Noname manuscript No.
(will be inserted by the editor)

A Generalized Theoretical Deterministic Particle Swarm Model

Christopher W. Cleghorn ·
Andries P. Engelbrecht

Received: date / Accepted: date

Abstract A number of theoretical studies of particle swarm optimization (PSO) have been done to gain a better understanding of the dynamics of the algorithm and the behaviour of the particles under different conditions. These theoretical analyses have been performed for both the deterministic PSO model, and more recently for the stochastic model. However, all current theoretical analyses of the PSO algorithm were based on the stagnation assumption, in some form or another. The analysis done under the stagnation assumption is one where the personal best and neighborhood best positions are assumed to be non-changing. While analysis under the stagnation assumption is very informative, it could never provide a complete description of a PSO's behavior. Furthermore, the assumption implicitly removes the notion of a social network structure from the analysis. This paper presents a generalisation to the theoretical deterministic PSO model. Under the generalised model, conditions for particle convergence to a point are derived. The model used in this paper greatly weakens the stagnation assumption, by instead assuming that each particle's personal best and neighborhood best can occupy an arbitrarily large number of unique positions. It was found that the conditions derived in previous theoretical deterministic PSO research could be obtained as a specialisation of the new generalised model proposed. Empirical results are presented to support the theoretical findings.

Keywords Deterministic Particle Swarm Optimization · Theoretical Analysis · Particle Convergence

Christopher W. Cleghorn
Department of Computer Science, University of Pretoria
E-mail: ccleghorn@cs.up.ac.za

Andries P. Engelbrecht
Department of Computer Science, University of Pretoria Tel.: +27 12 420 3578
E-mail: engel@cs.up.ac.za

1 Introduction

In the world of optimization there are many important factors to consider when deciding on the best approach to use. Some of the primary factors are the computational efficiency, the accuracy and the predictability of an approach. At present there are so many optimization approaches available, varying from classical approaches to the more modern computational intelligence approaches, that making an informed decision is difficult. It is for this reason that having a complete understanding about how an optimization technique behaves is imperative. While this fact may seem obvious, there are certain situations where, if this understanding is not available, the approach might never be considered, for example in a situation where human life is at risk. It is for this reason that theoretical work on optimization techniques is of great importance, as it provides the certainty that empirical findings will never be able to.

Particle swarm optimization (PSO) is a well known population-based search algorithm, originally developed by Kennedy and Eberhart (1995). The PSO has been utilized in a variety of application domains, providing a wealth of empirical evidence for its effectiveness as an optimizer. The PSO itself has undergone many alterations subsequent to its inception, some fundamental to the PSO's core behavior, others more application specific. The fundamental alterations to the PSO have to a large extent been a result of theoretical analysis of the PSO's particles' long term trajectory. The most obvious example is the need for velocity clamping in the original PSO. While there were empirical findings that suggested that each particle's velocity was increasing at a rapid rate, it was only once a solid theoretical study was preformed that the reason for the velocity explosion was understood.

The PSO model has undergone a fair amount of theoretical analysis. However, at present all theoretical PSO research is done under the stagnation assumption, in some form or another. An analysis done under the stagnation assumption is one where the personal best and neighbourhood best positions are assumed to be non-changing. While analysis under the stagnation assumption is very informative, it could never provide a complete description of a PSO's behaviour. It is for this reason that a generalisation of the model that allows the particles to occupy more than one personal best and neighbourhood best position during the course of a run is required. Specifically, in the model proposed in this paper, each particle is allowed to occupy an arbitrarily large finite number of unique personal best and neighborhood best positions during the course of a run (referred to as the weak chaotic assumption). The PSO model is investigated from a deterministic perspective, due to the inherent complexity of removing the stagnation assumption in the stochastic context. This complexity is one of the main reasons why the stagnation assumption is still present throughout theoretical PSO research. In an attempt to get to the point of a "grand" PSO model (meaning no simplifying assumptions), a stepping stone is needed, namely the model proposed in this paper. In addition, the study of the deterministic PSO with the weak chaotic assumption makes it possible to directly see what implications the weak chaotic assumption has on PSO's behavior, without the influence of the stochastic component obscuring the underlying behavior. This generalization and its theoretical implications are the focus of this paper.

The primary objectives of this paper can be summarized as follows:

- To justify why a generalization of the theoretical deterministic PSO model is needed.
- To provide a generalization of the theoretical deterministic PSO model that mitigates the old theoretical models' weaknesses, in particular a generalization that caters for the influence of the social network structure.
- To theoretically derive the conditions necessary for particle convergence to a point, under the new generalized model.
- To empirically support the findings of the theoretical derivation, by utilizing parameters sets that illustrate important particle behavior.

The model proposed is one under which an arbitrarily large finite number of unique positions for both personal best and neighbourhood best positions is assumed, as opposed to the commonly used assumption that these positions are fixed. Under the new proposed model the following conditions necessary for particle convergence are derived:

$$|w| < 1, \qquad 0 < \theta_1 + \theta_2 < 4, \qquad w > \frac{(\theta_1 + \theta_2)}{2} - 1, \tag{1}$$

where $\theta_1(t) = c_1 r_1(t)$, $\theta_2(t) = c_2 r_2(t)$, c_1, c_2, and w are the cognitive, social, and inertia weights respectively, and $r_1(t)$ and $r_2(t)$ are uniformly sampled values in the range $[0, 1]$.

The remainder of this paper is structured as follows: Section 2 provides a review of PSO and some of its alterations. A thorough discussion of the state of the theoretical research is also made. Section 3 presents a generalisation of the deterministic PSO model, and conditions for particle convergence to a point are derived under the generalised model. Section 4 presents an experiment designed to illustrate the convergence properties of the deterministic PSO under various social network structures in relation to the derived condition. Section 5 presents a summary of the findings of this paper, as well as a discussion of topics for future research.

2 Background

This section begins with a description of the PSO in subsection 2.1. This is followed by a review of the state of theoretical research in subsection 2.2, considering both the deterministic PSO theoretical results and the stochastic PSO theoretical results.

2.1 Particle swarm optimization

Particle swarm optimization (PSO) is a stochastic population-based search algorithm. The algorithm was originally developed to simulate the complex movement of bird flocks. The algorithm has however evolved into a simple and efficient optimization algorithm. The abstract premise of the algorithm is to have particles move through a potentially multi-dimensional search space in an attempt to find an optimal region. Each particle position represents a candidate solution. Each particle's movement is guided by its own experience in conjunction with that of its neighbours.

More technically, let $f : \mathbb{R}^k \rightarrow \mathbb{R}$ be the objective function that the PSO aims to find an optimum for. For the sake of simplicity, a minimization problem is assumed from this point forward. Let $\Omega(t)$ be a set of N particles in $\mathbb{R}^k$ at a discrete time step t. Then $\Omega(t)$ is said to be the particle swarm at time t. The position $\mathbf{x}_i$ of particle i, is updated using

$$\boldsymbol{x}_i(t+1) = \boldsymbol{x}_i(t) + \boldsymbol{v}_i(t+1) \tag{2}$$

where the velocity update, $\boldsymbol{v}_i(t+1)$, is defined as

$$\boldsymbol{v}_i(t+1) = \boldsymbol{v}_i(t) + c_1\boldsymbol{r}_1(t)(\boldsymbol{y}_i(t) - \boldsymbol{x}_i(t)) + c_2\boldsymbol{r}_2(t)(\hat{\boldsymbol{y}}_i(t) - \boldsymbol{x}_i(t)), \tag{3}$$

where $\boldsymbol{r}_1(t), \boldsymbol{r}_2(t) \sim U(0,1)^k$ for all t, where $U(0,1)$ returns a value from a uniform distribution in the range $[0,1]$; these random values introduce the stochastic aspect to PSO. The position $\boldsymbol{y}_i(t)$ represents the "best" position that particle i has visited, where "best" means the location where the particle has obtained the lowest objective function evaluation. The position $\hat{\boldsymbol{y}}_i(t)$ represents the "best" position that the particles in the neighbourhood of the i-th particle have visited. For a detailed overview of neighbourhoods and their impact on PSO performance the reader is referred to (Kennedy, 1999; Kennedy and Mendes, 2002; Peer et al., 2003; Engelbrecht, 2013a)

The velocity update consists of three components:

- **Inertia component**, $\boldsymbol{v}_i(t)$: The inertia component serves as the memory of the direction of travel in the particle's immediate past. The component prevents a particle's trajectory from wildly changing, much like a bird would be unable to perform a perfect ninety degree turn mid flight. The inertia often assists in pushing a particle out of a local optimum.
- **Cognitive component**, $c_1\boldsymbol{r}_1(t)(\boldsymbol{y}_i(t)-\boldsymbol{x}_i(t))$: The cognitive component serves as the memory of the best position by the particle thus far. The cognitive component causes particles to be drawn back to its best position found, much like a human's tendency to revisit places with which a positive association exists. The coefficient c_1 is referred to as the cognitive weight.
- **Social component**, $c_2\boldsymbol{r}_2(t)(\hat{\boldsymbol{y}}_i(t) - \boldsymbol{x}_i(t))$: The social component serves as the memory of the best position found by neighbouring particles (including itself). The component causes the particle to be drawn to a position where the most success was achieved within the neighbourhood. The behaviour is akin to individuals seeking to fit the group norm of their neighbourhood. The coefficient c_2 is referred to as the social weight.

The PSO algorithm using synchronous updates (Carlisle and Dozier, 2001; Engelbrecht, 2013b) is summarized in algorithm 1.

Shi and Eberhart (1998) proposed the inclusion of the inertia weight in the velocity update equation. The aim of the inclusion was two fold: Firstly, as a means of controlling the exploration and exploitation ability of the swarm, and secondly as a simple way of removing the need for velocity clamping (Eberhart et al., 1996). The inertia coefficient w is applied to the previous velocity of a particle, changing velocity update equation (3) to

$$\boldsymbol{v}_i(t+1) = w\boldsymbol{v}_i(t) + c_1\boldsymbol{r}_1(t)(\boldsymbol{y}_i(t) - \boldsymbol{x}_i(t)) + c_2\boldsymbol{r}_2(t)(\hat{\boldsymbol{y}}_i(t) - \boldsymbol{x}_i(t)) \tag{4}$$

where $w \in \mathbb{R}$; usually a positive value is used.

Algorithm 1 Standard PSO algorithm

```
Create and initialize a k-dimensional swarm, Ω(0), of N particles uniformly within a pre-
defined hypercube.
Let f be the objective function.
Let y_i represent the personal best position of particle i, initialized to x_i(0).
Let ŷ_i represent the neighbourhood best position of particle i, initialized to x_i(0).
Initialize v_i(0) to 0.
repeat
    for all particles i = 1, ··· , N do
        if f(x_i) < f(y_i) then
            y_i = x_i
        end if
        for all particles î that have particle i in their neighborhood do
            if f(y_i) < f(ŷ_î) then
                ŷ_î = y_i
            end if
        end for
    end for
    for all particles i = 1, ··· , N do
        update the velocity of particle i using equation (3)
        update the position of particle i using equation (2)
    end for
until stopping condition is met
```

The inclusion of the inertia weight successfully provides a simple control mechanism for the exploration and exploitation ability of the swarm. However, its secondary goal was not completely met. The need for velocity clamping was unfortunately still required in some cases (Van den Bergh, 2002; Van den Bergh and Engelbrecht, 2006; Trelea, 2003). It was only later, when theoretical studies were done by Van den Bergh and Engelbrecht (Van den Bergh, 2002; Van den Bergh and Engelbrecht, 2006) and Trelea (2003), that it was ascertained that velocity clamping was not required under specified coefficient choices. The details of this analysis are discussed in subsection 2.2. There exist many variations of the concept of an inertia coefficient, many of which alter the value of w dynamically as the time t changes (Peng et al., 2000; Naka et al., 2001; Ratnaweera et al., 2003; Peram et al., 2003; Venter and Sobieszczanski-Sobieski, 2003; Shi and Eberhart, 2001; Zheng et al., 2003), the details of which are beyond the scope of this paper. The use of velocity update equation (4) is referred to as the standard PSO in this paper.

2.2 The state of current theoretical research on particle swarm optimization

Theoretical research of the PSO has evolved since the original study done by Ozcan and Mohan (1998). The evolution has been a process of modelling ever more accurate simplifications of the actual PSO.

The idea of theoretically modelling an optimization algorithm is a relatively new phenomenon, in contrast to modelling a natural system. Historically, a model of a complex natural system is made such that the model fits the data that the system produces. The model of a natural system is almost by definition an approximation of the system, largely due to the fact that the exact underlying behaviour is somewhat unknown. However, in the case of modelling an algorithm, the exact model is already known (i.e., the algorithm itself). The problem, however, is that the non-simplified PSO model is too complex to derive meaningful results from. In the case of PSO, there are two primary complexities in the model: The inherent coupling between successive updates as a result of the non-fixed nature of $\boldsymbol{y}_i$ and $\hat{\boldsymbol{y}}_i$, and the influence of the stochastic component.

There are many aspects of PSO particle behaviour that are of interest from a theoretical and practical perspective. In this paper, we focus on convergence results, specifically, under what conditions will convergence of each particle in the swarm to a point be guaranteed. To clarify, convergence of a particle is defined as follows:

Definition 2.1. ***Particle convergence****: Let the dimension of the search space be k. Then, particle i is said to converge if there exists a $\boldsymbol{p} \in \mathbb{R}^k$ such that for every $\epsilon > 0$ there exists a $\hat{t}(\epsilon)$ such that if $t > \hat{t}(\epsilon)$, then*

$$\|\boldsymbol{x}_i(t) - \boldsymbol{p}\| \leq \epsilon, \tag{5}$$

where $\|\bullet\|$ is an arbitrary norm on $\mathbb{R}^k$.

The definition of convergence in definition 2.1 should not be confused with the convergence of a particle to an optimum. Rather, it indicates that the particle has come to rest at a point in the search space.

It is important to note that finding or deriving conditions that guarantee particle convergence is fundamental to the effective use of a PSO. From a practical perspective, the point to which a PSO particle will converge is not really that important, but rather that convergence will occur. This is particularly true when trying to avoid particle velocity explosion, as a guarantee on convergence is equally a guarantee that velocity explosion will not occur. The impact of convergence results is not only on the extremes of PSO behaviour, but in all scenarios, as the more information and guarantees about PSO behaviour becomes available, the better the PSO can be used, resulting in better performance.

The current theoretical research on PSO can be split into two sub-branches: The first is where the stochastic component is assumed to be fixed, resulting in a deterministic PSO, as discussed in subsection 2.2.1. The second is where this assumption is dropped, as discussed in subsection 2.2.2.

2.2.1 Deterministic particle swarm model

The first analysis of the PSO was preformed by Ozcan and Mohan (1998) on a very simplified version of the PSO excluding inertia weights. The analysis was performed in 1-dimension on the "best" particle's trajectory. This "best" particle i is one where the position of the personal best $\boldsymbol{y}_i(t)$ and that of the neighbourhood best position $\hat{\boldsymbol{y}}_i(t)$ are assumed to be equal for all t. There were two additional assumptions made in the analysis: The stochastic components of the update equations were assumed to be fixed, specifically:

Assumption 2.1. ***Deterministic assumption****: It is assumed that $\boldsymbol{\theta}_1 = \boldsymbol{\theta}_1(t) = c_1\boldsymbol{r}_1(t)$, and $\boldsymbol{\theta}_2 = \boldsymbol{\theta}_2(t) = c_2\boldsymbol{r}_2(t)$, for all t.*

The second assumption is that the personal best position remains fixed for all t, specifically:

Assumption 2.2. ***Best particle stagnation assumption:*** *It is assumed that* $\boldsymbol{y}_i(t) = \hat{\boldsymbol{y}}_i(t) = \boldsymbol{y}_i$*, for all* t.

Giving these assumptions and dropping the particle index, the update equations are reformulated as

$$x\,(t+1) = x\,(t) + v\,(t+1)\,, \tag{6}$$

$$v\,(t+1) = v\,(t) - \theta x\,(t) + \theta y, \tag{7}$$

where $\theta = \theta_1 + \theta_2$. This was then reformulated into the recurrence relation,

$$x\,(t+1) = (2-\theta)x(t) - x(t-1) + \theta y, \tag{8}$$

with initial conditions $x\,(0) = x_0$, $v\,(0) = v_0$ and $x\,(1) = x_0(1-\theta) + v_0 + \theta y$. From this recurrence relation the closed form of $x(t)$ is obtained, as

$$x(t) = \alpha\left(\frac{2-\theta+\delta}{2}\right)^t + \beta\left(\frac{2-\theta-\delta}{2}\right)^t + y, \tag{9}$$

where

$$\delta = \sqrt{\theta^2 - 4\theta}, \tag{10}$$

$$\beta = (x_0 - y)\,(\delta+\theta)\,2\delta - \frac{v_0}{\delta}, \tag{11}$$

$$\alpha = x_0 - y - \beta. \tag{12}$$

The analysis of the long term trajectory of a particle under equation (9) was split into two groups of cases: One where δ is real, the other where it is complex. This analysis was the first to give an idea of the trajectories of a particle in a PSO, and provided a theoretical bases for intelligent parameter selection. The exact cases are discussed in (Ozcan and Mohan, 1998).

The fundamental issue with the original research was that the “best” particle is not of practical interest, as the selection of which particle is “best” is very unlikely to remain constant through a whole run. This results in a poor approximation of the actual PSO, as the model effectively provides a prediction on the trajectory of a particle that is unlikely to exist in practice. This issue was subsequently addressed again by Ozcan and Mohan in a later paper (Ozcan and Mohan, 1999), where the trajectory of a “general” particle is considered. The assumptions remained the same as in their 1998 paper; however, $\boldsymbol{y}_i(t)$ and $\hat{\boldsymbol{y}}_i(t)$ were not assumed equal for all t. Instead, the best particle stagnation assumption was replaced with:

Assumption 2.3. ***Stagnation assumption:*** *It is assumed that* $\boldsymbol{y}_i(t) = \boldsymbol{y}_i$*, and* $\hat{\boldsymbol{y}}_i(t) = \hat{\boldsymbol{y}}_i$*, for all* t.

The analysis was also done in $\mathbb{R}^k$. However, it is worth noting that, because of the assumption that $\boldsymbol{y}_i(t)$ and $\hat{\boldsymbol{y}}_i(t)$ are fixed, an analysis in $\mathbb{R}^k$ and $\mathbb{R}$ are equivalent as the assumption inherently removes the coupling between dimensions. The paper derives a new recurrence relation,

$$x(t) = 2\,(2-\theta_1-\theta_2)\,x(t-1) - x(t-2) + \theta_1 y + \theta_2\hat{y}, \tag{13}$$

where $y(t) = y$, and $\hat{y}(t) = \hat{y}$ for all t, with the initial conditions $x(0) = x_0$, $v(0) = v_0$, and $x(1) = x_0(1 - \theta_1 - \theta_2) + v_0 + \theta_1 y + \theta_2 \hat{y}$. Then, using methods similar to that of (Ozcan and Mohan, 1998), the following closed form is derived:

$$x(t) = \eta\alpha^t + \iota\beta^t + \zeta, \tag{14}$$

where

$$\delta = \sqrt{(2 - \theta_1 - \theta_2)^2 - 4}, \tag{15}$$

$$\alpha = \frac{(2 - \theta_1 - \theta_2 + \delta)}{2}, \tag{16}$$

$$\beta = \frac{(2 - \theta_1 - \theta_2 - \delta)}{2}, \tag{17}$$

$$\eta = \left(\frac{1}{2} - \frac{\theta_1 + \theta_2}{2\delta}\right) x_0 + \frac{v_0}{\delta} + \frac{\theta_1 y + \theta_2 \hat{y}}{2\delta} - \frac{\theta_1 y + \theta_2 \hat{y}}{2\theta_1 + 2\theta_2}, \tag{18}$$

$$\iota = \left(\frac{1}{2} + \frac{\theta_1 + \theta_2}{2\delta}\right) x_0 - \frac{v_0}{\delta} - \frac{\theta_1 y + \theta_2 \hat{y}}{2\delta} + \frac{\theta_1 y + \theta_2 \hat{y}}{2\theta_1 + 2\theta_2}, \tag{19}$$

$$\zeta = \frac{\theta_1 y + \theta_2 \hat{y}}{\theta_1 + \theta_2}. \tag{20}$$

The analysis of particle trajectory is once again split into two groups of cases, one for δ real, and the other for δ complex. The exact cases are discussed in Ozcan and Mohan (1999).

While both (Ozcan and Mohan, 1998) and (Ozcan and Mohan, 1999) provided a good overview of the behaviour of the PSO trajectory, they did not provide the necessary or sufficient conditions on the PSO's coefficients which would ensure convergence. This is not surprising, as it was later theoretically proven by Clerc and Kennedy (2002) that the PSO using update equations (2) and (3) could not actually be guaranteed to converge under any coefficient choice, even if velocity clamping is used. This result led to the development of constriction coefficients (Clerc, 1999; Clerc and Kennedy, 2002), in an attempt to guarantee particle convergence. Even though the application of constriction coefficients has become less prevalent in favour of simply including the inertia weight, its development marks the first practical use of theoretical results with regards to PSO. The analysis done by Clerc and Kennedy (2002) utilized the same assumptions as the 1999 paper by Ozcan and Mohan (1999).

The theoretical research focus shifted to PSO models that included the inertia weight. There are three primary papers that address the deterministic PSO model with the inclusion of the inertia weight (Zheng et al., 2003; Van den Bergh and Engelbrecht, 2006; Trelea, 2003). All of this research was done utilizing the same assumptions as the paper of Ozcan and Mohan (1999). The work done by Zheng et al. (2003) was one of the first to study the convergence properties of the PSO with inertia weights. The study looked at a multitude of possible cases where convergence would or would not occur. However, the results did not provide a general requirement on the coefficients which would guarantee convergence. The work of Van den Bergh and Engelbrecht (2006), Van den Bergh (2002), and Trelea (2003) provided a more complete analysis than the findings of Zheng et al. (2003) by ascertaining general conditions on the PSO's coefficients under which particle convergence would be guaranteed.

The approach utilized by Van den Bergh (2002) mirrors that of Ozcan and Mohan (1999) in a more complex setting. Utilizing update equations (2) and (4), the second order recurrence relation

$$x\left(t+1\right)=\left(1-w-\theta_1-\theta_2\right)x\left(t\right)-wx\left(t-1\right)+\theta_1 y+\theta_2\hat{y} \tag{21}$$

is derived, with initial conditions defined as $x\left(0\right)=x_0$, $x\left(1\right)=x_1$, and $x\left(2\right)=\left(1+w-\theta_1-\theta_2\right)x_1-wx_0+\theta_1 y+\theta_2\hat{y}$. The following closed form is then derived:

$$x\left(t\right)=k_1+k_2\tau_1^t+k_3\tau_2^2, \tag{22}$$

where

$$\tau_1=\frac{1+w-\theta_1-\theta_2+\gamma}{2}, \tag{23}$$

$$\tau_1=\frac{1+w-\theta_1-\theta_2-\gamma}{2}, \tag{24}$$

$$\gamma=\sqrt{\left(1+w-\theta_1-\theta_2\right)^2-4w}, \tag{25}$$

$$k_1=\frac{\theta_1 y+\theta_2\hat{y}}{\theta_1+\theta_2}, \tag{26}$$

$$k_2=\frac{\tau_2\left(x_0-x_1\right)-x_1+x_2}{\gamma\left(\tau_1-1\right)}, \tag{27}$$

$$k_3=\frac{\tau_1\left(x_1-x_0\right)+x_1-x_2}{\gamma\left(\tau_2-1\right)}. \tag{28}$$

An analysis of the long term behaviour of the closed form in equation (22) yielded the conditions required to guarantee convergence, specifically,

$$0<w<1, \qquad c_1>0, \qquad c_2>0, \tag{29}$$

$$w>\frac{c_1+c_2}{2}-1. \tag{30}$$

The approach utilized by Trelea (2003) deviates from the approach used in earlier theoretical PSO research in that no closed form is derived. The analysis begins with fixing

$$r_1=r_2=\frac{1}{2}, \tag{31}$$

and by defining

$$\hat{\theta}=\frac{\theta_1+\theta_2}{2}, \tag{32}$$

$$p=\frac{\theta_1}{\theta_1+\theta_2}y+\frac{\theta_2}{\theta_1+\theta_2}\hat{y}. \tag{33}$$

This results in the update equations

$$v\left(t+1\right)=wv\left(t\right)+\hat{\theta}\left(p-x\left(t\right)\right) \tag{34}$$

and

$$x\left(t+1\right)=x\left(t\right)+v\left(t+1\right). \tag{35}$$

Equations (35) and (34) can be combined into the compact form

$$\begin{vmatrix} x(t+1) \\ v(t+1) \end{vmatrix} = M \begin{vmatrix} x(t) \\ v(t) \end{vmatrix} + \begin{vmatrix} \frac{c_1+c_2}{2} \\ \frac{c_1+c_2}{2} \end{vmatrix} p, \quad (36)$$

where

$$M = \begin{vmatrix} 1-\hat{\theta} & w \\ -\hat{\theta} & w \end{vmatrix}. \quad (37)$$

Trelea realized that the eigenvalues of the matrix M are of fundamental importance when considering the convergence behaviour of a particle. Specifically, the magnitude of the eigenvalues of M must be less than 1 in order to guarantee convergence. This restriction leads to the following conditions for convergence:

$$0 < w < 1, \qquad c_1 > 0, \qquad c_2 > 0, \quad (38)$$

$$w > \frac{c_1 + c_2}{4} - 1. \quad (39)$$

At first glance there is a discrepancy between the results of Trelea (2003) and that of Van den Bergh and Engelbrecht (2006). The difference between equations (30) and (39) results from a difference in the treatment of the stochastic component. Specifically, the conditions of Van den Bergh and Engelbrecht are more restrictive than those of Trelea's. This is seen in figure 1 where $c_1 + c_2$ is plotted against w, with the coloured region implying convergence of particles to a point. The reason for the difference is that the conditions of Van den Bergh and Engelbrecht cater for the fact that r_1 and r_2 could in fact equal 1 despite their expected value being 0.5, while Trelea's conditions are strictly derived with $r_1 = r_2 = 0.5$. This means that the conditions in equation (30) are a guarantee that every iteration will encourage convergence, whereas the conditions in equation (39) only guarantee that there is at worst a 50 percent chance that any given iteration will encourage convergence.

2.2.2 Stochastic particle swarm model

The theoretical research has at present shifted to almost exclusively analysing PSO convergence with the re-inclusion of the stochastic component. However, the stagnation assumption is still present.

The first analysis that included the stochastic component was done by Kadirkamanathan et al. (2006), where the analysis is performed on a PSO model in 1-dimension on the "best" particles with the inclusion of the inertia weight. The best particle stagnation assumption is made. The analysis of particle trajectory is done by modelling particles as a nonlinear feedback controlled system as formulated by Lure, discussed in (Dosoer and Vidyasagar, 1975; Vidyasagar, 2002). Utilizing the passivity theorem (Vidyasagar, 2002) and Lyapunov stability (Kisacanin and Agarwal, 2001), the following conditions were derived to guarantee asymptotic stability:

$$|w| < 1, \quad w \neq 0, \quad (40)$$

$$c_1 + c_2 < \frac{2\left(1 - 2|w| + w^2\right)}{1 + w}. \quad (41)$$

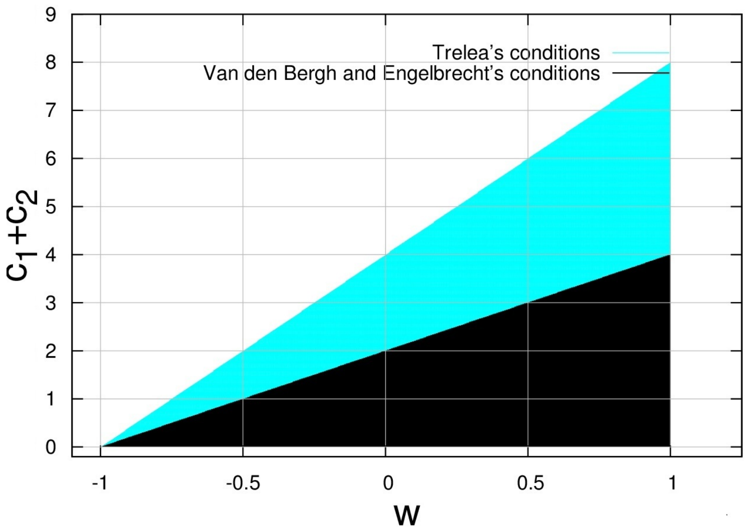

Fig. 1 Parameter regions in which each particle converges to a point, under Van den Bergh and Engelbrecht (2006), and Trelea (2003)'s conditions

While the conditions of equations (40) and (41) do guarantee convergence, the parameter restriction is unfortunately very conservative. The conservativeness is a direct result of utilizing Lyapunov stability, which is inherently a conservative approach, implying that there may in fact be a less strict parameter restriction that will guarantee convergence.

Shortly, after the work of Kadirkamanathan et al. (2006), Jiang et al. (2007) independently provided a more extensive analysis of the stochastic PSO. The analysis is done in both 1-dimension and k-dimensions on an arbitrary particle under a partial stagnation assumption. The partial stagnation being that $\hat{y}$ is held fixed, and y is allowed to vary in a restricted manner. The conditions derived for convergence are as follows:

$$c_1 \geq 0, \quad c_2 \geq 0, \quad 0 \leq w < 1, \tag{42}$$

$$c_1 + c_2 > 0, \quad 0 < h(1) < \frac{c_2^2 (1+w)}{6}, \tag{43}$$

where

$$h(1) = -(c_1+c_2)w^2 + \left(\frac{1}{6}c_1^2 + \frac{1}{6}c_2^2 + \frac{1}{2}c_1c_2\right)w + c_1 + c_2 - \frac{1}{3}c_1^2 - \frac{1}{3}c_2^2 - \frac{1}{2}c_1c_2 \tag{44}$$

It is worth noting that these conditions are not necessary for convergence but only sufficient.

The subsequent research by Martinez et al. (2008) utilized a novel approach by modelling the particle dynamics as being analogous to that of a damped mass-spring system. The analysis is done in 1-dimension under the stagnation assumption. Both a continuous and a discrete version of the mass spring system were considered, and the following condition for convergence was derived:

$$|w| < 1, \quad 0 < r_1(t)c_1 + r_2(t)c_2 < 2(w+1), \quad \forall t. \tag{45}$$

The conditions of equation (45) correspond with those of Van den Bergh and Engelbrecht (2006) in equations (29) and (30), as well as the condition of Trelea (2003) in equations (38) and (39). However, the range of coefficients allowed by the conditions of equation (45) are wider, due to the inclusion of possible negative c_1, c_2, and w.

There has been further research done under alterations of the definitions of stability, e.g., first order and second order stability (details can be found in (Poli and Broomhead, 2007; Poli, 2009)). However, no change to the conditions of equation (45) was needed.

The most recent contribution to the theoretical PSO research was by Gazi (2012), where an analysis in k-dimensions is made on the stochastic PSO dynamics of a best and non-best particle under the stagnation assumption. The work extends the method of Kadirkamanathan et al. (2006) by finding conditions less conservative than those of the standard Lyapunov stability. The following conditions were derived:

$$|w| < 1, \quad 0 < r_1(t)c_1 + r_2(t)c_2 < \frac{24\left(1 - 2\,|w| + w^2\right)}{7\,(1+w)}, \quad \forall t. \tag{46}$$

In the next section, issues with the current theoretical models' assumptions are discussed in detail. A new generalized theoretical deterministic PSO model is proposed, under which the conditions necessary for particles to converge to a point are derived and discussed.

3 Theoretical analysis

This section provides a detailed theoretical analysis of the deterministic PSO, with a justification for the theoretical model generalization. Conditions for particle convergence to a point are provided under this generalization.

The discussion begins in subsection 3.1 with a justification for why the theoretical model needs an extension. This is followed by subsection 3.2 with a brief overview of some elementary mathematical definitions and theorems that will be used in subsection 3.3. Subsection 3.2 may be used as a reference and need not be read first. Subsection 3.3 contains a detailed proof of the PSO particle convergence under the weak chaotic assumption, and an arbitrary social network structure. This is followed by a discussion of the coefficient space where particle convergence will occur in subsection 3.4, as derived from the results of subsection 3.3.

3.1 Generalization of the current theoretical particle swarm model

The theoretical studies of PSO convergence have evolved greatly over the last decade, and are discussed in great detail in subsection 2.2. Despite the great advances made, a truly accurate analysis of the non-simplified PSO has not yet been made from a theoretical standpoint. This is not surprising as the true PSO contains a large amount of coupling between particles and iterations, making a true model incredibly difficult to work with. All of the theoretical results available to date were derived under the stagnation assumption, with the closest to an exception being the work of Jiang et al. (2007) as discussed in subsection 2.2.2, where a slightly weakened version of stagnation is considered.

The stagnation assumption is the last remaining simplification still present in the current theoretical PSO models. In many regards it is the most simplifying of all the assumptions. If the "neighbourhood" best $\hat{y}$ of particle i is assumed fixed, it actually decouples all particles in the swarm from each other. This decoupling actually removes the notion of a swarm from the analysis. It has occasionally been stated that theoretical analyses thus far were done in the setting of a star topology (Van den Bergh and Engelbrecht, 2006; Van den Bergh, 2002; Jiang et al., 2007) (that is using a Gbest PSO). However, the notion of a social network structure is completely meaningless if $\hat{y}_i$ is fixed for each particle. Furthermore, the stagnation assumption removes the true memory that each particle has, specifically if the personal best y of particle i is assumed to be fixed, the "memory" is then fixed. It is clear that a fixed "memory" is a very poor approximation of an actual PSO, where the memory could potentially change after each iteration.

The stagnation assumption in many ways is unrealistic; most fundamentally, a stagnated PSO is no longer an optimizer. What this implies is that even if convergence of a PSO is proved in a stagnated PSO, the convergence of the particles has little practical implication, since if a PSO was in a true stagnant state subsequent iterations of a PSO algorithm would not alter the best located solution of the swarm, or that of an individual particle, making the iterations pointless. Now, if the PSO was guaranteed to enter a stagnant state, the analysis would be more meaningful. However, this has never been proved. Realistically, it might not even be possible to prove that a stagnant state will always occur, other than the case where stagnation and convergence are the same (each particle's position has stagnated, not only the personal and neighbourhood best positions). It is not hard to imagine a situation where stagnation will not occur, other than the case of convergence.

It is clear that, given the implications of the stagnation assumption, a theoretical study of the PSO without the assumption is of great importance. This is the primary issue that this paper addresses. Specifically, the focus is that of the deterministic PSO with the stagnation assumption removed. The assumption is replaced with a substantially weaker one. Instead of assuming that $y_i(t)$ and $\hat{y}_i(t)$ are fixed, the following assumption is made:

Assumption 3.1. ***Weak chaotic assumption**: It is assumed that both $y_i(t)$ and $\hat{y}_i(t)$ will occupy an arbitrarily large finite number of unique positions, ψ_i and $\hat{\psi}_i$, respectively.*

The weak chaotic assumption provides a closer approximation to an actual PSO, as it allows $y_i(t)$ and $\hat{y}_i(t)$ to change over time, an aspect of PSO that the stagnation assumption completely removes. The term "chaotic" is used to emphasize that $y_i(t)$ and $\hat{y}_i(t)$ can change in a seemingly chaotic manner. The term "weak" is used to emphasize that, while the positions of $y_i(t)$ and $\hat{y}_i(t)$ can change in a seemingly chaotic manner, it is not completely chaotic, as in reality $y_i(t)$ and $\hat{y}_i(t)$ could potentially occupy an infinite number of unique positions. It should be noted that from a mathematical perspective "weak" and "strong" chaos are quite different. However, computationally they are equivalent. This equivalence is due to the fixed level of precision and range on all modern computers.

Working under this assumption, the theoretical analysis provides a meaningful step towards the "grand" PSO model, with "grand" meaning no assumptions. The deterministic PSO model is the focus of this paper, so as to directly understand the implication of the weak chaotic assumption on the PSO's particle's behaviour. Future research should expand on this study to remove the stagnation assumption for the stochastic PSO model as depicted in figure 2. The detailed convergence analysis done under the weak chaotic assumption is presented in subsection 3.3.

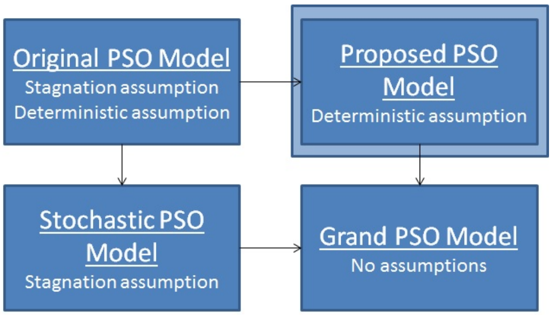

Fig. 2 Theoretical PSO models

3.2 Preliminary concepts

For the sake of completeness, a few elementary definitions and theorems are given in this section.

Definition 3.1. ***Metric space.***
A metric space is a pair (X, d), where X is a set and d is a metric on X. The metric is defined on $X \times X$ such that $\forall x, y, z \in X$ the following hold:

1. $d(x, y) \geq 0$, *d is real valued and finite.*
2. $d(x, y) = 0 \iff x = y$.
3. $d(x, y) = d(y, x)$.
4. $d(x, y) \leq d(x, z) + d(z, y)$.

Definition 3.2. ***Cauchy sequence.***
A sequence (x_n) in a metric space (X, d) is Cauchy if $\forall \epsilon > 0 \ \exists$ a $G(\epsilon) \in \mathbb{N}$, such that $d(x_m, x_n) < \epsilon \ \forall m, n > G(\epsilon)$.

Definition 3.3. ***Completeness.***
A metric space (X, d) is complete if every Cauchy sequence in X converges to a point within X.

Definition 3.4. ***Normed space.***
A normed space $(X, \|\bullet\|)$ is a vector space with a norm defined on it. A norm on a real valued vector space X is a real valued function defined such that $\forall x, y \in X$ and $\forall \alpha \in \mathbb{R}$ the following hold:

1. $\|x\| \geq 0$

2. $\|x\| = 0 \iff x = 0$
3. $\|\alpha x\| = |\alpha| \, \|x\|$
4. $\|x + y\| \leq \|x\| + \|y\|$

Also, a norm on X defines the metric d on X given by $d(x, y) = \|x - y\|$.

Definition 3.5. ***Banach space.***
A Banach space $(X, \|\bullet\|)$ is a complete normed space, where completeness is defined in definition 3.3.

Definition 3.6. ***Spectral radius.***
The spectral radius of a matrix A is defined as $\rho(A) = \max_{i=1,\cdots,dim(A)} |\lambda_i|$, where λ_i is the ith eigenvalue of A.

Definition 3.7. ***Convergent matrix.***
An $n \times n$ matrix A is said to be convergent if $\lim_{\kappa\to\infty} A^\kappa = \Theta$, where Θ is the zero matrix.

Theorem 3.1. ***Convergence of matrix.***
A is convergent if and only if $\rho(A) < 1$ (Burden and Faires, 2010).

Definition 3.8. ***Affine mapping.***
The mapping $F : \mathbb{R}^m \to \mathbb{R}^n$ is affine if there is a linear function $L : \mathbb{R}^m \to \mathbb{R}^n$ and a vector $\boldsymbol{b} \in \mathbb{R}^n$ such that $F(\boldsymbol{x}) = L(\boldsymbol{x}) + \boldsymbol{b} \; \forall \boldsymbol{x} \in \mathbb{R}^m$.

Definition 3.9. ***Contraction.***
Let $X = (X, d)$ be a metric space. A mapping $T : X \to X$ is called a contraction on X if there is a positive real number $\alpha < 1$ such that $\forall \boldsymbol{x}, \boldsymbol{y} \in X$, $d(Tx, Ty) \leq \alpha d(\boldsymbol{x}, \boldsymbol{y})$.

Theorem 3.2. ***Infinite composition of affine mapping.***
Let $\{G_h : h = 1, \cdots, U\}$ be affine mappings on $\mathbb{R}^k$, where $U \in \mathbb{N}$. If there exists an $s \geq 1$ such that for every $\sigma_1, \cdots, \sigma_s \in \{1, \cdots, U\}$ the composition $G_{\sigma_1} \circ \cdots \circ G_{\sigma_s}$ is a contraction, then for any infinite sequence of the form $\sigma_1 \in \{1, \cdots, U\}$, $\sigma_2 = \sigma_1$ or $\sigma_2 \in \{1, \cdots, U\} \setminus \{\sigma_1\}$, $\sigma_3 = \sigma_2$ or $\sigma_3 \in \{1, \cdots, U\} \setminus (\{\sigma_1\} \cup \{\sigma_2\}) \cdots$ and any $\boldsymbol{z} \in \mathbb{R}^k$, $\lim_{\xi\to\infty} G_{\sigma_1} \circ \cdots \circ G_{\sigma_\xi}(z)$ is convergent (Mate, 1999).

Stated less rigorously: for any arbitrary finite set $\hat{G}$ of affine mappings, if there exists an $s \geq 1$ such that for any s mappings selected from $\hat{G}$ the resulting composition will be a contraction, then for any infinite sequence of mappings matching the form above and selected from $\hat{G}$, the limit of the compositions will map a point in $\mathbb{R}^k$ to a point in $\mathbb{R}^k$.

3.3 Convergence of the particle swarm

It is assumed without loss of generality that the PSO algorithm is applied on $\mathbb{R}$, instead of $\mathbb{R}^k$ (this is a reasonable assumption, as the proof that follows is valid regardless of dimension, as will become clear during the proof). The original update equations become:

$$x_i(t+1) = x_i(t) + v_i(t+1), \tag{47}$$

$$v_i(t+1) = wv_i(t) + c_1 r_1(t)(y_i(t) - x_i(t)) + c_2 r_2(t)(\hat{y}_i(t) - x_i(t)). \tag{48}$$

These update equations can be reformulated as the following second order non-homogeneous non-autonomous recurrence relation:

$$x_i(t+1) = (1+w-(\theta_1(t)+\theta_2(t)))x_t(t) - wx_i(t-1) + \theta_1(t)y_i(t) + \theta_2(t)\hat{y}_i(t) \quad (49)$$

where $\theta_1(t) = c_1 r_1(t)$ and $\theta_2(t) = c_2 r_2(t)$. In the standard PSO model the initial position $x_i(0)$ and velocity $v_i(0)$ are given for $1 < i < N$. In the reformulation of equation (49) this is equivalent to the values of $x_i(0)$ and $x_i(1)$ being given for $1 < i < N$.

To simplify the discussion, the following definition of all PSO's state is used: The state S_t of the PSO algorithm at time step t is defined as containing all particle positions $x_i(t)$, as well as all the best personal positions $y_i(t)$, in addition to the best position so far within each particle's neighbourhood $\hat{y}_i(t)$ for each $1 < i < N$. More briefly, $S_t = \bigcup_{i=0}^{N} \{(x_i(t), y_i(t), \hat{y}_i(t), i)\}$.

Under the assumption that $x_i(0)$ and $x_i(1)$ for $1 < i < N$ are given, the initial states of the PSO, S_0 and S_1, are completely known.

Update equation (49) can then be reformulated as

$$\begin{vmatrix} x_i(t+1) \\ x_i(t) \end{vmatrix} = \begin{vmatrix} 1+w-(\theta_1(t)+\theta_2(t)) & -w \\ 1 & 0 \end{vmatrix} \begin{vmatrix} x_i(t) \\ x_i(t-1) \end{vmatrix} + \begin{vmatrix} \theta_1(t)y_i(t)+\theta_2(t)\hat{y}_i(t) \\ 0 \end{vmatrix}. \quad (50)$$

It is at this point that the stochastic component is removed by assuming that $\theta_1(t) = \theta_1$, and $\theta_2(t) = \theta_2$ (in effect r_1 and r_2 are fixed), producing the following deterministic form of equation (50):

$$\begin{vmatrix} x_i(t+1) \\ x_i(t) \end{vmatrix} = \begin{vmatrix} 1+w-(\theta_1+\theta_2) & -w \\ 1 & 0 \end{vmatrix} \begin{vmatrix} x_i(t) \\ x_i(t-1) \end{vmatrix} + \begin{vmatrix} \theta_1 y_i(t)+\theta_2\hat{y}_i(t) \\ 0 \end{vmatrix}. \quad (51)$$

This reformulation allows the calculation of the elements of $S(t+1)$ from those of $S(t)$ and $S(t-1)$. Specifically, the state can be calculated by first calculating particle positions using equation (51), and then the remaining information is readily calculated. At this point it is worth noting that the mapping is defined only for a specific particle i, and for a specific $S(t)$ and $S(t-1)$. To make this more clear, equation (51) is rather defined as an operator (it is customary not to use parentheses when working with operators) of the following form:

$$\begin{vmatrix} x_i(t+1) \\ x_i(t) \end{vmatrix} = T^i_{(t+1)} \begin{vmatrix} x_i(t) \\ x_i(t-1) \end{vmatrix} \quad (52)$$

$$= \begin{vmatrix} 1+w-(\theta_1+\theta_2) & -w \\ 1 & 0 \end{vmatrix} \begin{vmatrix} x_i(t) \\ x_i(t-1) \end{vmatrix} + \begin{vmatrix} \theta_1 y_i(t)+\theta_2\hat{y}_i(t) \\ 0 \end{vmatrix}, \quad (53)$$

where $T^i_{(t+1)}$ is an operator applied to $[x_i(t), x_i(t-1)]^T$.

Equation (52) is shortened to

$$T^i_{(t+1)} \begin{vmatrix} x_i(t) \\ x_i(t-1) \end{vmatrix} = A \begin{vmatrix} x_i(t) \\ x_i(t-1) \end{vmatrix} + \boldsymbol{b^i_t}, \quad (54)$$

where

$$A = \begin{vmatrix} 1+w-(\theta_1+\theta_2) & -w \\ 1 & 0 \end{vmatrix} \quad (55)$$

and

$$\boldsymbol{b_t^i} = \begin{vmatrix} \theta_1 y_i(t) + \theta_2 \hat{y}_i(t) \\ 0 \end{vmatrix}. \tag{56}$$

This means that, in order to calculate the position of $x_i(t)$, the operator $T^i_{(t)} \circ T^i_{(t-1)} \circ \ldots \circ T^i_{(3)} \circ T^i_{(2)}$ must be applied to $|x_i(1), x_i(0)|^T$. Let $\|\bullet\|_\infty$ be the max norm on $\mathbb{R}^2$ and let d be the induced metric. Focusing on the actual operators, it is noted that, for arbitrary $\boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^2$,

$$d(T^i_{(t+1)}\boldsymbol{y}, T^i_{(t+1)}\boldsymbol{z}) = \|T^i_{(t+1)}\boldsymbol{y} - T^i_{(t+1)}\boldsymbol{z}\|_\infty = \|A(\boldsymbol{y}-\boldsymbol{z})\|_\infty \le \|A\|_\infty \|\boldsymbol{y}-\boldsymbol{z}\|_\infty \tag{57}$$

If $\|A\|_\infty < 1$, then $T^i_{(t+1)}$ is a contractive mapping, and theorem 3.2 could be utilized with $s = 1$ to prove convergence. Unfortunately, A cannot be a contractive mapping given its form, as

$$\|A\|_\infty = \max\{|1 + w - (\theta_1 + \theta_2)| + |-w|, 1 + 0\}, \tag{58}$$

which clearly cannot be less than 1.

However, since the operator $T^i_{(t+1)}$ is a mapping from $\mathbb{R}^2$ to $\mathbb{R}^2$, there is still a way to utilize theorem 3.2 if the spectral radius $\rho(A)$ is less than 1. In order for $\rho(A) < 1$, $\max\{|\lambda_1|, |\lambda_2|\}$ must be less than 1, where $|\lambda_1|$ and $|\lambda_2|$ are the modulus of the eigenvalues of A. Simple calculation yields the following characteristic polynomial of A:

$$\theta_1\lambda + \theta_2\lambda - w\lambda + w + \lambda^2 - \lambda, \tag{59}$$

from which the following, possibly complex valued eigenvalues are obtained:

$$\lambda_1 = \frac{1}{2}(-\sqrt{(\theta_1 + \theta_2 - w - 1)^2 - 4w} - \theta_1 - \theta_2 + w + 1), \tag{60}$$

$$\lambda_2 = \frac{1}{2}(\sqrt{(\theta_1 + \theta_2 - w - 1)^2 - 4w} - \theta_1 - \theta_2 + w + 1). \tag{61}$$

If PSO parameters are chosen such that $\max\{|\lambda_1|, |\lambda_2|\} < 1$, applying theorem 3.1, the following is known:

$$\lim_{\kappa \to \infty} A^\kappa = \begin{vmatrix} 0 & 0 \\ 0 & 0 \end{vmatrix}, \tag{62}$$

which implies the existence of a K such if $A^K = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$ then

$$\max\{|a_{11}| + |a_{12}|, |a_{21}| + |a_{22}|\} < 1. \tag{63}$$

Let $\{\xi_1, \xi_2 \ldots, \xi_K\}$ be an arbitrary set, where $\xi_j \in \{1, 2, \ldots, U_i\}$ and $U_i = \psi_i \hat{\psi}_i$ is the number of unique vectors $\boldsymbol{b_t^i}$. Then, given an arbitrary $\boldsymbol{y}, \boldsymbol{z} \in \mathbb{R}^2$,

$$\begin{aligned}
&d(T^i_{(\xi_1)} \circ T^i_{(\xi_2)} \circ \ldots \circ T^i_{(\xi_{K-1})} \circ T^i_{(\xi_K)}\boldsymbol{y}, T^i_{(\xi_1)} \circ T^i_{(\xi_2)} \circ \ldots \circ T^i_{(\xi_{K-1})} \circ T^i_{(\xi_K)}\boldsymbol{z}) \\
&= \|A^K\boldsymbol{y} + A^{K-1}\boldsymbol{b}^i_{\xi_K - 1} + \ldots + \boldsymbol{b}^i_{\xi_1 - 1} - \big(A^K\boldsymbol{z} + A^{K-1}\boldsymbol{b}^i_{\xi_K - 1} + \ldots + \boldsymbol{b}^i_{\xi_1 - 1}\big)\|_\infty \\
&= \|A^K(\boldsymbol{y} - \boldsymbol{z})\|_\infty \\
&= \max\{|a_{11}(y_1 - z_1) + a_{12}(y_2 - z_2)|, |a_{21}(y_1 - z_1) + a_{22}(y_2 - z_2)|\} \\
&\le \max\{|y_1 - z_1|, |y_2 - z_2|\} \max\{|a_{11} + a_{12}|, |a_{21} + a_{22}|\} \\
&= \|\boldsymbol{y} - \boldsymbol{z}\|_\infty \max\{|a_{11} + a_{12}|, |a_{21} + a_{22}|\} \\
&= d(\boldsymbol{y}, \boldsymbol{z}) \max\{|a_{11} + a_{12}|, |a_{21} + a_{22}|\}.
\end{aligned} \tag{64}$$

From equations (63) and (64) it follows that the operator

$$T^i_{(\xi_1)} \circ T^i_{(\xi_2)} \circ \ldots \circ T^i_{(\xi_{K-1})} \circ T^i_{(\xi_K)} \tag{65}$$

is a contractive mapping based on definition 3.9, where $a = \max\{|a_{11} + a_{12}|, |a_{21} + a_{22}|\}$. Using theorem 3.2, there exists an operator $\hat{T}_i$ such that

$$\lim_{t\to\infty} T^i_{(t)} \circ T^i_{(t-1)} \circ \ldots T^i_{(3)} \circ T^i_{(2)}\boldsymbol{y} = \hat{T}_i\boldsymbol{y}. \tag{66}$$

Given equation (66) it follows that $\forall\epsilon > 0$ there exists an $M(\epsilon) \in \mathbb{N}$ such that

$$\|T^i_{(m)} \circ T^i_{(m-1)} \circ \ldots \circ T^i_{(3)} \circ T^i_{(2)}\boldsymbol{y} - T^i_{(n)} \circ T^i_{(n-1)} \circ \ldots \circ T^i_{(3)} \circ T^i_{(2)}\boldsymbol{y}\| \le \epsilon\, \forall n, m > M(\epsilon). \tag{67}$$

Shifting focus back to the PSO model and observing that for each i the following sequence of particle i's positions are defined:

$$\hat{x}_i(1) = \begin{vmatrix} x_i(1) \\ x_i(0) \end{vmatrix}, \quad \hat{x}_i(2) = \begin{vmatrix} x_i(2) \\ x_i(1) \end{vmatrix} = T^i_{(2)} \begin{vmatrix} x_i(1) \\ x_i(0) \end{vmatrix}, \quad \hat{x}_i(3) = \begin{vmatrix} x_i(3) \\ x_i(2) \end{vmatrix} = T^i_{(3)} \begin{vmatrix} x_i(2) \\ x_i(1) \end{vmatrix}$$
$$\ldots \quad \hat{x}_i(n) = \begin{vmatrix} x_i(n) \\ x_i(n-1) \end{vmatrix} = T^i_{(n)} \begin{vmatrix} x_i(n-1) \\ x_i(n-2) \end{vmatrix}, \tag{68}$$

using equation (67), the sequence $\{\hat{x}_i(n)\}$ is Cauchy, since

$$\begin{aligned} &\|\hat{x}_i(m) - \hat{x}_i(n)\| \\ &= \|T^i_{(m)} \circ T^i_{(m-1)} \circ \ldots \circ T^i_{(3)} \circ T^i_{(2)}\hat{x}_i(0) - T^i_{(n)} \circ T^i_{(n-1)} \circ \ldots \circ T^i_{(3)} \circ T^i_{(2)}\hat{x}_i(0)\| \\ &< \epsilon, \end{aligned} \tag{69}$$

for all $m, n > M(\epsilon)$. Since $\{\hat{x}_i(n)\}$ is Cauchy and $\mathbb{R}^2$ is a Banach space, $\{\hat{x}_i(n)\}$ converges to a point in $\mathbb{R}^2$, say $\hat{x}_i \in \mathbb{R}^2$. Therefore, the sequence $\{x_i(n)\}$ converges to a point say $x_i \in \mathbb{R}$ (as x_i is the first component of $\hat{x}_i$). Since i was arbitrary in $1 \le i \le N$, each particle's position converges to a point in $\mathbb{R}$.

3.4 Coefficient region guaranteeing convergence

The results of subsection 3.3 can be significantly simplified for practical application. As was proved in subsection 3.3, if $\max\{|\lambda_1|, |\lambda_2|\} < 1$, with λ_1 and λ_2 as defined in equations (60) and (61), convergence of each particle to a point is guaranteed. If the conditions are violated, the deterministic PSO will not converge. This condition can be simplified by overlaying the inequality plots of $|\lambda_1| < 1$ and $|\lambda_2| < 1$ to the following conditions:

$$|w| < 1, \qquad 0 < \theta_1 + \theta_2 < 4, \qquad w > \frac{(\theta_1 + \theta_2)}{2} - 1. \tag{70}$$

The region of parameters that would ensure particle convergence to a point is illustrated in figure 3, where $\theta = \theta_1 + \theta_2$.

The conditions presented in equation (70) are very positive as they match the conditions of the theoretical research on deterministic PSO done by Van den Bergh and Engelbrecht (2006), as well as the work of Trelea (2003), just with the inclusion of a possible negative inertia weight. In both of the theoretical studies, r_1 and r_2 are fixed to a specific value in the derivation of the conditions needed for

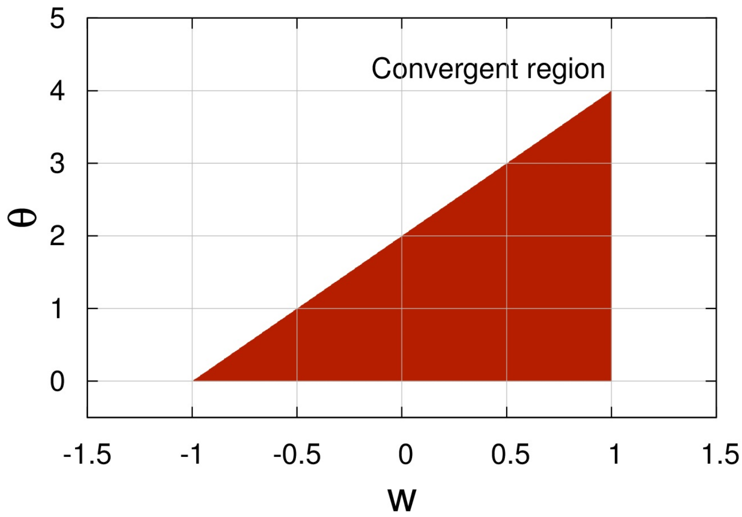

Fig. 3 Parameter region in which each particle converges to a point

particles to converge to a point. However, a slightly different approach to handling the stochastic component is used in this paper, namely the stochastic component is left unfixed throughout the analysis.

The relationship between the previous research and this paper can be seen by replacing θ_1 and θ_2 with $\frac{1}{2}c_1$ and $\frac{1}{2}c_2$ in equation (70) to match Trelea's conditions in equation (39). Similarly, by replacing θ_1 and θ_2 with c_1 and c_2 in equation (70), Van den Bergh and Engelbrecht's conditions in equation (30) are matched.

The conditions of equation (70) are left in their most general form, as it provides the most clarity with regards to the influence of the stochastic component. Specifically, it is clear how likely it is for an iteration to encourage convergence (follow the condition of equation (70)), for example, if $r_1 = r_2 = 0.2$, then an iteration that encourages convergence would at worst only have a 20 percent chance of occurring (if the coefficients used are on the edge of the allowable convergent parameter region). This is obviously not a comment on the long term convergence of particles to a point for the stochastic PSO. It is, however, useful when choosing parameters to have knowledge about the impact that the choice has on the individual iterations.

The similarity between the derived conditions of equation (70) and the condition of Van den Bergh and Engelbrecht (2006), as well as Trelea (2003), is not very surprising. This is because if the assumption is made that each particle's personal best and neighbourhood's best position are in fact fixed (the stagnation assumption), then in subsection 3.3 U_i is 1 for all particles, and the same conditions as in equation (70) can be derived. This is a clear indication that the proof utilized in subsection 3.3 is in fact a generalization of a proof needed to prove convergence of particles to a point for the deterministic PSO under the stagnation assumption.

The conditions of equation (70) are very useful as they provide guarantees about particle convergence derived under a more complete model. The more complete the model, the more accurate the conclusions derived from the model will be.

4 Empirical analysis

This section contains the details of an experiment designed to illustrate the convergence/divergence properties of the PSO algorithm under certain parameter choices in relation to the theoretical findings of the previous section. The section begins with the experimental setup in subsection 4.1, followed by the obtained results in subsection 4.2. The section ends with a discussion of the results and how they relate to the conditions derived in subsection 3.4.

4.1 Experimental set up

In theoretical research on deterministic PSOs, if any experimental results are presented, they are often generated from the theoretical model itself. However, the theoretical model itself is based on a situation where the stagnation assumption is forced. While the results obtained in subsection 3.4 were derived under the weak chaotic assumption, the experiment preformed in this section are on an actual PSO with the only simplification being the deterministic assumption, where $\boldsymbol{\theta}_1 = \mathbf{1}c_1$, and $\boldsymbol{\theta}_1 = \mathbf{1}c_2$. This approach provides a better real world approximation, as it includes the impact of the more complicated interactions between particles in the analysis.

There is an inherent difficulty in empirically analysing the convergence behaviour of PSO particles, specifically with regards to understanding the influence of the underlying objective functions landscape on the PSO algorithm. In an attempt to try and mitigate this issue, a specific objective function that will make it "hard" for PSO to become stagnant is used. This type of function is ideal, as it makes analysis of the PSO's convergence properties more difficult, as a non-stagnate PSO is by definition more chaotic. The function used is

$$f(\boldsymbol{x}) \in U\left(-1000, 1000\right). \tag{71}$$

The objective function in equation (71) is constructed on initialization, and remains static from this point onwards. What the objective function in equation (71) provides is an environment that is rife with discontinuities (actually, it is discontinuous almost everywhere), resulting in a search space where finding the global optimum is very difficult. The objective function was chosen to illustrate that even in the "worst case" scenario, utilisation of the derived conditions of equation (70) will ensure convergence.

The experiment utilizes the following static parameters:

- Population size: 64.
- Iterations: 2000.
- Search space: Dimension $k = 100$.
- Population initialisation: Particle's positions are instantiated within the hypercube $(-1000, 1000)^k$. Particle's velocities are instantiated to $\mathbf{0}$.
- Objective function: Equation (71) is utilized as the objective function.

The measure of convergence is as follows:

$$\Delta\left(t+1\right) = \frac{1}{k}\sum_{i=1}^{i=k}\|\boldsymbol{x}_i\left(t+1\right) - \boldsymbol{x}_i\left(t\right)\|_2. \tag{72}$$

Convergence of each particle to a point will occur if and only if Δ approaches zero.

Four social network structures were utilized in the experiment: star, ring, 2-D von Neumann, and 3-D von Neumann social network structures. The choice of a population of size 64 was made to ensure that both the 2-D von Neumann and the 3-D von Neumann social network structures are completely formed, i.e., a complete grid or cube respectively. The convergence of the PSO algorithm is tested under each of the parameter sets in table 1.

Table 1 Parameter sets used in experiments

No	w	c_1	c_2
1	1	2	2
2	0.7298	1.49618	1.49618
3	0.9	1.85	1.85
4	-0.5	0.4	0.4
5	1	0.5	0.5
6	0.9	-0.1	-0.1
7	0.9	2.01	2.01
8	0.5	1.5	1.5

Each of the parameter sets were chosen in such a way as to illustrate the impact that certain parameter choices have on the long term convergence behaviour of the PSO, or to utilize common or historically chosen parameters. The justification for each parameter set is given in subsection 4.2, along with the detailed results of this experimentation. Each reported experiment is the result of averaging over 50 independent runs.

4.2 Experimental results and discussion

This section presents the results of the experiment as described in subsection 4.1. For each parameter set in table 1, a figure illustrating the variation of Δ as t increases is provided. To ensure that the figures are as informative as possible, two aspects of the figures vary. Firstly, the number of iterations displayed has been chosen to best illustrate the more volatile parts of the convergence behaviour. Secondly, in the cases where radical divergence occurred, Δ has been log scaled in order to avoid later iterations obscuring the results of earlier ones. If a figure contains a log scaled Δ, this is made explicit in the figure's legend.

Parameter set 1 in table 1 are the parameters used in the original PSO (Kennedy and Eberhart, 1995). This parameter set has known issues as mentioned in subsection 2.2.1. The parameter set clearly does not satisfy the conditions of equation (70); specifically, it fails to satisfy all three of the conditions. The results of running the PSO algorithm under this parameter set are illustrated in figure 4. The PSO algorithm was clearly divergent, with Δ rapidly increasing to the extent where log scaling was needed to even analyse the divergence. While there was a slight difference in the rate of divergence between different social network structures, the divergent property was apparent throughout.

Parameter set 2 was originally proposed by Eberhart and Shi (2000). Checking the conditions of equation (70), it is seen that $|0.7298| < 1$, $0 < 2.99236 < 4$, and

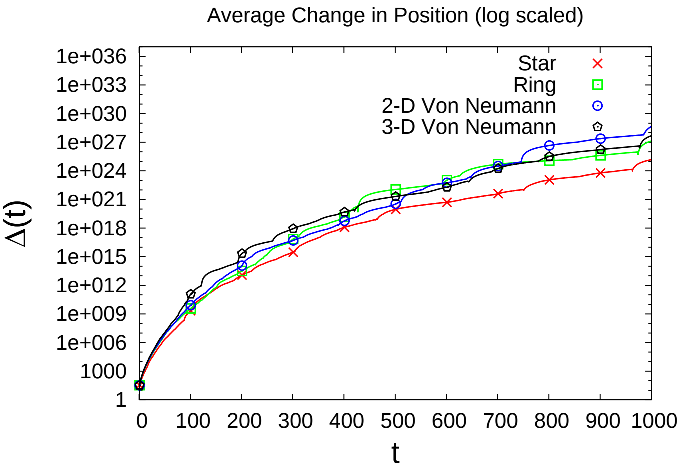

Fig. 4 Convergence behavior of the PSO algorithm under parameter set 1: $w = 1$, $c_1 = 2$, and $c_2 = 2$

$0.7298 > 1.49618 - 1 = 0.49618$, so all the conditions for convergence are satisfied. Figure 5 shows that convergence did in fact occur, with the PSO algorithm using the star social network structure converging the fastest. The long term convergent property of the PSO algorithm under parameter set 2 was, however, not effected by the social network structure as is clear from figure 5; once again only the rate of convergence seems to be affected.

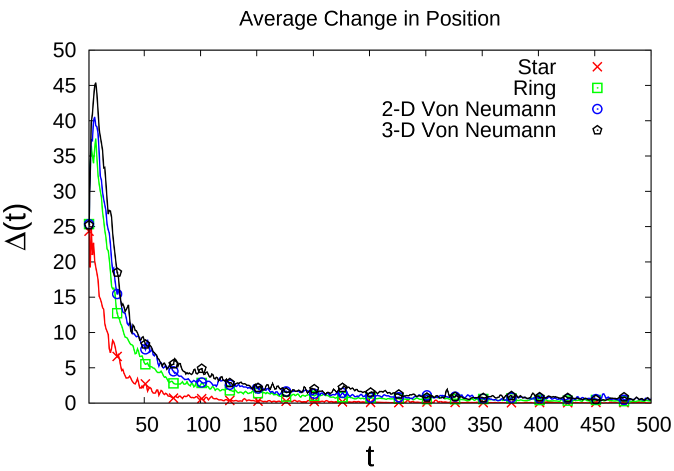

Fig. 5 Convergence behavior of the PSO algorithm under parameter set 2: $w = 0.7298$, $c_1 = 1.49618$, and $c_2 = 1.49618$

Parameter set 3 was chosen as it comes very close to violating every condition in equation (70), but does not actually violate any. The parameter set results in the evaluation $|0.9| < 1$, $0 < 3.6 < 4$, and $0.9 > 1.8 - 1 = 0.8$, which should ensure convergence. In figure 6 convergence of Δ to 0 occurred under all social network structures with the star social network structure providing the quickest and most stable (smallest variations between iterations) convergence to 0, and the 3-D von Neumann social network structure being the slowest to converge in addition to being the least stable.

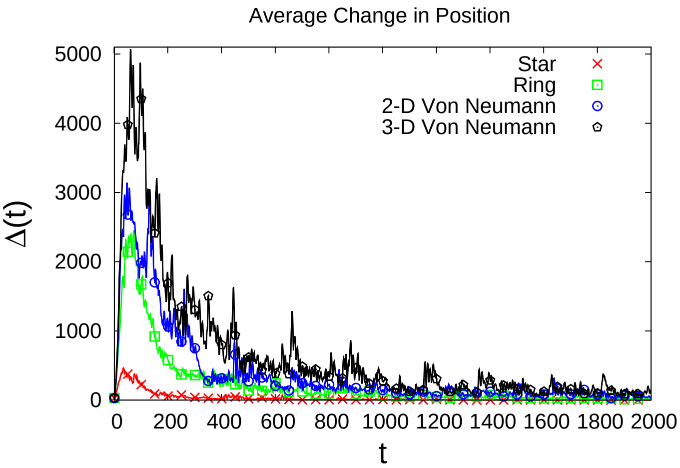

Fig. 6 Convergence behavior of the PSO algorithm under parameter set 3: $w = 0.9$, $c_1 = 1.85$, and $c_2 = 1.85$

Parameter set 4 was chosen as it utilizes a negative inertia weight, a case not often considered. The parameter set results in the evaluation $|-0.5| < 1$, $0 < 0.8 < 4$, and $-0.5 > 0.4 - 1 = -0.6$, which should ensure convergence. Once again, convergence of Δ to 0 occurred under all social network structures as can be seen in figure 7. The convergence behaviour of all the social network structures were virtually identical under parameter set 4.

Parameter set 5 was chosen as it violates only the first condition of equation (70), and is on the boundary of the convergence region of figure 3. The parameter set results in the evaluation $|1| \not< 1$, $0 < 1 < 4$, and $1 > 0$, so the theory does not guarantee convergence. In figure 8 the value of Δ is clearly not converging to 0 under any social network choice. However, Δ was not necessarily diverging to ∞, as was the case with parameter set 1. The long term behaviour of Δ more closely resembles convergence of Δ to a non-zero constant. While this is a more stable behaviour, it is not necessarily a desirable one, as a constant Δ could still result in a particle travelling infinitely far from the desired search space.

Parameter set 6 was chosen as it violates only condition two. The parameter set results in the evaluation $|0.9| < 1$, $0 \not< -0.2 < 4$, and $0.9 > -1.1$, so the theory does not guarantee convergence. Also, parameter set 7 was chosen as it violates condition two from above, and therefore also condition three. The parameter set 7 results in the evaluation $|0.9| < 1$, $0 < 4.02 \not< 4$, and $0.9 \not> 2.1 - 1 = 1.1$, so once again the theory does not guarantee convergence. Under the utilization of both parameter sets 6 and 7, rapid divergence occurred, as seen in figures 9 and

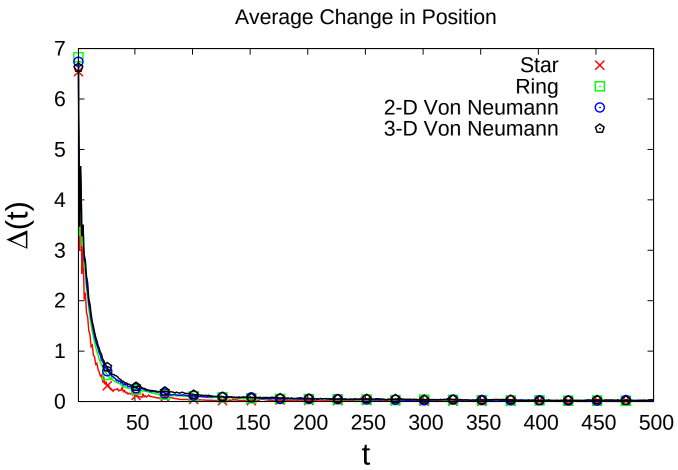

Fig. 7 Convergence behavior of the PSO algorithm under parameter set 4: $w = -0.5$, $c_1 = 0.4$, and $c_2 = 0.4$

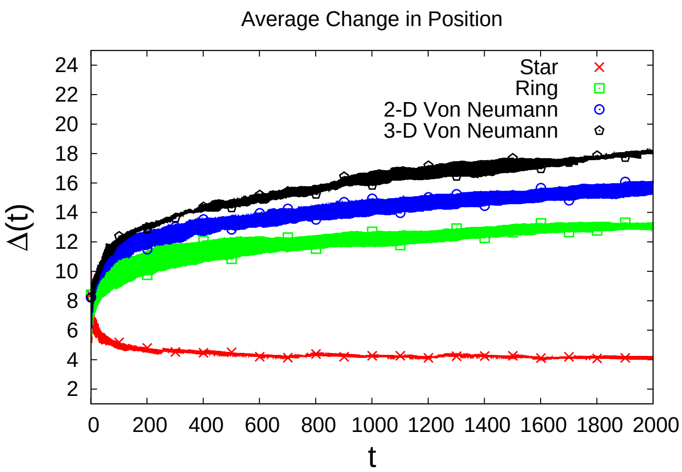

Fig. 8 Convergence behavior of the PSO algorithm under parameter set 5: $w = 1$, $c_1 = 0.5$, and $c_2 = 0.5$ (the shaded appearance of the lines is due to rapid oscillation)

10, with Δ causing overflow before 500 iterations have occurred. The divergence of Δ occurred regardless of social network choice. The choice of social network structure appears to only slightly affect the rate at which Δ diverges.

Parameter set 8 was chosen as it violates only condition three. The parameter set results in the evaluations $|0.5| < 1$, $0 < 3 < 4$, and $0.5 \not> 1.5 - 1 = 0.5$, so the theory does not guarantee convergence. The long term behaviour of Δ can be seen in figure 11. Parameter set 8 resulted in one of the more interesting behaviours. All social network structures except for the star network structure clearly exhibited near linear divergence. However, the divergence of the star social network structure was exceptionally slow in contrast to the other social network

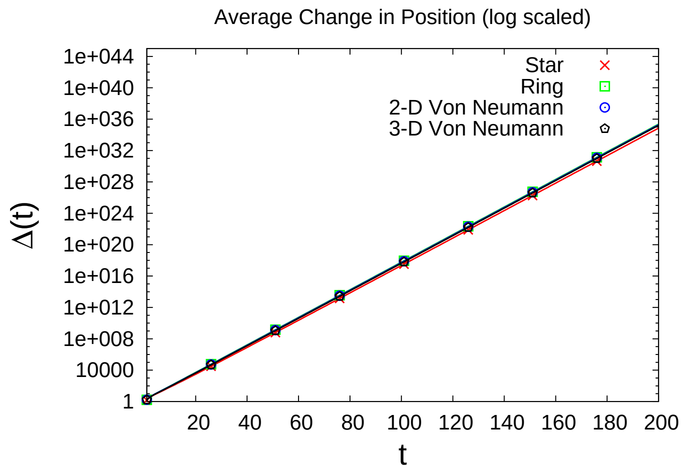

Fig. 9 Convergence behavior of the PSO algorithm under parameter set 6: $w = 0.9$, $c_1 = -0.1$,and $c_2 = -0.1$

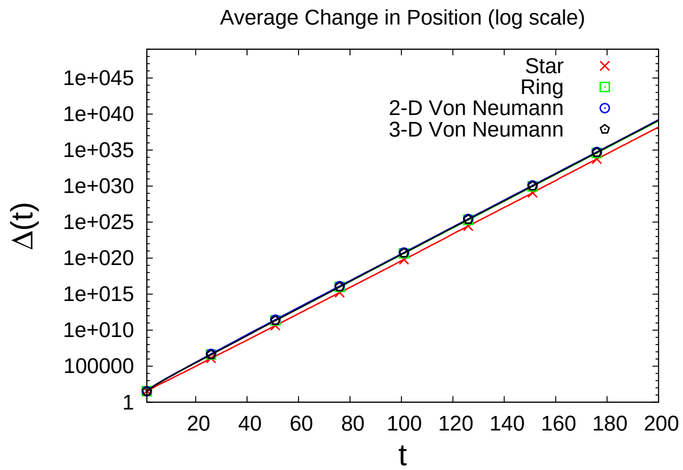

Fig. 10 Convergence behavior of the PSO algorithm under parameter set 7: $w = 0.9$, $c_1 = 2.01$, and $c_2 = 2.01$

structures. Parameter set 8 provided the best indications that the social network choice does in fact have an effect on the rate on divergence. However, the social network choice did not have a large enough effect to cause a convergent Δ.

5 Conclusions

This paper analysed current theoretical research on PSO in both the stochastic and deterministic context. The primary focus was on the derivation of conditions necessary for each particle in a swarm to converge to a point. It was found that

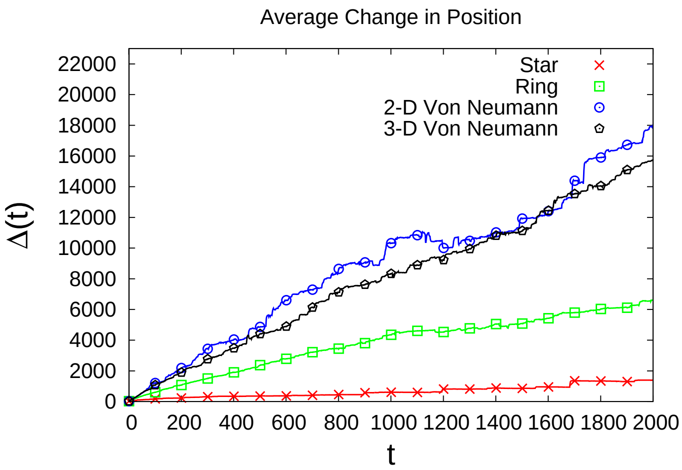

Fig. 11 Convergence behavior of the PSO algorithm under parameter set 8: $w = 0.5$, $c_1 = 1.5$, and $c_2 = 1.5$

all theoretical models used in the research had one common shortcoming: all of the models derived their conditions under the stagnation assumption as discussed in section 2. While the theoretical PSO models utilized in the research have provided a wealth of insight into the PSO's underlying behaviour, the presence of the stagnation assumption has removed important aspects of the PSO algorithm's behaviour.

Section 3 begun with a detailed explanation of the aspects of the PSO algorithm that the presence of the stagnation assumption had inherently removed from the previous theoretical models. A new theoretical model for the deterministic PSO algorithm was proposed under the weak chaotic assumption, a substantially weaker assumption than the stagnation assumption. Under the proposed model, the conditions necessary for each particle in a swarm to converge to a point were derived. The conditions were derived under an arbitrary social network structure and swarm size, as well as under an arbitrary objective function.

Section 4 provided an experiment designed to test the theoretically derived condition for particle convergence to a point. The experiment was run using an actual PSO algorithm under multiple social network structures, with only the stochastic component of the update equations fixed.

It was found that the PSO algorithm behaved in the theoretically expected manner with regards to long term behaviour under all the parameter sets. The results also illustrated that the choice of social network structure does in fact affect the particles' convergence. However the effect is limited to the rate at which convergence or divergence occurred. This conclusion matches the results of the theoretical analysis done in section 3, namely that long term convergence will occur if the parameters used satisfy the conditions of equation (70) regardless of the social network structure used. This empirical study on the PSO algorithm differs from other studies done, as it directly analyses the effect that the social network structure choice has on the swarm's ability to have each particle converge to a point.

In summary, a generalization of the theoretical deterministic PSO was made which better reflected the actual PSO algorithm. Under the new model, conditions to guarantee particle convergence to a point were derived and supported by empirical findings.

There are two very clear directions for future work which are directly related to the work of this paper. The first would be to derive the conditions necessary for each particle in a swarm to converge to a point under the theoretical stochastic PSO model with the weak chaotic assumption, instead of the usual stagnation assumption. The second would be to derive the conditions necessary for each particle in a swarm to converge to a point, without any assumptions on the nature of the personal best and neighbourhood best particles at all. This derivation might not actually be achievable, nevertheless it should be investigated.

References

Burden, R. and Faires, J. (2010). *Numerical Analysis, 9th edition*. Cengage Learning, Hampshire, UK.

Carlisle, A. and Dozier, G. (2001). An off-the-shelf pso. In *Proceedings of the Workshop on Particle Swarm Optimization*, pages 1–6, Indianapolis, IN, USA. IUPUI.

Clerc, M. (1999). The swarm and the queen: Towards a deterministic and adaptive particle swarm optimization. In *proceedings of the IEEE Congress on Evolutionary Computation*, volume 3, pages 1951–1957, Piscataway, NJ, USA. IEEE Press.

Clerc, M. and Kennedy, J. (2002). The particle swarm-explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, 6(1):58–73.

Dosoer, C. and Vidyasagar, M. (1975). *Feedback Systems: Input Output Properties*. Academic Press, New York, NY, USA.

Eberhart, R. and Shi, Y. (2000). Comparing inertia weights and constriction factors in particle swarm optimization. In *Proceedings of the IEEE Congress on Evolutionary Computation*, volume 1, pages 84–88, Piscataway, NJ. IEEE Press.

Eberhart, R., Simpson, P., and Dobbins, R. (1996). Computational intelligence pc tools. *Academic Press Professional, first edition*.

Engelbrecht, A. (2013a). Particle swarm optimization: Global best or local best. In *1st BRICS Countries Congress on Computational Intelligence*, Piscataway, NJ, USA. IEEE Press.

Engelbrecht, A. (2013b). Particle swarm optimization: Iteration strategies revisted. In *1st BRICS Countries Congress on Computational Intelligence*, Piscataway, NJ, USA. IEEE Press.

Gazi, V. (2012). Stochastic stability analysis of the particle dynamics in the PSO algorithm. In *Proceedings of the IEEE International Symposium on Intelligent Control*, pages 708–713, Dubrovnik, Croatia. IEEE Press.

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

Kadirkamanathan, V., Selvarajah, K., and Fleming, P. (2006). Stability analysis of the particle dynamics in particle swarm optimizer. *IEEE Transactions on Evolutionary Computation*, 10(3):245–255.

Kennedy, J. (1999). Small worlds and mega-minds: effects of neighborhood topology on particle swarm performance. In *Proceedings of the IEEE Congress on Evolutionary Computation*, volume 3, pages 1931–1938, Piscataway, NJ, USA. IEEE Press.

Kennedy, J. and Eberhart, R. (1995). Particle swarm optimization. pages 1942–1948, Piscataway, NJ, USA. IEEE Press.

Kennedy, J. and Mendes, R. (2002). Population structure and particle performance. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pages 1671–1676, Piscataway, NJ, USA. IEEE Press.

Kisacanin, B. and Agarwal, G. (2001). *Linear Control Systems: With Solved Problems and Matlab Examples*. Springer.

Martinez, F., Gonzalo, G., and Alvarez, F. (2008). Theoretical analysis of particle swarmtrajectories through a mechanical analogy. *International Journal of Computational Intelligence Research*, 4(2):93–104.

Mate, L. (1999). On infinite composition of affine mappings. *Fundamenta Mathematicae*, 159:85–90.

Naka, S., Genji, T., Yura, T., and Fukuyama, Y. (2001). Practical distribution state estimation using hybrid particle swarm optimization. In *IEEE Power Engineering Society Winter Meeting*, volume 2, pages 815–820, Piscataway, NJ, USA. IEEE Press.

Ozcan, E. and Mohan, C. (1998). Analysis of a simple particle swarm optimization system. *Intelligent Engineering Systems through Artificial Neural Networks*, volume 8:253–258.

Ozcan, E. and Mohan, C. (1999). Particle swarm optimization: Surfing the waves. In *Proceedings of the IEEE Congress on Evolutionary Computation*, volume volume 3, Piscataway, NJ, USA. IEEE Press.

Peer, E., Van den Bergh, F., and Engelbrecht, A. (2003). Using neighborhoods with the guaranteed convergence PSO. In *Proceedings of the IEEE Swarm Intelligence Symposium*, pages 235–242, Piscataway, NJ, USA. IEEE Press.

Peng, J., Chen, Y., and Eberhart, R. (2000). Battery pack state of charge estimator design using computational intelligence approaches. In *Proceedings of the Annual Battery Conference on Applications and Advances*, pages 173–177, Piscataway, NJ, USA. IEEE Press.

Peram, T., Veeramachaneni, K., and Mohan, C. (2003). Fitness-distance-ratio based particle swarm optimization. In *Proceedings of the IEEE Swarm Intelligence Symposium*, pages 174–181, Piscataway, NJ, USA. IEEE Press.

Poli, R. (2009). Mean and variance of the sampling distribution of particle swarm optimizers during stagnation. *IEEE Transactions on Evolutionary Computation*, 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 *Genetic and Evolutionary Computation Conference*, pages 134–141, New York, NY, USA. ACM Press.

Ratnaweera, A., Halgamuge, S., and Watson, H. (2003). Particle swarm optimization with self-adaptive acceleration coefficients. In *Proceedings of the First International Conference on Fuzzy Systems and Knowledge Discovery*, pages

264–268, Los Vaqueros, CA, USA. IEEE Press.

Shi, Y. and Eberhart, R. (1998). A modified particle swarm optimizer. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pages 69–73, Piscataway, NJ, USA. IEEE Press.

Shi, Y. and Eberhart, R. (2001). Fuzzy adaptive particle swarm optimization. In *Proceedings of the IEEE Congress on Evolutionary Computation*, volume 1, pages 101–106, Piscataway, NJ, USA. IEEE Press.

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

Van den Bergh, F. (2002). An analysis of particle swarm optimizers. *PhD thesis, Department of Computer Science, University of Pretoria, Pretoria, South Africa*.

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

Venter, G. and Sobieszczanski-Sobieski, J. (2003). Particle swarm optimization. *Journal for the American Institute of Aeronautics and Astronautics*, 41(8):1583–1589.

Vidyasagar, M. (2002). *Nonlinear Systems Analysis*. SIAM, Philadelphia, PA, USA.

Zheng, Y., Ma, L., Zhang, L., and Qian, J. (2003). On the convergence analysis and parameter selection in particle swarm optimization. In *the International Conference on Machine Learning and Cybernetics*, volume 3, pages 1802–1807, Piscataway, NJ, USA. IEEE Press.