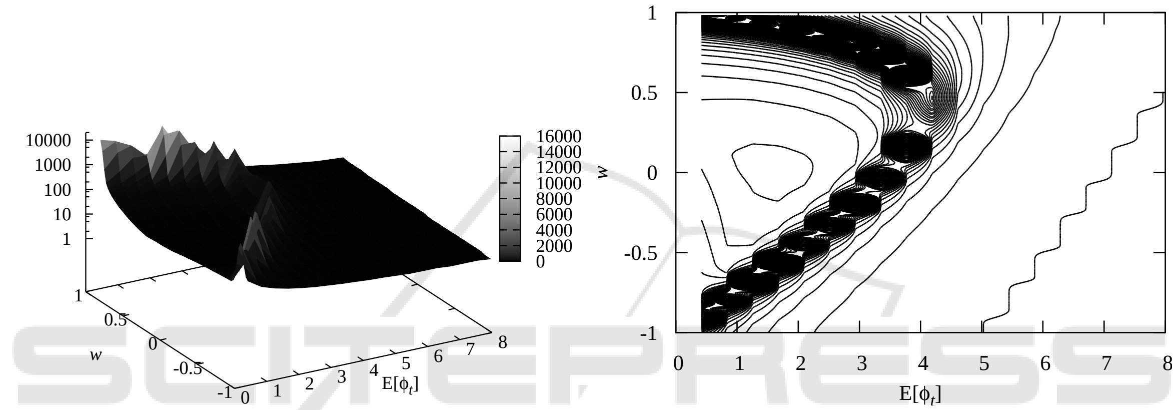


Figure 4: Recorded convergence times of the particle location variance $pvwct(E[\phi_t],w)$ for example starting conditions: $e_0 = -9$ and $e_1 = -5$; 3D shape with logarithmic scale for $pvwct(E[\phi_t],w)$ (left graph), and isolines from 0 to 20000 with step 10 (right graph).

In both figures the configurations generating $pwcet > 100000$ have assigned a constant value of 100000. It is also assumed that $c_1 = c_2 = \phi_{max}/2$.

5 CONVERGENCE OF VARIANCE OF PARTICLE LOCATION DISTRIBUTION

Convergence of the expected value of the particle location still does not guarantee the convergence of the particle position. This is the case, for example, where the particle oscillates symmetrically and the oscillations do not fade. In (Poli, 2009) author studied convergence of the variance and standard deviation of the particle location and obtained region (Ineq. (3)) of the order-2 stability of the system. In the studied model with the best particle stagnation assumption described by Eq. (6) the variance of the particle location converges to zero for the configurations originating from the order-2 stability region (Ineq. 3).

It is interesting to show how fast the variance of a particle location fades. Formally, we are interested in evaluation of the particle location variance convergence time. Below, d_t denotes variance of particle location in time t, that is

$$d_t = Var[x_t] = m_t - e_t^2. \qquad (27)$$

Definition 5.1 (The particle location variance convergence time). *Let δ be a given positive number. The particle location variance convergence time $pvct(\delta)$ is the minimal number of steps necessary to get the variance of particle location lower than δ for all subsequent time steps, that is*

$$pvct(\delta) = min\{t:\ d_s < \delta\ for\ all\ s \geq t\}. \qquad (28)$$

Empirical evaluation of *pvct* is difficult, so, we introduce the less restrictive measure, that is, a particle location variance weak convergence time.

Definition 5.2 (The particle location variance weak convergence time). *Let δ be a given positive number. The particle location variance weak convergence time $pvwct(\delta)$ is the minimal number of steps necessary to* *get the variance of particle location lower than* δ*, that is*

$$pvwct(\delta) = min\{t: \ d_t < \delta\}. \tag{29}$$

As in the case of $pwcet(\delta)$ and $pwct(\delta)$ it is also obvious that $pvwct(\delta) \le pvct(\delta)$ and equality generally does not hold.

When $pvwct(\delta)$ has to be calculated according to Def. 5.2, it is important to select appropriately initial values of the algorithm parameters: h_1 and m_1. To do this, lets first note that Eq. (1) can be converted to the form:

$$\begin{cases} v_{t+1} = w \cdot v_t + \varphi_t(y - x_t), \\ x_{t+1} = x_t + v_{t+1}. \end{cases} \tag{30}$$

When we substitute zero for t in Eq. (30) we obtain Eq. (31):

$$x_1 = x_0 + w \cdot v_0 + \varphi_0(y - x_0). \tag{31}$$

Let us assume, that x_0 and v_0 are independent random variables. Applying the expectation operator to both sides of Eq. (31) we get

$$e_1 = e_0(1-f) + w \cdot s_0 + fy, \tag{32}$$

where $s_0 = Ev_0$. From Eq. (32) we obtain

$$s_0 = \frac{e_1 - e_0(1-f) - fy}{w}. \tag{33}$$

Multiplying both sides of Eq. (31) by x_0 we get

$$x_1 x_0 = x_0^2(1 - \varphi_0) + w x_0 v_0 + x_0 \varphi_0 y. \tag{34}$$

Applying expectation operator to both sides of the Eq. (34) we obtain

$$h_1 = m_0(1-f) + w e_0 s_0 + e_0 f y, \tag{35}$$

and substituting expression from Eq. (33) for s_0

$$h_1 = m_0(1-f) + e_0(e_1 - e_0(1-f) - fy) + e_0 f y. \tag{36}$$

Eventually, above formula can be simplified to the form

$$h_1 = (m_0 - e_0^2)(1-f) + e_0 e_1, \tag{37}$$

or equivalent

$$h_1 = d_0(1-f) + e_0 e_1. \tag{38}$$

Next, we raise both sides of Eq. (31) to the second power and obtain

$$\begin{aligned} x_1^2 =& x_0^2(1-\varphi_0)^2 + w^2 v_0^2 + \varphi_0^2 y^2 \\ &+ 2w x_0(1-\varphi_0) v_0 + 2x_0(1-\varphi_0) v_0 \\ &+ 2x_0(1-\varphi_0)\varphi_0 y + 2w v_0 \varphi_0 y. \end{aligned} \tag{39}$$

Applying the expectation operator to both sides of Eq.(39) and because of the statistical independence of x_0, φ_0 and v_0 we get

$$\begin{aligned} m_1 =& m_0(1-2f+g) + w^2 s_2 + g y^2 + 2w e_0(1-f) s_0 \\ &+ 2e_0(f-g)y + 2w s_0 f y, \end{aligned} \tag{40}$$

where $s_2 = Ev_0^2$. Expression from Eq. (33) can be substituted for s_0 in Eq. (40). This way we obtain

$$\begin{aligned} m_1 =& m_0(1-2f+g) + w^2 s_2 + g y^2 \\ &+ 2(e_1 - e_0(1-f) - fy)(e_0(1-f) + fy). \end{aligned} \tag{41}$$

Let $d_0 = Var[x_0]$ and $l_o = Var[v_0]$ are given. Then we can calculate

$$m_0 = e_0^2 + d_0$$

and

$$s_2 = s_0^2 + l_0,$$

what can be written in view of Eq. (33) as

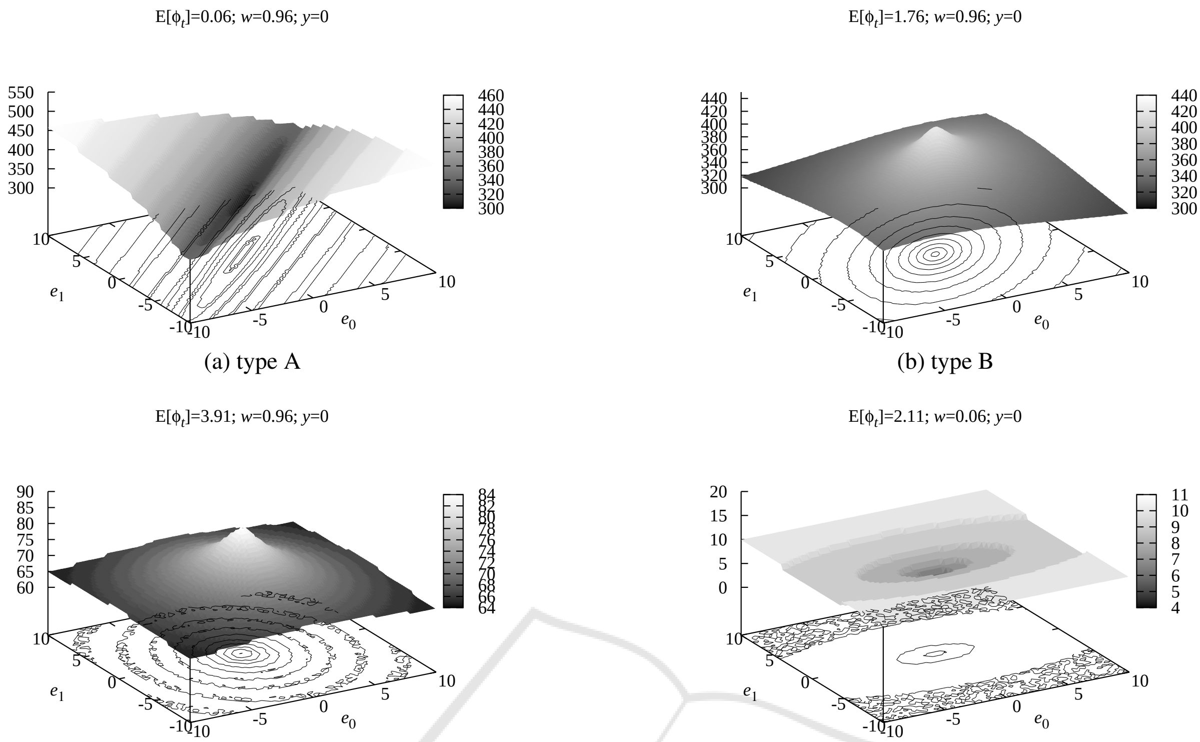
$$s_2 = \frac{(e_1 - e_0(1-f) - fy)^2}{w^2} + l_0. \tag{42}$$

Expression from Eq. (42) can be substituted for s_2 in Eq. (41). This way one can obtain the final version of equation for m_1:

$$\begin{aligned} m_1 =& m_0(1-2f+g) + w^2 l_0 + g y^2 \\ &+ e_1^2 - (e_0(1-f) + fy)^2. \end{aligned} \tag{43}$$

Algorithm 2: Particle location variance weak convergence time evaluation procedure.

```
Initialize: T_max = 1e+5, two successive expected
    locations e_0 and e_1, variance of initial location
    and velocity, for example, d_0 = 0 and l_0 = 1 re-
    spectively, and an attractor of a particle, for ex-
    ample, y = 0.
f = (c_1 + c_2)/2;
g = (c_1)^2/12 + (c_2)^2/12 + ((c_1 + c_2)/2)^2;
m_0 = e_0^2 + d_0.
m_1 = m_0(1 - 2f + g) + w^2 l_0 + g y^2 + e_1^2 - (e_0(1 -
    f) + fy)^2.
h_1 = d_0(1 - f) + e_0 e_1.
d_1 = m_1 - e_1^2.
t = 1
repeat
    h_{t+1} = (1 + w - f) m_t - w h_t + f y e_t
    e_{t+1} = (1 + w - f) e_t - w e_{t-1} + f y
    m_{t+1} = m_t((1 + w)^2 - 2(1 + w) f + g) +
    m_{t-1} w^2 - 2 h_t w (1 + w - f) + 2 e_t y (f(1 + w) - g) -
    2 e_{t-1} w y f + y^2 g
    d_{t+1} = m_{t+1} - e_{t+1}^2
    t = t + 1
until (d_t > δ) ∧ (d_t < 1e+10) ∧ (t < T_max)
if d_t < 1e+10 then
    return t
else
    return T_max
end if
```



(a) type A (b) type B (c) type C (d) type D

Figure 5: Graphs of recorded values of the particle location variance $pvwct(E[\phi_t], w)$ for selected configurations $(E[\phi_t], w)$.

Empirical characteristics of the particle location variance weak convergence time (*pvwct*) are given in Fig. 4 and Fig. 5.

As in the case of empirical characteristics of *pwcet*, Fig. 4 shows the values of *pvwct* also obtained for a grid of configurations $(\phi_{\max}, w)$ starting from $[\phi_{\max} = 0.0,\ w = -1.0]$ and changing with step 0.02 but in both directions (which also gave 200×100 points). The configurations generating $pvwct > 100000$ also have assigned a constant value of 100000 and it is assumed that $c_1 = c_2 = \phi_{\max}/2$.

Fig. 5 presents the values of *pvwct* as a function of e_0 and e_1 where $E[\phi_t]$ and w are fixed. The grid of pairs $[e_0, e_1]$ consists of 40000 points (200×200) varying from -10 to 10 for both e_0 and e_1.

The characteristics depicted in Fig. 4 and Fig. 5 were obtained with Algorithm 2 for selected values of variance of initial location $d_0 = 0$ and velocity $l_0 = 1$.

6 CONCLUSIONS

In the presented research for the stochastic model of PSO with inertia weight we propose new measures inspired by the measure of particle convergence time earlier defined for the deterministic model of PSO. The proposed measures are based on the order-1 and order-2 analysis of PSO dynamics.

The order-1 equivalent of particle convergence time (*pct*) is the particle convergence expected time $pcet(\delta)$ which represents the minimal number of steps necessary for the expected particle location to obtain equilibrium. As in the deterministic case, the upper bound formula ($pcetb(\delta)$) is also derived.

For the order-2 analysis of the PSO model the particle location variance convergence time $pvct(\delta)$ is proposed as a minimal number of steps necessary to get variance of particle location lower than δ for all subsequent time steps.

Weak versions of $pcet(\delta)$ and $pvct(\delta)$, that is, $pwcet(\delta)$ and $pvwct(\delta)$ are also proposed as more convenient for experimental evaluation. Empirical characteristics of $pwcet(\delta)$ and $pvwct(\delta)$ are presented. The issue of appropriate selection of initial parameters for the $pvwct(\delta)$ evaluation procedure is discussed.

ACKNOWLEDGMENTS

The authors would like to thank Prof. Mohammad Reza Bonyadi for inspiration and valuable comments and directions.

REFERENCES

Bonyadi, M. R. and Michalewicz, Z. (2016a). Analysis of stability, local convergence, and transformation sensitivity of a variant of particle swarm optimization algorithm. *IEEE T. Evolut. Comput.*, 20(3):370–385.

Bonyadi, M. R. and Michalewicz, Z. (2016b). Particle swarm optimization for single objective continuous space problems: a review. *Evol. Comput.* Date of Publication: March 8, 2016.

Cleghorn, C. W. and Engelbrecht, A. P. (2014). A generalized theoretical deterministic particle swarm model. *Swarm Intelligence*, 8(1):35–59.

Cleghorn, C. W. and Engelbrecht, A. P. (2015). Particle swarm variants: standardized convergence analysis. *Swarm Intelligence*, 9(2):177–203.

Jiang, M., Luo, Y. P., and Yang, S. Y. (2007). Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm. *Information Processing Letters*, 102(1):8–16.

Lehre, P. K. and Witt, C. (2013). *Finite First Hitting Time Versus Stochastic Convergence in Particle Swarm Optimisation*, pages 1–20. Springer New York.

Liu, Q. (2015). Order-2 stability analysis of particle swarm optimization. *Evol. Comput.*, 23(2):187–216.

Poli, R. (2008a). Analysis of the publications on the applications of particle swarm optimisation. *J. Artif. Evol. App.*, 2008:4:1–4:10.

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

Poli, R. (2009). Mean and variance of the sampling distribution of particle swarm optimizers during stagnation. *IEEE T. Evolut. Comput.*, 13(4):712–721.

Poli, R. and Broomhead, D. (2007). Exact analysis of the sampling distribution for the canonical particle swarm optimiser and its convergence during stagnation. In *GECCO '07: Proceedings of the 9th annual conference on Genetic and evolutionary computation*, volume 1, pages 134–141. ACM Press.

Trelea, I. C. (2003). The particle swarm optimization algorithm: convergence analysis and parameter selection. *Inform. Process. Lett.*, 85(6):317 – 325.

Trojanowski, K. and Kulpa, T. (2015). Particle convergence time in the PSO model with inertia weight. In *Proceedings of the 7th International Joint Conference on Computational Intelligence (IJCCI 2015) - Volume 1: ECTA*, pages 122–130.

van den Bergh, F. and Engelbrecht, A. P. (2006). A study of particle swarm optimization particle trajectories. *Inform. Sciences*, 176(8):937–971.

Witt, C. (2009). Why standard particle swarm optimisers elude a theoretical runtime analysis. In *Proceedings of the Tenth ACM SIGEVO Workshop on Foundations of Genetic Algorithms*, FOGA '09, pages 13–20. ACM.