Particle Swarm Optimizer: The Impact of Unstable Particles on Performance

Christopher W Cleghorn  
Department of Computer Science  
University of Pretoria  
Pretoria, South Africa  
Email: ccleghorn@cs.up.ac.za

Andries Engelbrecht  
Department of Computer Science  
University of Pretoria  
Pretoria, South Africa  
Email: engel@cs.up.ac.za

Abstract—There exists a wealth of theoretical analysis on particle swarm optimization (PSO), specifically the conditions needed for stable particle behavior are well studied. This paper investigates the effect that the stability of the particle has on the PSO’s actually ability to optimize. It is shown empirically that a majority of PSO parameters that are theoretically unstable perform worse than a trivial random search across 28 objective functions, and across various dimensionalities. It is also noted that there exists a number of parameter configurations just outside the stable-2 region which did not exhibit poor performance, implying that a minor violation of the conditions for order-2 stability is still acceptable in terms of overall performance of the PSO.

I. INTRODUCTION

Particle swarm optimization (PSO) is a stochastic population-based search algorithm developed by Kennedy and Eberhart [1]. PSO has been effectively utilized to solve numerous real world optimization problems, a summary of which can be found in the work of Poli [2].

PSO has undergone a substantial amount of theoretical analysis, the majority of which focus on the conditions that are both necessary and sufficient for particle stability. Some of the more recent works are [3], [4], [5], [6], [7], [8]. Currently the conditions needed for order-1 and order-2 stability are well known. Order-1 and order-2 stability are respectively defined as convergence in expectation of the particle’s position and convergence of the particle’s variance.

This paper investigates the effect of selecting parameters that theoretically yield unstable particle behavior (non order-2 behavior). This paper considers a wide range of objective functions across various dimensionalities. Performance is also considered across differing iteration counts.

A description of PSO is given in section II. Section III contains an discussion about the theoretical PSO results directly relevant to this paper. The experimental setup is presented in section IV, followed by the experimental results and a discussion thereof in section V. Section VI presents a summary of the findings of this paper.

II. PARTICLE SWARM OPTIMIZER

Particle swarm optimization (PSO) was originally developed by Kennedy and Eberhart [1] to simulate the complex movement of birds in a flock. The standard variant of PSO this section focuses on includes the inertia coefficient proposed by Shi and Eberhart [9].

The PSO algorithm is defined as follows: Let \( f : \mathbb{R}^d \rightarrow \mathbb{R} \) be the objective function that the PSO algorithm aims to find an optimum for, where \( d \) is the dimensionality of the objective function. For the sake of simplicity, a minimization problem is assumed from this point onwards. Specifically, an optimum \( o \in \mathbb{R}^d \) is defined such that, for all \( x \in \mathbb{R}^d \), \( f(o) \leq f(x) \). The analysis of this paper focuses on objective functions where the optima exist. Let \( \Omega(t) \) be a set of \( N \) particles in \( \mathbb{R}^d \) at a discrete time step \( t \). Then \( \Omega(t) \) is said to be the particle swarm at time \( t \). The position \( x_i \) of particle \( i \) is updated using

\[
x_i(t + 1) = x_i(t) + v_i(t + 1),
\]

where the velocity update, \( v_i(t + 1) \), is defined as

\[
v_i(t + 1) = w v_i(t) + c_1 r_1(t) \odot (y_i(t) - x_i(t)) + c_2 r_2(t) \odot (y_j(t) - x_i(t)),
\]

where \( r_{1,j}(t), r_{2,j}(t) \sim U(0, 1) \) for all \( t \) and \( 1 \leq j \leq k \). The operator \( \odot \) is used to indicate component-wise multiplication of two vectors. The position \( y_i(t) \) represents the “best” position that particle \( i \) has visited, where “best” means the location where the particle had obtained the lowest objective function evaluation. The position \( y_j(t) \) represents the “best” position that the particles in the neighborhood of the \( i \)-th particle have visited. The coefficients \( c_1, c_2, \) and \( w \) are the cognitive, social, and inertia weights, respectively.

A primary feature of the PSO algorithm is social interaction, specifically the way in which knowledge about the search space is shared amongst the particles in the swarm. In general, the social topology of a swarm can be viewed as a graph, where nodes represent particles, and the edges are the allowable direct communication routes. The social topology chosen has a direct impact on the behaviour of the swarm as a whole [10], [11], [12]. Some of the most frequently used are the star, ring, and Von Neumann topologies.

III. THEORETICAL CONSIDERATIONS

This section presents a brief description of the theoretical PSO results that are directly related to this paper are pre-
Algorithm 1 PSO algorithm

Create and initialize a swarm, \( \Omega(0) \), of \( N \) particles uniformly within a predefined hypercube red of dimension \( k \). Let \( f \) be the objective function.
Let \( y_i \) represent the personal best position of particle \( i \), initialized to \( x_i(0) \).
Let \( y_i \) represent the neighborhood best position of particle \( i \), initialized to \( x_i(0) \).
Initialize \( v_i(0) \) to 0.

\[
\begin{align*}
\text{repeat} & \\
& \text{for all particles } i = 1, \ldots, N \text{ do} \\
& \quad \text{if } f(x_i) < f(y_i) \text{ then} \\
& \quad \quad y_i = x_i \\
& \quad \text{end if} \\
& \quad \text{for all particles } i \text{ with particle } i \text{ in their nhb do} \\
& \quad \quad \text{if } f(y_i) < f(y_j) \text{ then} \\
& \quad \quad \quad y_j = y_i \\
& \quad \text{end if} \\
& \quad \text{end for} \\
& \text{end for} \\
& \text{for all particles } i = 1, \ldots, N \text{ do} \\
& \quad \text{update the velocity of particle } i \text{ using equation (2)} \\
& \quad \text{update the position of particle } i \text{ using equation (1)} \\
& \text{end for} \\
\text{until} \text{ stopping condition is met}
\end{align*}
\]

The technical detail would detract from the discussion. Now, unwinding of equation (3) leads to

\[
z_t = M^{n-1}z_1 + \sum_{j=0}^{n-2} M^j b
\]

Since \( M \) is symmetric it is possible to represent any \( z_1 \) as a weighted sum of \( M \)'s eigenvectors, specifically \( z_1 = \sum_{i=0}^{q} \eta_i e_i \), where \( e_i \) are the eigenvectors of \( M \), which have the corresponding eigenvalues \( \lambda_i \). Consider the first term of equation (5):

\[
M^{n-1}z_1 = M^{n-1} \sum_{j=0}^{q} \eta_j e_j = \sum_{j=0}^{q} \eta_j \lambda_i^{n-1} e_i
\]

Since \( \rho(M) > 1 \), at least one eigenvalue of \( M \) is greater than 1 so at least one term in the summation in equation (6) will diverge (assuming the corresponding \( \eta_i \neq 0 \)). Not only will the term diverge, but it will do so exponentially.

The exponential divergence is very important to note as small increases in \( \rho(M) \) could drastically increase the long term particle trajectory. The immediate question is how tolerant the PSO is to having control parameters selected that have slightly larger than 1 spectral radius. It would be ideal to derive the required conditions for different spectral radius's of \( M \). While the criteria of equation (4) appear relatively simple, they are not directly derivable from the condition \( \rho(M) < 1 \) even with assistance of a symbolic solver. Instead, necessary conditions where derived using a technique first used by Blackwell [5], and then empirically verified to hold for \( \rho(M) < 1 \) [8]. This approach was needed as the individual eigenvalues of \( M \) require over a 1000 characters to express. Unfortunately, this indicates that finding the explicit condition for \( \rho(M) < s \), where \( s \) is any non-negative number is intractable.

IV. E MPirical Setup

This section summarizes the experimental procedure used for this paper. The experiment aims to illustrate that most unstable parameter configurations actually result in such poor...
performance of PSO that a random search can outperform them.

The performance of the PSO was measured for each parameter configuration across the following region:

$$ w \in [-1.1, 1.1] \text{ and } c_1 + c_2 \in (0, 5.5] $$ \hspace{1cm} (7)

where step sizes of 0.1 were used for $w$ and $c_1 + c_2$. This results in 1264 parameter configurations, of which 761 are unstable and 504 are stable according to equation (4). A fully connected star neighborhood topology was used. Velocities where initialized to 0. A population size of 20 was used. The results for each configuration were derived from 35 independent runs.

The performance of PSO is compared to that of a random search, with the premise that if a specific configuration of PSO does worse than a random search it is not effectively optimizing. The random search algorithm used, samples uniformly within a given objective function’s defined domain. For the sake of comparison each iteration of the PSO algorithm is seen as comparable to 20 random samples of the search space, one for each of the particles in the PSO swarm.

The objective functions used in this paper used are presented in table I. Full definitions of the objective functions can be found in following works, [14], [15], [16]. Each objective function is tested in 5, 10, and 30 dimensions. The performance is measured at 500, 1000, 2000, and 5000 iterations.

V. Empirical Results and Discussion

This section presents the results of the experiments described in section IV.

For each parameter configuration Mann-Whitney U tests, using a confidence level of 95%, were performed to determine, whether a given parameter configuration resulted in a PSO that was in fact better or worse than a random search, or if the PSO showed no statistical difference with the performance of the random search. This information is summarized into four categories for each test case:

- **CP_BR**: The percentage of parameters that are theoretically stable that resulted in the PSO performing better than random search.
- **CP_NDR**: The percentage of parameters that are theoretically stable that resulted in the PSO performing with no statistical difference to random search.
- **DP_BR**: The percentage of parameters that are theoretically unstable that resulted in the PSO performing better than random search.
- **DP_NDR**: The percentage of parameters that are theoretically unstable that resulted in the PSO performing with no statistical difference to random search.

For each dimensionality tested, a table summarizing the performance information is given. In which the performance results for 500, 1000, 2000, and 5000 iterations are reported.

The first thing to note is that, over all the results shown in tables II, III, and IV for 5, 10, and 30 dimensions respectively, the percentage of unstable parameter choices that were able to outperform random search were significantly low. The highest percentage of unstable parameter configurations able to outperform random search was 34.67% for Griewank in 5 dimensions, at 500 iterations. The performance on Griewank decreases with a increase in iteration count, down to 27.2% at 5000 iterations. The exact parameter configurations failing to beat random search can be seen clearly in figure 2. What is interesting is that all the unstable parameter configurations that did in fact outperform random search are in a region which appears to be a natural extension of the stability boundary of equation (4), specifically a region with a slightly larger spectral radius. Similar results can be seen for both Rosenbrock and Solomon in 5-dimensions at 5000 iterations as illustrated figures 4 and 3 respectively.

It might appear as if there is a fair degree of tolerance on the stability boundary of equation (4). However, a quick scan of tables II, III, and IV, shows numerous cases where the performance of unstable parameters is terrible, for example, only 6.18% of unstable parameters where able to outperform random search on Michalewicz in 30 dimensions at 1000 iterations, whereas 100% of stable parameter configurations outperformed random search. It can be seen in figure 5 that there is nearly a perfect relationship between convergent parameter configurations and the ability of the PSO to outperform random search when optimizing Michalewicz.

There are actually two objective functions that seem to illustrate how finely tuned PSO sometimes needs to be,
namely Egg Holder and Elliptic. Over all tested dimensions even stable parameter configurations were more often than not outperformed by random search. This requirement of fine tuning can be seen in figure 6 for the Elliptic objective function in 5 dimensions at 5000. What is worth noting in figure 6 is that there are some unstable parameter configurations that outperform some stable parameters. However, all unstable parameter configurations that were successful are still near the apex of the stable region of equation (4). The performance of PSO was very poor for Egg Holder, with only less than 12% of parameter configurations (stable and unstable) outperforming random search in 5 dimensions at 500 iterations, and decreasing to less than 11%. Looking at figure 7 it is very clear that only a very small number of parameter configurations are effective at optimizing the Egg Holder objective function. But, what is interesting is that once again the optimal parameter configurations appear to be clustered around the boundary of the stable region, however in the case of the Egg Holder objective function most are just slightly outside the stable region.

There is a very clear trend throughout all the data, namely that stable parameter configurations improve as the dimensionality of the problem increases. This behavior is observed on all tested objective functions except of the Happy Cat objective function. This is not surprising since a random search’s performance is likely to degrade quickly as dimensionality increase. However, what is interesting is that the performance across nearly all the objective functions worsens for unstable parameter configurations. This implies that the higher the dimensionality, the more important selecting stable parameter configurations becomes.

VI. CONCLUSION

It was shown that the majority of parameter configurations that are theoretically unstable perform worse than random search on all objective functions tested. It was also shown that there is a degree of tolerance from which parameters can be selected just outside of the convergent region without extreme performance degradation. However, the degree to which parameter values can be selected outside of the stable region is very problem dependent. For most tested

Fig. 2: Griewank, 5-dimensions, 5000 iterations
1 = performed better than random search, 2 = no statistical difference, 3 = random search performed better

Fig. 3: Salomon, 5-dimensions, 5000 iterations
1 = performed better than random search, 2 = no statistical difference, 3 = random search performed better

Fig. 4: Rosenbrock, 5-dimensions, 5000 iterations
1 = performed better than random search, 2 = no statistical difference, 3 = random search performed better
cases it was found that convergent parameter configurations drastically increased the chance of outperforming random search. However, selecting parameter values within the stable region near the apex was the best strategy to ensure that PSO was always superior to random search. It was also observed that the higher the dimensionality of the problem, the more important selecting stable parameter configurations became.

Future work could include an investigation into defining a region between two spectral radius values that encompasses the region around the apex where the PSO outperforms random search in all the tested objective function. While this region may be problem dependent to a degree, there appears
Fig. 7: Egg Holder, 5-dimensions, 5000 iterations

1 = performed better than random search, 2 = no statistical difference, 3 = random search performed better

to be enough of a pattern to warrant further investigation.

REFERENCES


<table>
<thead>
<tr>
<th></th>
<th>500 Iterations</th>
<th>1000 Iterations</th>
<th>2000 Iterations</th>
<th>5000 Iterations</th>
</tr>
</thead>
<tbody>
<tr>
<td>F1</td>
<td>100.00</td>
<td>30.75</td>
<td>4.73</td>
<td>100.00</td>
</tr>
<tr>
<td>S2</td>
<td>100.00</td>
<td>28.38</td>
<td>6.44</td>
<td>100.00</td>
</tr>
<tr>
<td>F3</td>
<td>100.00</td>
<td>6.70</td>
<td>1.05</td>
<td>100.00</td>
</tr>
<tr>
<td>F4</td>
<td>100.00</td>
<td>17.86</td>
<td>2.76</td>
<td>100.00</td>
</tr>
<tr>
<td>F5</td>
<td>100.00</td>
<td>7.23</td>
<td>2.73</td>
<td>100.00</td>
</tr>
<tr>
<td>S6</td>
<td>52.18</td>
<td>7.34</td>
<td>2.76</td>
<td>100.00</td>
</tr>
<tr>
<td>F7</td>
<td>99.80</td>
<td>9.20</td>
<td>3.55</td>
<td>99.80</td>
</tr>
<tr>
<td>F8</td>
<td>100.00</td>
<td>23.39</td>
<td>6.83</td>
<td>100.00</td>
</tr>
<tr>
<td>F9</td>
<td>100.00</td>
<td>22.86</td>
<td>6.70</td>
<td>100.00</td>
</tr>
<tr>
<td>F10</td>
<td>67.26</td>
<td>10.32</td>
<td>4.34</td>
<td>67.26</td>
</tr>
<tr>
<td>F11</td>
<td>100.00</td>
<td>31.27</td>
<td>4.99</td>
<td>100.00</td>
</tr>
<tr>
<td>F12</td>
<td>100.00</td>
<td>16.16</td>
<td>6.96</td>
<td>100.00</td>
</tr>
<tr>
<td>F13</td>
<td>100.00</td>
<td>5.91</td>
<td>2.76</td>
<td>100.00</td>
</tr>
<tr>
<td>F14</td>
<td>100.00</td>
<td>5.91</td>
<td>2.76</td>
<td>100.00</td>
</tr>
<tr>
<td>F15</td>
<td>99.80</td>
<td>9.20</td>
<td>3.55</td>
<td>99.80</td>
</tr>
<tr>
<td>F16</td>
<td>100.00</td>
<td>23.39</td>
<td>6.83</td>
<td>100.00</td>
</tr>
<tr>
<td>F17</td>
<td>100.00</td>
<td>22.86</td>
<td>6.70</td>
<td>100.00</td>
</tr>
<tr>
<td>F18</td>
<td>100.00</td>
<td>18.27</td>
<td>5.44</td>
<td>100.00</td>
</tr>
<tr>
<td>F19</td>
<td>100.00</td>
<td>27.86</td>
<td>5.78</td>
<td>100.00</td>
</tr>
<tr>
<td>F20</td>
<td>100.00</td>
<td>31.27</td>
<td>4.99</td>
<td>100.00</td>
</tr>
<tr>
<td>F21</td>
<td>99.80</td>
<td>4.20</td>
<td>2.50</td>
<td>99.80</td>
</tr>
<tr>
<td>F22</td>
<td>100.00</td>
<td>22.86</td>
<td>5.65</td>
<td>100.00</td>
</tr>
<tr>
<td>F23</td>
<td>100.00</td>
<td>17.61</td>
<td>6.44</td>
<td>100.00</td>
</tr>
<tr>
<td>F24</td>
<td>75.20</td>
<td>19.64</td>
<td>3.42</td>
<td>75.20</td>
</tr>
<tr>
<td>F25</td>
<td>100.00</td>
<td>30.75</td>
<td>6.31</td>
<td>100.00</td>
</tr>
<tr>
<td>F26</td>
<td>100.00</td>
<td>31.41</td>
<td>5.65</td>
<td>100.00</td>
</tr>
<tr>
<td>F27</td>
<td>78.97</td>
<td>17.26</td>
<td>5.65</td>
<td>78.97</td>
</tr>
<tr>
<td>F28</td>
<td>100.00</td>
<td>26.28</td>
<td>7.23</td>
<td>100.00</td>
</tr>
</tbody>
</table>

TABLE IV: Performance of PSO versus Random Search: 30-Dimensions

---

