2014 IEEE Congress on Evolutionary Computation (CEC)
July 6-11, 2014, Beijing, China

# Particle Swarm Convergence: An Empirical Investigation

Christopher W Cleghorn and Andries P Engelbrecht

***Abstract*— This paper performs a thorough empirical investigation of the conditions placed on particle swarm optimization control parameters to ensure convergent behavior. At present there exists a large number of theoretically derived parameter regions that will ensure particle convergence, however, selecting which region to utilize in practice is not obvious. The empirical study is carried out over a region slightly larger than that needed to contain all the relevant theoretically derived regions. It was found that there is a very strong correlation between one of the theoretically derived regions and the empirical evidence. It was also found that parameters near the edge of the theoretically derived region converge at a very slow rate, after an initial population explosion. Particle convergence is so slow, that in practice, the edge parameter settings should not really be considered useful as convergent parameter settings.**

## I. Introduction

PARTICLE SWARM OPTIMIZATION (PSO) is a stochastic population-based search algorithm that has been effectively utilized to solve numerous optimization problems [1]. Despite PSO's widespread use, there are many aspects of the PSO algorithm that are not fully understood. The focus of this paper is on the conditions necessary for particles within the swarm to converge to an arbitrary point.

While there have been many theoretical studies on the convergence of PSO particles [2], [3], [4], [5], [6], [7], [8], [9], [10], [11], [12], their conclusions are often unfortunately obscured by rigorous mathematics. Almost all recent theoretical studies provide a region of parameter space that will ensure particle convergence, however, these derived regions do not match. This discrepancy in the theoretical results is largely due to the differing use of initial assumptions and the proof techniques used. At present there exists no assumption free theoretical PSO analysis that produces the guaranteed regions for particle convergence and divergence. While from a theoretical stand point each derived region is of importance, in practice a PSO user is left with a multitude of "reasonable" regions to select from. The question then becomes, which region should be used?

This paper aims to supplement the theory with an empirical study in order to provide a more transparent picture of the PSO's behavior, given certain parameter choices. This is achieved by measuring the average change in velocity of particles within a swarm for a multitude of initial parameters at each iteration. This empirical study is carried out over a region slightly larger than that needed to contain all the relevant theoretically derived regions.

A brief description of PSO is given in section II. A discussion of the derived parameter regions sufficient for particle convergence is given in section III. The experimental set up and results are given in section IV and V respectively. Section VI presents a summary of the findings of this paper, as well as a discussion of topics for future research.

## II. Particle swarm optimization

Particle swarm optimization (PSO) was originally developed by Kennedy and Eberhart [13] to simulate the complex movement of birds in a flock. The PSO algorithm this paper focuses on includes the inertia coefficient proposed by Shi and Eberhart [14].

The PSO algorithm is defined as follows: Let $f : \mathbb{R}^k \to \mathbb{R}$ 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 onwards. 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

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

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

$$\begin{aligned}\boldsymbol{v}_i(t+1) \quad = \quad & w\boldsymbol{v}_i(t) + c_1\boldsymbol{r}_1(t)(\boldsymbol{y}_i(t) - \boldsymbol{x}_i(t)) \\ & + c_2\boldsymbol{r}_2(t)(\hat{\boldsymbol{y}}_i(t) - \boldsymbol{x}_i(t)),\end{aligned} \tag{2}$$

where $\boldsymbol{r}_1(t), \boldsymbol{r}_2(t) \sim U(0,1)^k$ for all $t$. 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 neighborhood of the $i$-th particle have visited. For a full explanation of various neighborhood (NBHD) choices and their impact on performance, the reader is referred to [15], [16], [17], [18]. The coefficients $c_1$, $c_2$, and $w$ are the cognitive, social, and inertia weights respectively. The PSO algorithm is summarized in algorithm 1.

## III. Theoretical Background

This section presents each theoretically derived region that is sufficient for particle convergence, along with the corresponding assumptions utilized in the region's derivation. While there is a large body of theoretically work on PSO particle behavior, only the studies that actually contain a parameter region for particle convergence will be discussed in this section. Interested readers are referred to [2], [3], [4], [5].

The assumptions that occur commonly in the theoretical PSO research are as follows:

**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 $t$.

**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 $t$.

**Algorithm 1** PSO algorithm

Create and initialize a $k$-dimensional swarm, $\Omega(0)$, of $N$ particles uniformly within a predefined hypercube.
Let $f$ be the objective function.
Let $\boldsymbol{y}_i$ represent the personal best position of particle $i$, initialized to $\boldsymbol{x}_i(0)$.
Let $\hat{\boldsymbol{y}}_i$ represent the neighborhood best position of particle $i$, initialized to $\boldsymbol{x}_i(0)$.
Initialize $\boldsymbol{v}_i(0)$ to $\mathbf{0}$.
**repeat**
  **for all** particles $i = 1, \cdots, N$ **do**
    **if** $f(\boldsymbol{x}_i) < f(\boldsymbol{y}_i)$ **then**
      $\boldsymbol{y}_i = \boldsymbol{x}_i$
    **end if**
    **for all** particles $\hat{i}$ with particle $i$ in their NBHD **do**
      **if** $f(\boldsymbol{y}_i) < f(\hat{\boldsymbol{y}}_{\hat{i}})$ **then**
        $\hat{\boldsymbol{y}}_{\hat{i}} = \boldsymbol{y}_i$
      **end if**
    **end for**
  **end for**
  **for all** particles $i = 1, \cdots, N$ **do**
    update the velocity of particle $i$ using equation (2)
    update the position of particle $i$ using equation (1)
  **end for**
**until** stopping condition is met

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

Under both the deterministic and stagnation assumption Van den Bergh and Engelbrecht [6], [19] derived the following region for particle convergence:

$$c_1+c_2 < 2(1+w), \quad c_1 > 0, \quad c_2 > 0, \quad 0 < w < 1. \tag{3}$$

Under both the deterministic and stagnation assumption Trelea [7] derived the following region for particle convergence:

$$c_1+c_2 < 4(1+w), \quad c_1 > 0, \quad c_2 > 0, \quad 0 < w < 1. \tag{4}$$

More recently, under the deterministic and weak chaotic assumption Cleghorn and Engelbrecht [8] derived the same region as equation (3) or (4) (depending on the treatment of the stochastic variables $\boldsymbol{r}_1$ and $\boldsymbol{r}_2$) just with $|w| < 1$. The extended versions of equations (3) and (4) with $|w| < 1$ are illustrated in figure 1, as the triangle AFB and ACB respectively.

Kadirkamanathan et al [9], only under the stagnation assumption, derived the following region for particle convergence:

$$\begin{cases} c_1 + c_2 < 2(1+w) & w \in (-1, 0] \\ c_1 + c_2 < \frac{2(1-w)^2}{1+w} & w \in (0, 1). \end{cases} \tag{5}$$

Gazi [10] expanded the derived region of equation (5), also under the stagnation assumption only, resulting in the following region:

$$\begin{cases} c_1 + c_2 < \frac{24(1+w)}{7} & w \in (-1, 0] \\ c_1 + c_2 < \frac{24(1-w)^2}{7(1+w)} & w \in (0, 1). \end{cases} \tag{6}$$

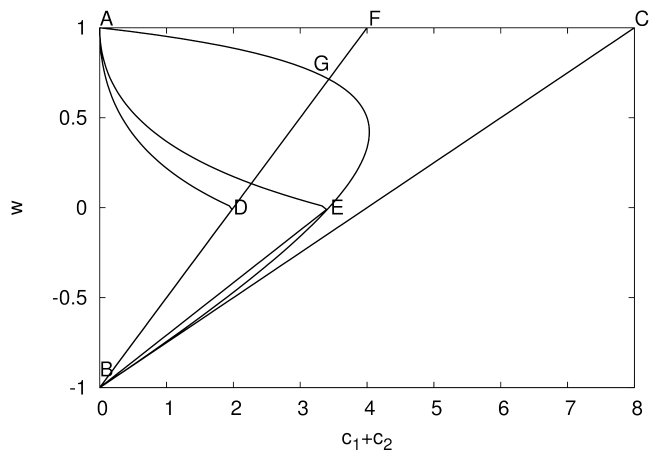

Fig. 1. Theoretically derived regions sufficient for particle convergence

Unfortunately, both equations (5) and (6) were derived utilizing the Lyapunov condition, resulting in conservative regions [20]. The regions corresponding to equation (5) and (6) are illustrated in figure (1) as triangle like regions ADB and AEB respectively.

Lastly, Poli [11], [12] under the stagnation assumption only, but without the use of the Lyapunov condition, derived the following region:

$$c_1 + c_2 < \frac{24\left(1-w^2\right)}{7-5w}. \tag{7}$$

The region defined by equation (7) is illustrated in figure 1 as the curved line segment AB.

## IV. Experimental Setup

The experiment conducted in this paper is designed to illustrate under what parameter settings the PSO algorithm will actually exhibit convergent behavior. There is an inherent difficulty in empirically analyzing the convergence behavior 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, the following objective function that will make it "hard" for PSO to become stagnant, is used:

$$f(\boldsymbol{x}) \in U(-1000, 1000). \tag{8}$$

The objective function in equation (8) is constructed on initialization, and remains static from that point onwards. What the objective function in equation (8) 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 provides a scenario where convergence would be difficult. If particles are seen to be convergent (not necessarily to the same point) with such a degenerate objective function as equation (8), it is reasonable to assume that convergent behavior will hold with less degenerate objective functions too.

The experiment utilizes the following static parameters: Population size of 64, 2000 iterations, a 50-dimensional search space, the star NBHD structure (gbest). Particle's positions are instantiated within $(-1000, 1000)^k$ and velocities are instantiated to $\mathbf{0}$. Equation (8) is utilized as the objective function. A population size of 64 is utilized to allow for easier future comparison of differing PSO NBHD structures, namely a population size of 64 will allow for a complete 2-D and 3-D von Neumann NBHD structures.

The measure of convergence is as follows:

$$\Delta(t+1) = \frac{1}{N}\sum_{i=1}^{N} \|\boldsymbol{x}_i(t+1) - \boldsymbol{x}_i(t)\|_2. \tag{9}$$

The test is conducted over the following parameter region:

$$w \in [-1.1, 1.1] \text{ and } c_1 + c_2 \in (0, 4.4], \tag{10}$$

where $c_1 = c_2$, with a sample point every 0.1 along $w$ and $c_1 + c_2$. A total of 1012 sample points are used. The region defined by equation (10) is slightly larger than that needed to include all the regions defined in section III, except the region defined by equation (4) which is partially omitted. The sub-region

$$c_1 + c_2 < 4(1+w), \quad c_1 + c_2 > 4.4, \tag{11}$$

of the region defined by equation (4) is omitted from the analysis due to the lack of any convergent behavior within the omitted region (this fact will become obvious in section V). The results reported in Section V are the averages over 35 independent runs for each sample point.

## V. Experimental Results and Discussion

This section presents the results of the experiment described in section IV. A snapshot of all parameter configurations' resulting convergence measure values are presented at the PSO's 10th, 500th and 2000th iteration. For each iteration, a snapshot figure with the convergence measure value capped at 100, 500, and 2000 is presented, so as to prevent very large convergence measure values from obscuring the important observations.

At iteration 10 a relatively clear picture of particle behavior is already developing. Even with the low convergence measure limit of 100, figure 2 illustrates parameter settings that are more conducive to convergence behavior, even at this low iteration count of 100. More specifically, the curved region with the apex at $w = 0.2$ and $c_1 + c_2 = 2.3$. In figure 3, the region with the most convergent behavior is already starting to show similarity to the region of equation (7), as illustrated in figure 5. In figure 4 there is a large number of parameter settings exhibiting convergent like behavior, this is not surprising given the early iteration count and the substantial convergence measure cap. What is quite surprising, however, is the large number of parameter settings that have resulted in convergence measure values in excess of 2000 after only 10 iterations. As an illustrative example, the exact convergence measure value at $w = -0.8$, $c_1 + c_2 = 3.5$ in figure 4 is 153101. Such large convergence measure values could be a serious hindrance on PSO's search capability, and in the extreme case, the PSO's search is surely useless [21].

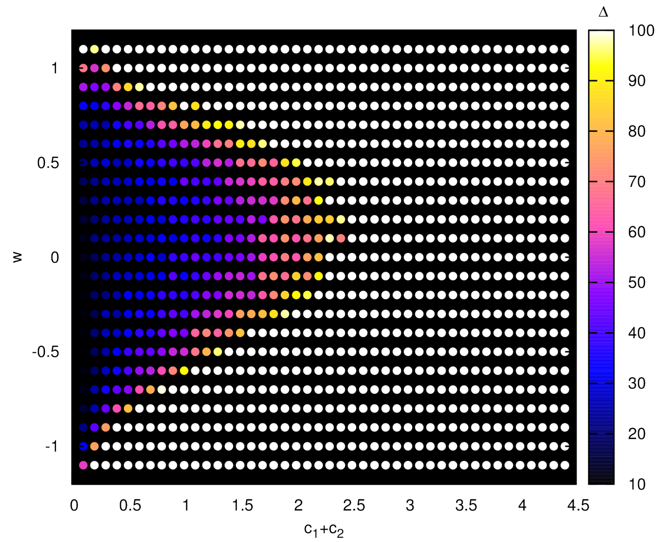


Fig. 2. Recorded convergence measure values after 10 iterations with a $\Delta$ cap of 100

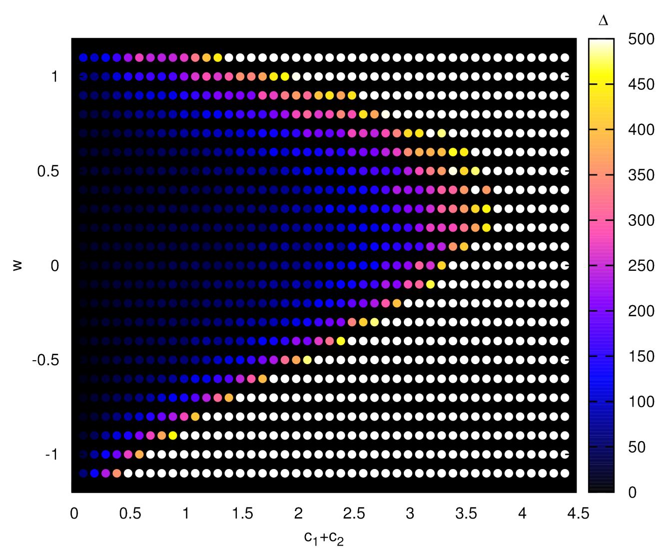


Fig. 3. Recorded convergence measure values after 10 iterations with a $\Delta$ cap of 500

At iteration 500 the particle behavior is already substantially more stable than at iteration 10. This is illustrated by the close similarity between figures 7 and 8. In figure 6 the region of convergent behavior is slightly narrower than that of figure 2 but the apex is further out. However, the number

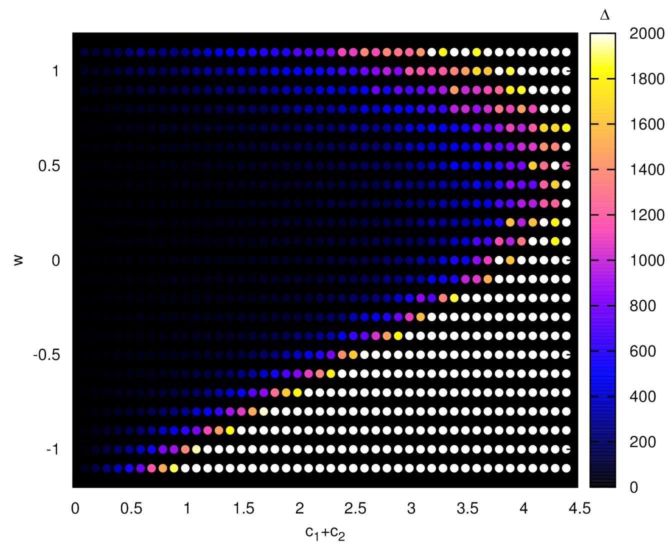


Fig. 4. Recorded convergence measure values after 10 iterations with a $\Delta$ cap of 2000

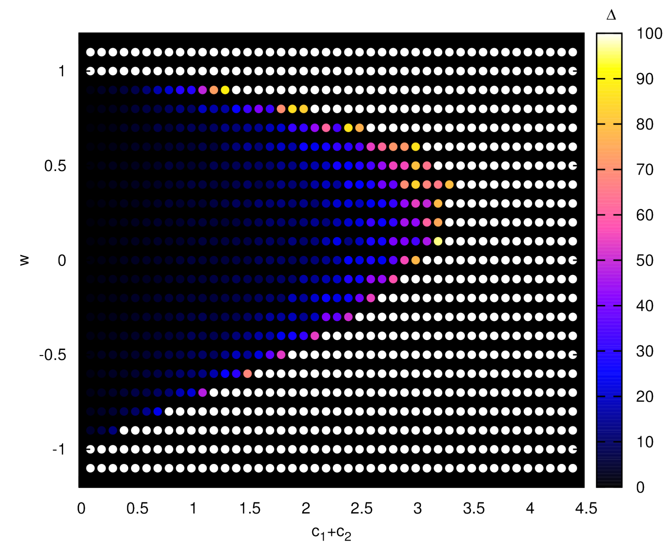


Fig. 6. Recorded convergence measure values after 500 iterations with a $\Delta$ cap of 100

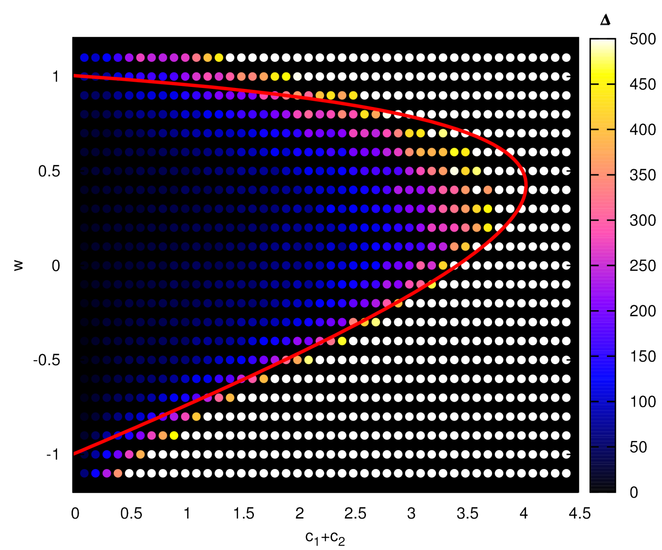


Fig. 5. Recorded convergence measure values after 10 iterations with a $\Delta$ cap of 500 with equation (7) overlaid

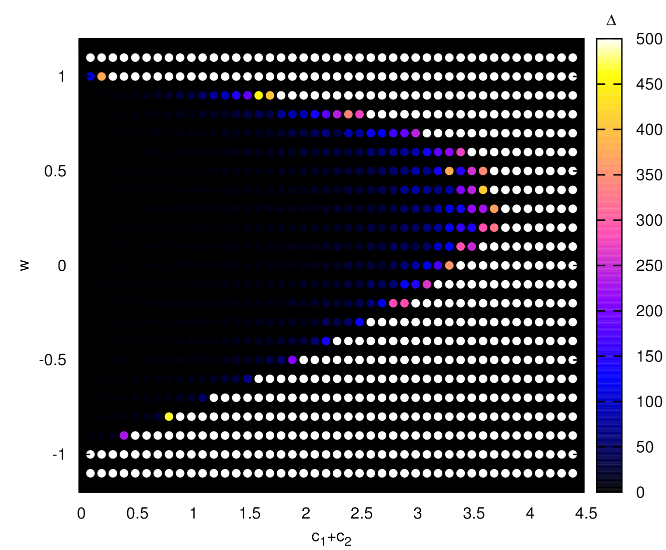


Fig. 7. Recorded convergence measure values after 500 iterations with a $\Delta$ cap of 500

of particles with a convergence measure value below 100 has increased over the course of 490 iterations. In figures 7 and 8 the correspondence between equation (7) and the convergent behavior is becoming very clear, as illustrated in figure 9. The only discrepancy is that the apex of the convergent region of figures 7 and 8 is slightly less than that of equation (7). However, with this correlation in mind, there are still convergence measure values of over 2000 corresponding to parameters that are technically in the region of equation (7).

At iteration 1000 the region of convergent behavior is nearly identical to the corresponding figures for iteration 500, as can bee seen in figures 10, 11 and 12. There is, however, a small decrease in convergence measure values across the region of convergent behavior. For example, when $w = 0.5$ and $c_1 + c_2 = 3$, the convergence measure changed from 79.689 at 500 iterations to 70.8303 at 1000 iterations. This decrease in convergence measure values is not negligible. However, the rate of decrease is particularly slow.

At iteration 2000 the same phenomenon as from 500 to 1000 iterations occurred, namely, figures 13, 14, and 15 are nearly identical to the corresponding figures for 1000 iterations. Again, a small decrease in convergence measure values across the region of convergent behavior occurred. For example, when $w = 0.5$ and $c_1 + c_2 = 3$, the convergence

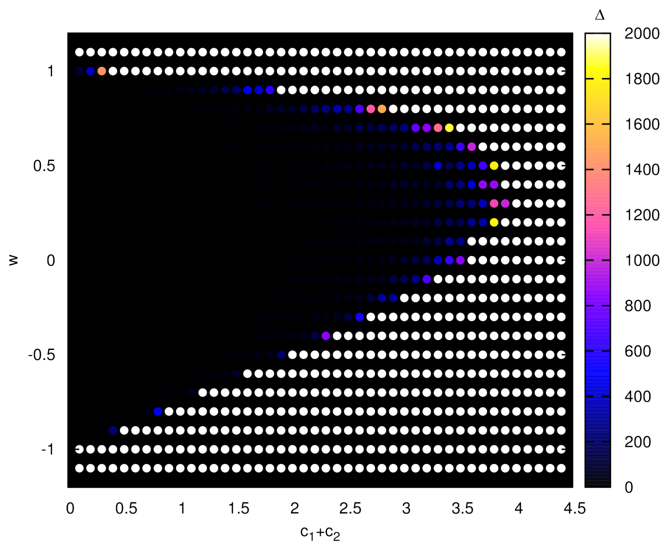


Fig. 8. Recorded convergence measure values after 500 iterations with a $\Delta$ cap of 2000

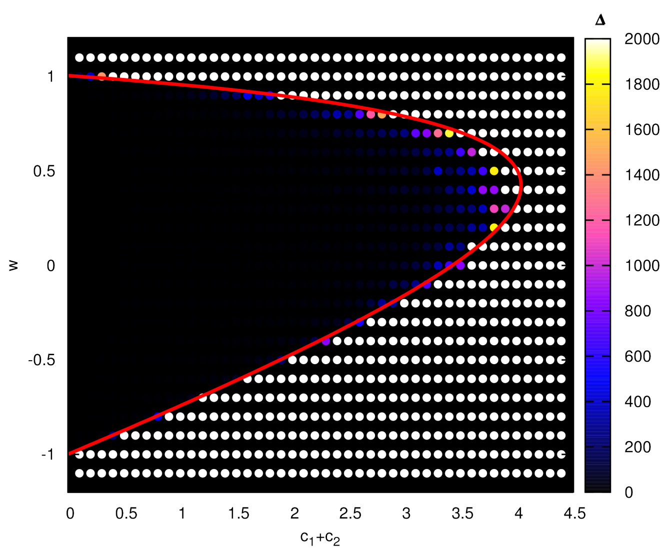


Fig. 9. Recorded convergence measure values after 500 iterations with a $\Delta$ cap of 2000 with equation (7) overlaid

measure changed from 70.8303 at 1000 iterations to 53.8461 at 2000 iterations. As illustrated in figure 16, equation (7) matches almost perfectly with the convergent region of figure 15.

In general, the convergence results correspond well with the derived region of equation (7). The regions defined by equations (5) and (6) both result in convergent behaviour, as they are subsets of the region defined equation (7). However, the regions defined by equations (5) and (6) excludes a large number convergent parameter settings. The extended region define by equation (4) (with $|w| < 1$) is far larger than the actual observed convergent region of parameter setting.

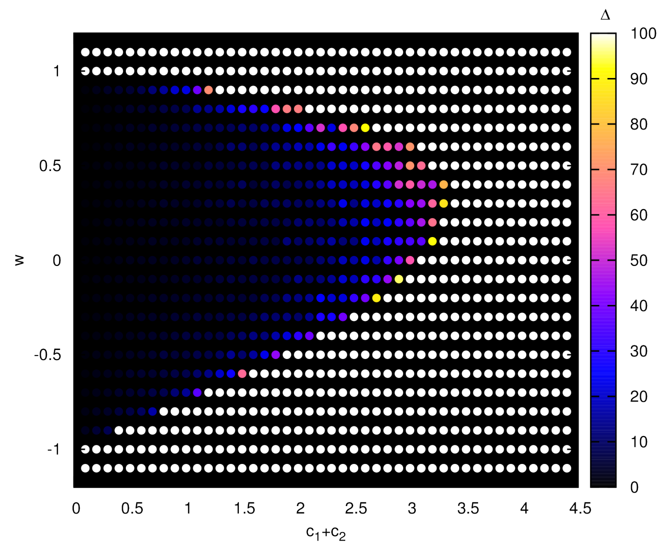


Fig. 10. Recorded convergence measure values after 1000 iterations with a $\Delta$ cap of 100

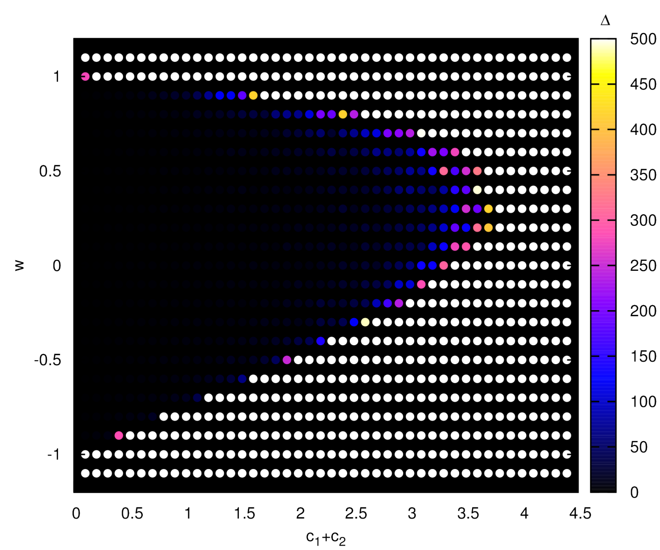


Fig. 11. Recorded convergence measure values after 1000 iterations with a $\Delta$ cap of 500

The extended region define by equation (3) (with $|w| < 1$), contains a large number of convergent parameter settings. However, there are a number of excluded parameters, namely those within the region BEG of figure 1. The extended region define by equation (3), also includes a small number of parameters that do not fall within the actual convergent region, namely, those within the region AGF of figure 1.

There are, however, two additional observations: Firstly, parameters that reside very near to the apex of the region defined in equation (7) do not exhibit a fast convergent trend. For example, the parameter setting $w = 0.5$ and $c_1+c_2 = 3.9$ which is within the derived region of equation (7), has a

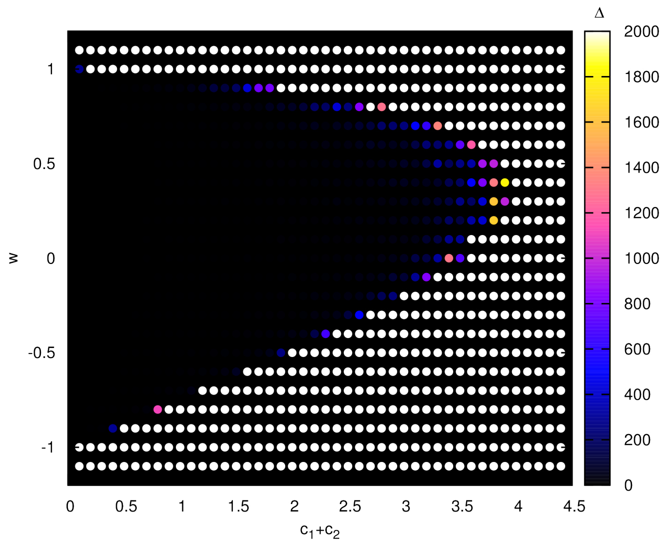


Fig. 12. Recorded convergence measure values after 1000 iterations with a $\Delta$ cap of 2000

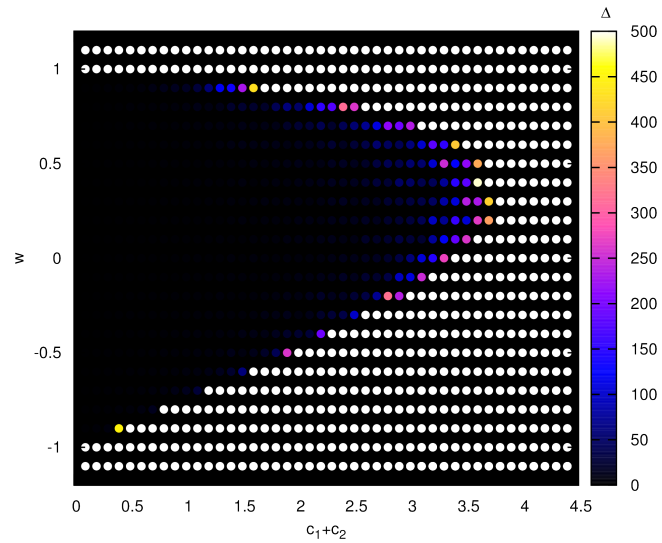


Fig. 14. Recorded convergence measure values after 2000 iterations with a $\Delta$ cap of 500

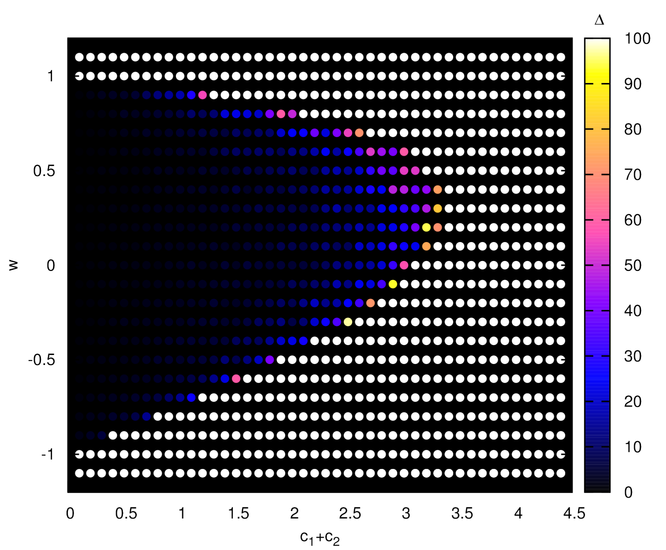


Fig. 13. Recorded convergence measure values after 2000 iterations with a $\Delta$ cap of 100

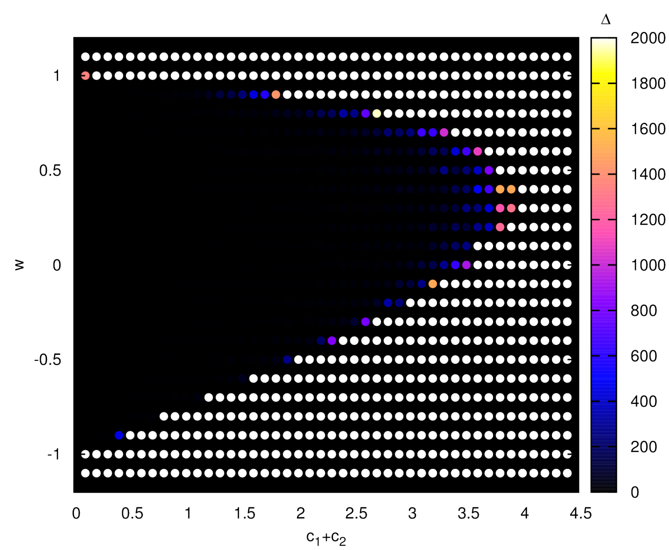


Fig. 15. Recorded convergence measure values after 2000 iterations with a $\Delta$ cap of 2000

convergence measure value of $2576.85$ after $2000$ iterations. While this parameter configuration may actually result in convergence, the rate of convergence is prohibitively slow in practise. The second observation is that the closer $c_1+c_2$ is to zero and $w$ to roughly $0.4$, the quicker particle convergence occurs.

Given these observations, when utilizing PSO in practice, selecting parameters from a region of the same form as that of equation (7), but excluding configurations within $0.1$ of equation (7)'s boundary, will yield a reasonable convergence rate.

## VI. Conclusions

The aim of this study was to perform an experiment that clearly shows which theoretically derived convergent region is most applicable to practical PSO use. It was found that there is a very strong correlation between the convergence behavior of the PSO and the parameter region defined by Poli [11], [12]. Despite this very strong correlation, the empirical results of section V also shows that, when PSO parameter settings are near the edge of the region defined by equation (7), convergence is incredibly slow, with some parameter settings still having an average particle movement of over 2500 after 2000 iterations. From these observations it is

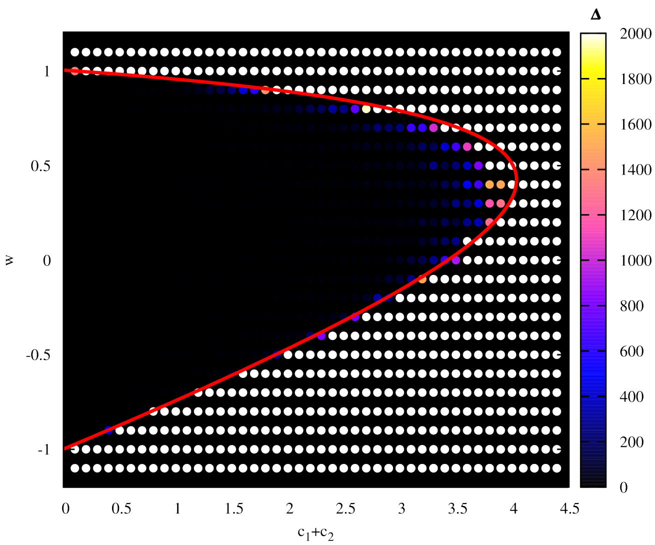


Fig. 16. Recorded convergence measure values after 2000 iterations with a $\Delta$ cap of 2000 with equation (7) overlaid

concluded that, in practice, PSO parameter setting should be selected from a slightly smaller region than that of equation (7), so as to avoid the unreasonably slow rate of particle convergence.

Future work, expanding from this paper, could cater for the impact of the social network used by the PSO. In addition, a theoretical adaptation to the work of Poli [12] could be made to provide not only a convergent parameter region, but rather a way of calculating the expected rate of particle convergence or divergence given a certain PSO parameter setting. Empirical investigations such as the one performed in this paper could be performed on complex PSO variants, where theoretical analysis has proved prohibitively complex.

## REFERENCES


[1] R. Poli, "Analysis of the publications on the applications of particle swarm optimisation," *Journal of Artificial Evolution and Applications*, vol. 2008, pp. 1–10, 2008.

[2] E. Ozcan and C. Mohan, "Analysis of a simple particle swarm optimization system," *Intelligent Engineering Systems through Artificial Neural Networks*, vol. volume 8, pp. 253–258, 1998.

[3] ——, "Particle swarm optimization: Surfing the waves," in *Proceedings of the IEEE Congress on Evolutionary Computation*, vol. volume 3. Piscataway, NJ, USA: IEEE Press, July 1999.

[4] M. Clerc and J. Kennedy, "The particle swarm-explosion, stability, and convergence in a multidimensional complex space," *IEEE Transactions on Evolutionary Computation*, vol. 6, no. 1, pp. 58–73, 2002.

[5] Y. Zheng, L. Ma, L. Zhang, and J. Qian, "On the convergence analysis and parameter selection in particle swarm optimization," in *Proceedings of the International Conference on Machine Learning and Cybernetics*, vol. 3, Xi'an, China, 2003, pp. 1802–1907.

[6] F. Van den Bergh and A. Engelbrecht, "A study of particle swarm optimization particle trajectories," *Information Sciences*, vol. 176, no. 8, pp. 937–971, 2006.

[7] I. Trelea, "The particle swarm optimization algorithm: Convergence analysis and parameter selection," *Information Processing Letters*, vol. 85, no. 6, pp. 317–325, 2003.

[8] C. Cleghorn and A. Engelbrecht, "A generalized theoretical deterministic particle swarm model," *Swarm Intelligence*, vol. 8, pp. 35–59, 2014.

[9] V. Kadirkamanathan, K. Selvarajah, and P. Fleming, "Stability analysis of the particle dynamics in particle swarm optimizer," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 3, pp. 245–255, 2006.

[10] V. Gazi, "Stochastic stability analysis of the particle dynamics in the PSO algorithm," in *Proceedings of the IEEE International Symposium on Intelligent Control*. Dubrovnik, Croatia: IEEE Press, 2012, pp. 708–713.

[11] R. Poli and D. Broomhead, "Exact analysis of the sampling distribution for the canonical particle swarm optimiser and its convergence during stagnation," in *Genetic and Evolutionary Computation Conference*. New York, NY, USA: ACM Press, 2007, pp. 134–141.

[12] R. Poli, "Mean and variance of the sampling distribution of particle swarm optimizers during stagnation," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 4, pp. 712–721, 2009.

[13] J. Kennedy and R. Eberhart, "Particle swarm optimization." Piscataway, NJ, USA: IEEE Press, 1995, pp. 1942–1948.

[14] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *Proceedings of the IEEE Congress on Evolutionary Computation*. Piscataway, NJ, USA: IEEE Press, 1998, pp. 69–73.

[15] J. Kennedy, "Small worlds and mega-minds: effects of neighborhood topology on particle swarm performance," in *Proceedings of the IEEE Congress on Evolutionary Computation*, vol. 3. Piscataway, NJ, USA: IEEE Press, 1999, pp. 1931–1938.

[16] J. Kennedy and R. Mendes, "Population structure and particle performance," in *Proceedings of the IEEE Congress on Evolutionary Computation*. Piscataway, NJ, USA: IEEE Press, 2002, pp. 1671–1676.

[17] E. Peer, F. Van den Bergh, and A. Engelbrecht, "Using neighborhoods with the guaranteed convergence PSO," in *Proceedings of the IEEE Swarm Intelligence Symposium*. Piscataway, NJ, USA: IEEE Press, 2003, pp. 235–242.

[18] A. Engelbrecht, "Particle swarm optimization: Global best or local best," in *1st BRICS Countries Congress on Computational Intelligence*. Piscataway, NJ, USA: IEEE Press, 2013.

[19] F. Van den Bergh, "An analysis of particle swarm optimizers," *PhD thesis, Department of Computer Science, University of Pretoria, Pretoria, South Africa*, 2002.

[20] B. Kisacanin and G. Agarwal, *Linear Control Systems: With Solved Problems and Matlab Examples*. New York, NY, USA: Springer, 2001.

[21] A. Engelbrecht, "Roaming behavior of unconstrained particles," in *1st BRICS Countries Congress on Computational Intelligence*. Piscataway, NJ, USA: IEEE Press, 2013.