Diversity Preservation Using Excited Particle Swarm Optimisation

Shannon S. Pace
Faculty of ICT
Swinburne University of Technology
space@swin.edu.au

Clinton J. Woodward
Faculty of ICT
Swinburne University of Technology
cwoodward@swin.edu.au

ABSTRACT

The particle swarm optimisation (PSO) algorithm suffers from the possibility of premature convergence. This problem has historically been addressed *ab intra* – manipulating velocity and swarm topology – yet the judicious addition of external mechanisms has been shown to adjust search behaviour to yield significantly improved results across many problems. This paper introduces an addition to the canonical particle swarm algorithm, designed to preserve the diversity typically lost by attraction to suboptimal positions. The proposed excited PSO method stimulates exploration upon the discovery of a candidate solution by manipulating the position to which particles are attracted. It is shown to maintain a suitable degree of diversity for the duration of an experiment, as well as an ability for self-scaling. Comparisons to the canonical PSO algorithm demonstrate improved solutions in both unimodal and multimodal spaces.

Track

Ant Colony Optimization and Swarm Intelligence

Categories and Subject Descriptors

I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods, and Search—*Heuristic Methods*

General Terms

Algorithms

Keywords

Particle swarm optimisation, Preserving diversity, Stimulating exploration, Excited search

1. INTRODUCTION

The particle swarm optimisation (PSO) algorithm has the ability to dynamically scale the range of a swarm's search [9]. Its capacity to perform widespread exploration, yet eventually converge and exploit a found optimum, within suitable parameters [12, 2], has made it applicable to a large number of problems. However, the possibility exists for the swarm to converge on a local optimum prematurely. The particle swarm has no precaution against such an event and a poor solution may result. PSO has undergone significant research regarding the role of parameters [13] and network topology [7, 10] so as to determine search behaviour that may ensure suitable exploration occurs before convergence. Even so, the tendency for a swarm to rapidly lose diversity has not been sufficiently addressed.

The dynamics responsible for search in the canonical PSO algorithm can be considered a negative feedback mechanism with respect to diversity. The relationship between particle velocity and said particle's distance to certain positions may excessively limit the region in which it may freely travel. Diversity can be engineered into a swarm through suitably restrictive interaction between particles via topologies [7]. Nevertheless, a small number of well performing particles have the potential to bias swarm search in a particular region. Typical diversity preserving mechanisms seek to maintain a large portion of the former diversity of a swarm by insulating against changes in velocity or provincialising candidate solution information. Even so, they are often insufficient for offsetting the rate of diversity loss in PSO, and it is often difficult for a swarm to recover from the ensuing converged state. A viable solution may be to intentionally but judiciously introduce diversity during search.

This paper presents the *excited* PSO algorithm (EPSO) which stimulates exploration slightly upon the discovery of a candidate solution. Excitation occurs through the manipulation of *lbest* – the best known position attributed to a swarm neighbourhood. Specifically, the position is projected a prescribed distance in the direction of locally improving fitness. To avoid persistently deceiving particles, should the stimulated exploration prove fruitless, the predicted position is made more similar to the actual *lbest* as more iterations pass since its discovery. It is anticipated that the excitation will stimulate diversity such that the prospect of premature convergence is reduced, while continuing the typical convergence characteristics and self-scaling ability of canonical PSO.

To confirm the desired effect of EPSO it is compared to canonical PSO in three high-dimension problems. The experiment consists of multiple experiments in each algorithm-problem combination, performed under identical conditions except for the *lbest* determination method. In addition to

GECCO'11, July 12–16, 2011, Dublin, Ireland.

the final result quality, the fitness and diversity of each swarm is captured per iteration for the duration of experiments. Results are discussed and form the basis of a hypothesis concerned with the dynamics of EPSO.

This paper first considers existing, related techniques before introducing the EPSO algorithm. The effect of the inertia weight, constriction factor and swarm topology on diversity preservation are discussed. The experiment is then detailed, including justification for the selection of objective functions. In the Results section, algorithm performance is compared and analysed. A discussion then articulates the strengths and weaknesses of EPSO compared to canonical PSO, and considers the search dynamics that contribute to the difference in performance. The paper concludes by suggesting methods by which the algorithm may be improved, in light of the findings.

2. BACKGROUND

2.1 The PSO Algorithm

A particle's search behaviour in canonical PSO is informed by two positions, *lbest* and *pbest* [8, 13]. The highest fitness position discovered by a particle p is referred to as *pbest*, while *lbest* is typically the highest fitness position discovered by p's neighbourhood – the particles with which p shares *pbest* information. Commonly, *lbest* is replaced with *gbest*, the highest fitness *pbest* yet discovered by the swarm. The path traced through the problem space by a particle is a weighted stochastic sum of the vectors pointing towards the *lbest* and *pbest* positions, referred to as the *social* and *cognitive* components of velocity respectively. A swarm's objective is to find the ideal *pbest*.

The velocity $\vec{v}$, at time t, generating the motion of a particle p is equal to:

$$\vec{v}_i^{\,p}(t) = \omega\vec{v}_i^{\,p}(t-1) + \vec{m}_i(t) \tag{1}$$

$$\vec{m}_i(t) = \phi_s r_s \vec{s}_i(t) + \phi_c r_c \vec{c}_i(t) \tag{2}$$

$$\vec{s}_i(t) = \vec{g}_i^{\,p} - \vec{x}_i^{\,p}(t) \tag{3}$$

$$\vec{c}_i(t) = \vec{b}_i^{\,p} - \vec{x}_i^{\,p}(t) \tag{4}$$

applied to each vector dimension i, where ω is the inertia weight introduced by Shi and Kennedy [13]. The position of p is denoted by $\vec{x}$, and $\vec{g}$ and $\vec{b}$ are p's *lbest* and *pbest* positions respectively. Parameter ϕ is the prescribed component weight of the social s and cognitive c influences respectively, and r is a random number in the range [0, 1], also associated with said components. Note that new r values are generated for each vector element transformation. The social and cognitive components of velocity are $\vec{s}$ and $\vec{c}$ respectively. Velocity $\vec{v}$ is added to the position vector $\vec{x}$ of the concerned particle at each iteration. As such, a unit of time typically corresponds to an iteration.

2.2 Controlling Search Behaviour

When a particle is distant to *pbest* and *lbest* (which themselves are proximate), $\vec{c} \cdot \vec{s} \to 1$. In the velocity calculation Equation (1) the vectors "cooperate" to drag the particle into a region R, in which $\vec{c} \cdot \vec{s} < 0$. Within R, and in moving towards *pbest* and *lbest*, both $|\vec{c}|$ and $|\vec{s}|$ become smaller. Thus a particle slows down and the chance that the next found *pbest* will be local to the previous is increased. Furthermore, the $|\vec{s}|$ of neighbour particles will not be significantly affected by the new discovery. The reduced velocity of individual particles and decreased diversity of swarms in this manner increasingly confines the region in which swarms search. In the extreme case, R is recursively subdivided by each new *pbest* discovery.

Fortunately the inertia weight setting ω and the swarm topology employed can partially mitigate the tendency to search within increasingly small regions. A high ω significantly preserves the velocity of a particle, and thus it may escape R by a greater distance and for longer periods, rather than immediately converging on the region. Inertia weight, introduced by Shi and Kennedy, has been found to improve exploration characteristics [13] and is a simple means to prevent greediness. However, the velocity of a particle is still largely dependent on the magnitude of $\vec{c}$ and $\vec{s}$ as determined in Equations (3) and (4). Furthermore, the inertia weight may prolong periods of high velocity, but makes particles less able to recover once aggressive exploitation begins.

Kennedy and Shi have demonstrated that the constriction factor [2] allows swarms to easily switch between exploration and exploitation, but the algorithm has been found more susceptible to premature convergence than the judicious application of inertia weights and capped velocity [9]. Time to convergence is reduced, but that swarms are more likely to converge on suboptimal peaks suggests greediness. It should be seen that scaling velocity, by either the inertia weight or constriction factor, as a means to preserve diversity is only a partial solution. The PSO algorithm has shown itself more than capable in rapidly reducing $|\vec{c}|$ and $|\vec{s}|$, and scaling said vectors does not sufficiently mitigate this behaviour.

The swarm topology employed can help preserve diversity, but this capacity is tempered by the possibility for a less informed search. When the network topology is such that the particles in any given connected pair, i and j, have dissimilar neighbourhoods, all particles experience conflicting *lbest* advertisements. For i, an *lbest* originating from j (the location of which will have been influenced by j's neighbours, of which i has no knowledge) will be tempered by an *lbest* inevitably advertised by another member of i's neighbourhood – the particles of which, in turn, have not been directly influenced by j. The diversity arising from swarm topology, then, is perhaps better referred to as competition. A suitable topology prevents any one *pbest* from dominating the search. It is only when an optimum in the problem space persistently yields *pbests* for a number of particles that the swarm is sufficiently influenced to converge upon it.

Mendes and Kennedy have demonstrated that, in employing network topologies, there is a tradeoff between the utilisation of collective intelligence and the preservation of diversity – determined by the sharing and provincialisation of *lbest* information respectively [10]. The authors note that the toroidal Von Neumann network topology, notably a structure affording equally limited social influence among particles, yields consistently high performance across many objective problems. Diversity arising from such a network is to be expected. Even so, it should be anticipated that one family of spatially related *lbests* will eventually come to prominence. Subsequent solutions will be biased to that region, and the swarm will inevitably experience a rapid loss of diversity in future iterations. While this convergence is ultimately desirable, the rate at which it occurs is potentially detrimental.

Each of the methods discussed in this section maintain diversity in an unpredictable manner. They also tend to prolong the desirable, early configurations of particles – which are often associated with high mean velocity and spatial diversity – but do little to offset the possible negative feedback loop that can aggressively remove diversity, causing swarm collapse. It seems apparent that methods for intentionally introducing diversity to swarms may be utilised to prevent such outcomes, yet such techniques must not excessively interfere with desired swarm search behaviour. Such methods should be compatible with the strengths of PSO, permitting convergence, self-scaling and search efficiency.

3. EXCITED PSO (EPSO)

Algorithms such as the fully-informed particle swarm [11] have demonstrated improved search ability via manipulation of the *lbest* position. Given that this position influences the magnitude of particle velocity, and thus swarm diversity, judiciously modifying *lbest* may contribute to reduced incidence of premature convergence. The excited PSO algorithm attempts to effect this outcome by stimulating particles when it is apparent that they are within a high fitness region, so as to mitigate stagnation. This section describes how EPSO determines an *lbest* position by which excitation is achieved.

In the velocity update of a particle p, its current *lbest* ($\vec{g}$ in Equation (3)) is replaced with the position at which the next *pbest* of the reporting particle q is expected to be found. This position $\vec{e}$, as it applies to particle p at time t_c, is determined by:

$$\vec{e}(t_c) = \vec{b}_n^{\,q} + d \cdot (1 - \frac{t_c - t_u}{g})^a \cdot \vec{u}^{\,q}, \;\; t_c - t_u \leq g \qquad (5)$$

where

$$\vec{u}^{\,q} = \vec{b}_n^{\,q} - \vec{b}_{n-1}^{\,q} \qquad (6)$$

where b is the index set of *pbest* positions known to the particle denoted by the superscript, of which n is the most recently added element. The iteration at which $\vec{b}_n^{\,q}$ was discovered is noted by t_u, while g, d and a are prescribed values. Note that q is simply the particle whose *pbest* represents *lbest* for p, and so p and q are neighbours, yet p may also be made equal to q if desired.

As $\vec{e}$ is not the actual *lbest*, failing to improve its accuracy for prolonged periods could potentially misinform search. Parameter a and g both influence the rate at which $\vec{e}$ becomes more similar to $\vec{b}_n^{\,q}$: the former determines the degree to which $\vec{e}$ becomes similar to $\vec{b}_n^{\,q}$ as *lbest* ages, the period of which is determined by the latter parameter g. The magnitude of d, however, determines the magnitude of excitation. As such, it is likely that EPSO search performance is more sensitive to d than a and g.

Position $\vec{e}$ may actually be a low fitness position for much of the period described by g. It may, for example, overshoot an optimum, or the direction of stimulation $\vec{u}^{\,q}$ may be tangential to an optimum. Even so, though EPSO stimulates velocity in the direction of $\vec{e}$, it does not "scout ahead" at that position. The benefit of $\vec{e}$ is anticipated to be in the holistic exploration stimulated in the region local to q's *pbest*, inasmuch as the position suggests a region of high fitness.

Table 1: PSO parameter settings, corresponding to Equations (1) and (2), for all objective functions.

Parameter	Value
ϕ_s	2.0
ϕ_c	2.0
ω	0.9
Particles	36
Network topology	*lbest* [11]
Iterations	9,000

Table 2: EPSO parameter settings, corresponding to Equation 5, for each objective function.

	d	g	a
F1	2.5	45	1
F2	2	45	1
F3	1	45	1

4. EXPERIMENT

4.1 Overview

EPSO was compared to canonical PSO across three objective functions, all in 100 dimensions:

- Ackley's function [1] (F1):

$$f(x) = \left(-20 \cdot \exp(-0.2 \cdot \sqrt{\frac{1}{n}\sum_{i=1}^{n} x_i^2}) - \exp(\frac{1}{n}\sum_{i=1}^{n}\cos(2\pi x_i)) + 20 + \exp(1) \right)$$

where $x \in [-32.768, 32.768]$.

- Griewangk's function [3] (F2):

$$f(x) = \frac{1}{4000}\sum_{i=1}^{n} x_i^2 - \prod_{i=1}^{n}\cos(\frac{x_i}{\sqrt{i}}) + 1$$

where $x \in [-600, 600]$.

- Axis-parallel hyperellipsoid function (F3):

$$f(x) = \sum_{i=1}^{n}[i \cdot x_i^2], \; x \in [-5.12, 5.12]$$

Note that the subscript of x denotes the ith vector element. In each function the global optimum of zero resides at the zero vector. As minimisation problems, high fitness corresponds to low function evaluation values. The experiment consisted of 100 experiments on each objective function with both the PSO and EPSO algorithms. When initialising swarms, particles were randomly placed within the x domain of each problem space with zero velocity. Parameters common between the algorithms are detailed in Table 1. The parameters specific to EPSO for each objective function are detailed in Table 2. Also note that "self" was included in the *lbest* determination method of EPSO. That is, in Equation (5), q also responds to its own *pbest* discovery — $p = q$ in such cases.

During the course of experiments it was found that the PSO variant most successful across all objective functions is that described by Hendtlass [4], where the sum of social and cognitive components of the velocity vector are modified such that:

$$\vec{v}(t) = \omega\vec{v}(t-1) + (1-\omega)\cdot\vec{m}(t)$$

where $\vec{m}$ is described by Equation (2). This version of the PSO algorithm was used for the experiments.

It should be noted that the selection of parameters for PSO and EPSO is not exhaustive, and more successful parameters, for both algorithms, may exist. The raw performance of the algorithms is not as relevant as their comparative performance for this investigation. Note that, as EPSO is simply PSO with a modified *lbest* position, any significant difference in results can be attributed to the effect EPSO has on search behaviour.

In addition, *lbest*, *gbest* and Von Neumann topologies were tested. Only the results for *lbest* are presented, as PSO was found to be most successful with this swarm topology on the objective functions considered. A duration of 9000 iterations was found to be adequate to differentiate EPSO and PSO performance.

4.2 Objective Problem Selection

Ackley's function is deceptive to algorithms prone to greediness, as attractive local optima may prevent the discovery of the global optimum. The relative steepness of the local optima exacerbate this effect. This function should be considered a specific test of the effectiveness of EPSO diversity excitation in mitigating premature convergence.

The Griewangk function significantly increases in difficulty during fine-grained search, as the influence of the summation component drops near the global optimum. As such, the quality of solutions in this problem are dependent on the algorithm's ability to suitably maintain particle vigour during exploitation. Diversity stimulation methods risk not being able to effectively scale their influence to permit such behaviour.

The hyperellipsoid function is unimodal and thus premature convergence is not possible. Given the relative simplicity of the function, it is not arrogant to expect the optimum be found quickly. However, the potential still exists for particle velocity to drop significantly should mutual attraction reduce the magnitude of the social component of velocity, and thus particle speed. Slow search may result.

5. RESULTS AND OBSERVATIONS

5.1 Raw Performance

Tables 3 and 4 present the final fitness results for EPSO and PSO respectively, where Q represents a quartile denoted by the subscript. Immediately apparent is premature convergence in the PSO F1 algorithm-function pair. Note that such behaviour is not demonstrated on any problem by EPSO. EPSO also appears more capable in performing fine-grained search, evidenced by the difference in F2 and F3 solution quality. Also notable is the size of the quartile range ($Q_3 - Q_1$) for each algorithm-function pair with respect to the fitness associated with a satisfactory solution. This is most conspicuous in PSO F1 data, where poor solution quality suggests a high degree of greediness in the canonical PSO algorithm.

The number of iterations each algorithm took to yield the presented solutions should also be considered. Tables 5 and 6 present the iterations required for each algorithm-problem pair to reach the median fitness values presented in tables 3 and 4. In particular, note the $\frac{Q_3-Q_1}{Q_2}$, Θ and Δ results.

Table 3: Final fitness values for EPSO on each function

	Q_1	Q_2	Q_3	Mean
F1	$1.48e^{-3}$	$1.76e^{-3}$	$2.45e^{-3}$	$2.02e^{-3}$
F2	$6.26e^{-10}$	$1.37e^{-9}$	$3.87e^{-9}$	$9.47e^{-4}$
F3	$1.16e^{-11}$	$2.61e^{-11}$	$6.21e^{-11}$	$1.01e^{-10}$

Table 4: Final fitness values for canonical PSO on each function

	Q_1	Q_2	Q_3	Mean
F1	2.63	2.88	3.13	2.86
F2	$2.93e^{-5}$	$9.94e^{-3}$	$3.97e^{-2}$	$4.39e^{-2}$
F3	$1.31e^{-4}$	$4.21e^{-4}$	$1.42e^{-3}$	$5.52e^{-3}$

Value $\frac{Q_3-Q_1}{Q_2}$ suggests the degree of consistency in search, for which high values correspond to low consistency and *vice versa*. The consistently high values of PSO in this metric suggest greediness; convergence occurs rapidly after inevitably encountering a local optimum, for which F1 is contrived. The only problem in which this doesn't occur to a noticeable degree in PSO is F3, yet this problem has a single optimum. As such, the comparison between PSO's F1 and F3 performance highlights the performance characteristics associated with greediness. If high consistency corresponds to marginalised greediness, then EPSO swarms can be said to enjoy considerable moderation in all problems. The consistency of fitness solutions and the time taken to yield them suggests EPSO is generally resilient to premature convergence.

The Θ column in Table 5 represents the iteration in which the median of EPSO swarms' fitness, measured per iteration, surpassed that of PSO swarms. Note that no EPSO swarm was subsequently overtaken by a PSO swarm after the Θ iteration was passed. In each problem it can be seen that search takes a greater period than Θ iterations to yield a satisfactory solution, thus EPSO search is a more efficient algorithm under the experiment conditions. In this context, efficiency refers to the rate of fitness improvement over iterations.

Another insight into relative search efficiency is the comparison of the median number of iterations each algorithm took to reach PSO's median fitness value. The ratio of these values is presented in the Δ column of Table 5. For F2 and

Table 5: Iterations to median fitness, averaged over the EPSO experiments. Θ: iteration in which mean fitness of EPSO swarms first exceeded that of PSO swarms. Δ: ratio of the median number of iterations taken by EPSO to that of PSO to reach PSO's median fitness solution.

	Q_1	Q_2	Q_3	Mean	$\frac{Q_3-Q_1}{Q_2}$	Θ	Δ
F1	8712	8843	8898	8789	0.02	1626	1.18
F2	8471	8672	8835	8567	0.04	982	0.65
F3	8599	8771	8873	8720	0.03	572	0.51

Table 6: Iterations to median fitness, averaged over the PSO experiments

	Q_1	Q_2	Q_3	Mean	$\frac{Q_3-Q_1}{Q_2}$
F1	1040	1434	2373	2022	0.93
F2	4337	5572	6382	5437	0.37
F3	7785	8240	8500	8209	0.09

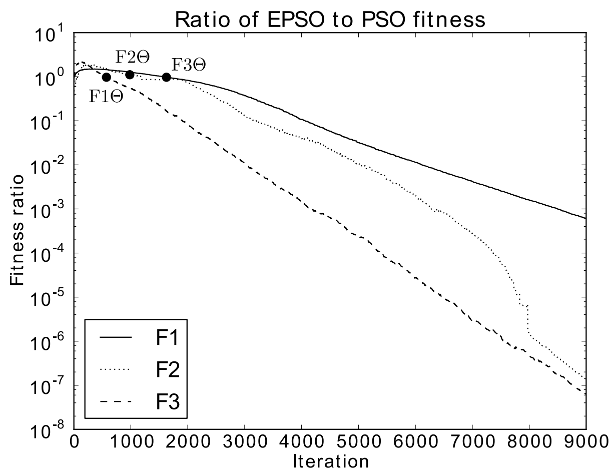

Figure 1: Comparison of median *gbest* fitness of EPSO to PSO swarms per iteration. Note that for iterations beyond Θ, EPSO has higher mean fitness than PSO.

F3, where PSO search was not excessively affected by premature convergence, the Δ values are small. In both problems there is a clear efficiency advantage of EPSO. In F3, this is an interesting result for which the cause will be expounded upon in subsequent sections. The characteristics of the F1 problem space exacerbate the greediness in predisposed algorithms. This certainly benefits the efficiency of PSO solutions, yet only until the peak of the local optimum is reached and solution quality fails to improve. It is deceptive, then, to compare algorithm efficiency in this problem.

5.2 Performance Across Experiments

Consider the comparison of swarm fitness across the duration of experiments. Figure 1 shows the ratio of EPSO mean fitness per iteration to that of PSO. The iteration at which the fitness ratio becomes less than one corresponds to the Θ values in Table 5. Note that the first 100 iterations are excluded in the calculation of Θ, as the fitness ratio in this period is significantly influenced by the initial configuration of the swarm rather than algorithm behaviour. In iterations before Θ, EPSO search was generally less efficient than PSO.

In all three profiles, shown in Figure 1, the advantage of EPSO's moderated search is evident. Some important features to note are the periods during which PSO swarms first converged, which generally correspond to the Q_2 values in Table 6. The consistent improvement in EPSO fitness for iterations thereafter is thus due to the inactivity of PSO swarms. This is most clear in F1, where the smoothness of the profile is due to comparison to an effectively fixed value, since PSO solution fitness fails to improve in the later period of all experiments. This does not obfuscate the qualities of EPSO search, however. Not only does EPSO exceed the solution quality of PSO, but the consistency in which it does so suggests that modest excitation yields a more effective search strategy.

The rapid relative improvement to fitness ratio in F2 is due to the collapse of PSO swarms across the period largely indicated by the quartile range in Table 6. Once again, a consistent fitness improvement is evident in EPSO swarms thereafter. While F2 is not as deceptive as F1, there nevertheless appears to be a general fitness limit beyond which PSO swarm performance cannot improve. It should be recognised that near the global optimum in F2, local optima become more deceptive as the summation component of the function contributes less to fitness. It appears that the prominence of the product component in the fitness evaluation of F2 at this search scale increases the difficulty of search such that the resulting frequency of *pbest* discovery is too low for PSO search to effectively persist.

In F3, the lack of local optima may have conceivably resulted in a similar search by both algorithms, but it is apparent that EPSO is consistently more efficient. This problem is notable given that for all positions, except the global optimum, there is a local region containing positions of higher fitness. However, this region becomes smaller upon approach to the global optimum, and thus particles may be less able to discover this region quickly when search becomes too fine-grained. This appears to be the case in PSO. EPSO, by consistently stimulating diversity, maintains higher velocity, which partially accounts for greater efficiency. While the partial convergence of PSO swarms should be expected given their governing dynamics, in F3 it is evident that this may occur even in non-deceptive problem spaces.

It should be noted that EPSO swarms are conspicuously inefficient compared to PSO swarms in the first 250 iterations of the experiments. This may be a result of large $|\vec{u}|$ values (see Equation (5)) in this period due to the initially excessive diversity of particles. As high diversity corresponds to high velocity, and thus high $|\vec{u}|$, the ensuing excitation by EPSO may be so extreme as to be detrimental.

5.3 Relative Diversity

Having considered performance across experiments, the factors that influenced such results are now discussed. The diversity Φ of a swarm of n particles at iteration t in experiment i is calculated as the mean distance of particles from the swarm centroid:

$$\Phi_i(t) = \frac{1}{n}\sum_{j=1}^{n}\left|\vec{x}_j(t) - \frac{1}{n}\sum_{k=1}^{n}\vec{x}_k(t)\right|$$

where $\vec{x}$ is the position vector of the particle denoted by a subscript. This was measured per iteration to create a "diversity profile" of the experiment. Profiles were then averaged across all experiments for each algorithm-function pair. This gives some insight into the diversity characteristics of the search. The integral of the mean diversity across experiments, per iteration, is given by:

$$\Omega = \sum_{t=0}^{9000}\left[\frac{1}{|s|}\sum_{i=1}^{|s|}\Phi_i(t)\right]$$

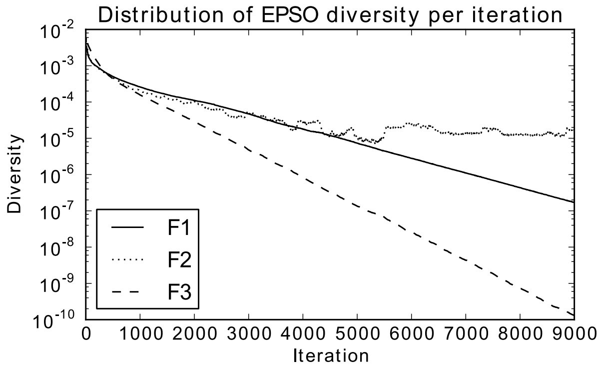

Figure 2: Distribution of diversity per iteration for EPSO on each problem

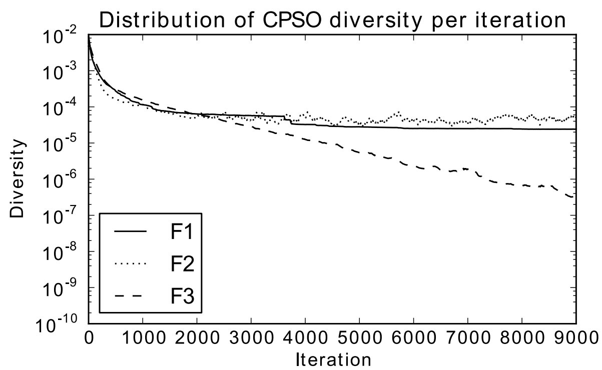

Figure 3: Distribution of diversity per iteration for CPSO on each problem

where i iterates over the set of experiments s. Plotting

$$\frac{\frac{1}{|s|}\sum_{i=1}^{|s|}\Phi_i(t)}{\Omega}$$

for $t \in [0, 9000]$ in each algorithm-function pair yields the profiles in figures 2 and 3.

Each profile represents the polarisation between exploration and exploitation experienced by swarms, as indicated by the profile's gradient. A large range suggests a moderated search. An ideal example of this is EPSO F3, which demonstrates a smooth, consistent scaling of diversity; from that due to the initial random placement of particles, to fine-grained search in the late period of experiments. It should be observed that despite the lack of local optima in F3, which may have otherwise exacerbated premature convergence, PSO swarms do not have a similarly equitable distribution of diversity to EPSO swarms. This is corroborated in previous results where PSO is shown to be less efficient in this search space, which is consistent with the evidence that reduced diversity correlates to reduced velocity.

It is instructive to consider the opposite case to EPSO F3. PSO F1 is marked by a sharp decline in the first 1000 iterations followed by two periods of effectively no change in the diversity distribution. A sharp decline followed by constancy is symptomatic of the premature convergence demonstrated in previous results, as the profile indicates that exploration almost immediately transitioned to swarm collapse. It should be noted that the swarm cannot converge to zero diversity due to the conditions of *pbest* discovery, but the swarm has effectively done so as the magnitude of diversity is not sufficient to allow other optima to be discovered.

PSO F2 should be observed as the algorithm-function pair in which search was most strongly polarised. Despite this, PSO can be said to have effectively solved F2, though it should be noted that F2 is not difficult in 100 dimensions. For the majority of search in F2, a swarm may be expected to improve performance despite little diversity given the pronounced global fitness trend. However, once F2 fitness evaluations become sensitive to the product component of the function, it is apparent that the diversity dynamics of PSO are not sufficient to allow search to persist.

At approximately 5500 iterations, EPSO swarms increase diversity and maintain it thereafter. This corresponds to swarm fitness proceeding below 0.0001, thus swarms occupy the region in which the product component dominates fitness evaluation in F2. The feedback from the fitness surface is polarised in this region, as beneficial change in a single dimension results in less fitness improvement, but movement in a generally beneficial direction can be expected to significantly (relatively) improve fitness. EPSO should thus be expected to yield higher quality solutions than PSO in F2 due to higher diversity. The fixed duration that such excitation applies, controlled by g and a (Equation (5)), is also likely to benefit search performance. More generally, it should be seen that EPSO swarms were able to significantly modify search behaviour, as evident in Figure 1, as the characteristics of the local fitness surface, or the fitness feedback from the space, changed.

6. DISCUSSION

6.1 EPSO Dynamics

It may seem contradictory to stimulate exploration when exploitation appears to be most valuable, as EPSO does. However, the corollary suggests that immediate convergence is an acceptable search strategy. Implicit in the presentation of the EPSO algorithm is the suggestion that particles should be excited upon discovery of a high fitness region, and continue to be stimulated as long as the region continues to yield *pbests*. EPSO demonstrates such behaviour.

In a multimodal problem space, any given optimum discovered is unlikely to be the global. However, a *pbest* discovery generally suggests a region of high fitness, an optimum; let this be region R. It may also indicate a global fitness trend, or be near to higher fitness optima, and so it might be more effective to search a region larger than R; let this be S, and $R \in S$. If diversity, and thus velocity, is stimulated such that a swarm can briefly explore S, higher fitness solutions may be discovered without having to first exploit or migrate through R, upon which swarms might also otherwise prematurely converge. PSO's polarisation in diversity suggests that search is typically confined to R once an exploitation phase begins. Even so, if exploration in S fails to yield a new *pbest*, convergence on R is appropriate, as it is a known region of high fitness, within which swarms may refine solutions. It seems that, if diversity excitation is proportional to the frequency of *pbest* discovery, a suitable degree of diversity can be maintained, appropriate to the local fitness surface, thus increasing search efficiency.

EPSO appears to demonstrate this behaviour. It is readily apparent that excitation stimulates diversity upon *pbest* discovery, as the manner of *lbest* modification frequently yields a greater $|\vec{s}|$ (Equation (3)), especially where $d > 1$ (Equation (5)). Thus a greater $|\vec{v}|$ (Equation (1)) results, inevitably increasing spatial diversity. A greater $|\vec{v}|$ also increases the distance between fitness evaluation positions. Should *pbest* discovery immediately occur in this period of excitation, then, a larger $|\vec{u}|$ (Equation (5)) can be expected. This increases $|\vec{s}|$, and so the process repeats. This potential for a positive feedback mechanism regarding diversity may significantly offset the negative diversity feedback mechanism apparent in PSO.

The dynamics arising from $|\vec{u}|$ have an identifiable effect on results. Consider in F1 that the early period of search is dominated by exploitation of an attractive yet ultimately poor optima, to which PSO swarms succumbed in a short period. The similar fitness of EPSO swarms in this period suggests that some exploitation of local optima occurred. However, given that this optima is attractive, one may expect frequent *pbest* discovery to occur. Should this activate the positive feedback mechanism in $|\vec{u}|$ to some degree, it is likely that EPSO swarms would be excited to an extent that perhaps ensured the discovery of higher fitness optima. It is notable that no EPSO swarm prematurely converged in this problem space. While the positive feedback mechanism need not strictly apply for EPSO to avoid premature convergence, the frequent discovery of *pbests* is nevertheless evident in EPSO's moderated diversity loss in this problem.

Should excitation fail to yield a new *pbest*, the time-based nature of the EPSO effect can be seen to permit convergence under canonical PSO. Specifically, where $t_c - t_u = g$, $\vec{e} = \vec{b}_n$ (see Equation (5)), and thus EPSO behaviour is identical to canonical PSO in such cases. Note that the condition $t_c - t_u \leq g$ prevents exploration from being stimulated when the period of excitation g elapses. In relation to the feedback mechanism described above, it can be seen that excitation is only reinforced if it continues to be useful, otherwise typical exploitation occurs. Furthermore, during typical convergence, after the period of excitation elapses, the next *pbest* should be expected to be found near to the previous. If this occurs $|\vec{u}|$ will be small, thus the magnitude of EPSO influence is decreased, reducing its interference in exploitative behaviour.

This residual influence of $|\vec{u}|$ still appears to serve a valuable role, however. Consider that persistent fine-grained search relies on the discovery of a new *pbest* in an increasingly small region of higher fitness. As such, the probability that a particle's velocity will carry it into this region is reduced as exploitation progresses and convergence occurs. Should a particle fail to find the region immediately, it may be expected that it will do so as it slows and approaches the previous *pbest*, as positions immediately local to this former *pbest* are likely to have similar fitness. But in doing so, a particle significantly loses velocity and the efficiency of search is reduced.

The temporary yet fixed duration of particle excitation under EPSO may maintain the diversity that PSO swarms would otherwise lose in this scenario. This increases the chance that the region will be found while a particle has relatively high velocity, without first requiring excessive convergence. This may be responsible for the inherent advantage of EPSO in search efficiency on each problem, after the Θ period of iterations. Specifically, in F3 swarms were persistently involved in the scenario described above, and consistently higher efficiency is observed in EPSO. Results in F2 suggest that the fineness required for search to persist eventually became too great for PSO swarms to continue to improve performance, for whom collapse would have been exacerbated by infrequent *pbest* discovery. Slight, temporary diversity stimulation during exploitation appears to have permitted EPSO swarms to further refine solutions in such conditions. That this behaviour is permitted at such a scale is also a positive reflection on the scaling effect of $|\vec{u}|$, despite the fixed d parameter used in experiments.

6.2 Further Experimentation

The experiment performed in this investigation has concentrated on three objective functions for the purpose of confirming the viability of EPSO as a diversity preserving mechanism. It has been shown to meet this aim, but the selected problems are only a small subset of the objective functions typically employed to test heuristic performance. EPSO has demonstrated similar results on the rotated hyperellipsoid function and Rastrigin's function, yet careful selection of objective function in future investigations may yield a greater understanding of the dynamics of EPSO search.

While the experiment detailed in this paper has suggested evidence for the dynamics assumed to be responsible for EPSO search behaviour, additional experiments have not yet been designed to confirm the hypothesis. Preliminary testing has confirmed that EPSO particles have consistently higher mean velocity than their PSO counterparts, corresponding to increased diversity, during all periods of experiments. Whether this affects the distance between consecutive *pbests* has not yet been investigated, but should be considered a natural consequence of higher mean particle velocities. Nevertheless, a specific experiment may confirm this and also provide more insight into the nuanced aspects of EPSO search behaviour.

Another potential line of enquiry observes that the EPSO algorithm interferes minimally with the structure of the canonical PSO algorithm. As such, it could be subjected to the same mechanisms as canonical PSO, such as inertia weight tweaking, choice of swarm topology and integration with other particle behaviours such as WoSP [5]. Sensitivity to canonical PSO parameter choice has not been tested in EPSO. A focused investigation may improve understanding of the difference between EPSO and PSO dynamics by highlighting differences in their performance as a result of common parameter adjustment.

6.3 Algorithm Refinement

Two main areas are recognised for algorithm refinement: (A) determination of the direction in which exploration is stimulated and (B) mitigation of the counteractive behaviour evident in the early period of search. In (A), this refers to the composition of position $\vec{u}$ in Equation (5). There exist similar methods of similar computational complexity that may serve the same role. However, the relationship between the current determination method of $\vec{u}$ and EPSO's ability to self-scale should also be considered. $|\vec{u}|$ may be critical to this, but the direction of $\vec{u}$ may not. A survey of $\vec{u}$ determination methods and the sensitivity of scaling to $|\vec{u}|$ would be prudent.

In Figure 1 there is a period, before the Θ iterations, in which EPSO's performance is suboptimal to PSO. It was argued that this is due to EPSO's attempt to leverage local information, which is less useful when particles occupy distant regions of the problem space – presumably only benefiting search when the fitness surface is varied in such cases. Although EPSO soon overtakes PSO in efficiency, one should consider that fitness evaluations are nevertheless taking place during EPSO's period of suboptimality, and such evaluations may dominate the processing load of the experiment [6].

Addressing concern (B), a candidate solution is to scale the distance between the advertised *lbest* — $\vec{e}$ in Equation (5) — and the actual *lbest* according to the proximity of particles, so as to advertise an accurate *lbest* to distant particles and a modified position when neighbour particles are proximate. This, however, may also interfere with the ability of EPSO to effectively scale diversity, as externally modifying the position $\vec{e}$ indirectly scales $|\vec{u}|$. A better solution may be to place a limit on the upper magnitude of $\vec{u}$. This may allow the swarm to effectively scale, but extreme exploration arising from an excessive $|\vec{u}|$ is avoided.

7. CONCLUSIONS

The excited PSO algorithm has been shown to be a viable method for reducing the rate at which diversity in swarms is lost in some problem domains, with respect to solution quality. Furthermore, the algorithm has demonstrated a greater ability to self-scale than canonical PSO on said problems. In the results presented, EPSO has exhibited the properties of a robust heuristic; appropriate diversity, consistency among experiments, reduced risk of premature convergence and relatively high search efficiency. Particle excitation differs from traditional diversity preservation methods in that modifying the social component vector of particle velocity reduces the dependence of diversity at any given iteration from that of the immediately previous iteration. Opportunity exists for further investigations to confirm the hypothesis arising from the observations made in this paper. Additionally, the practicalities and applicability of EPSO should be determined with respect to canonical PSO parameter choice across a broader range of objective functions. Finally, given that this is the initial presentation of the algorithm, it may yet benefit from modification to the intricacies of its construction, but the manner of diversity introduction has nevertheless been demonstrated to be useful.

8. REFERENCES

[1] D. H. Ackley. *A Connectionist Machine for Genetic Hillclimbing*. Kluwer Academic Publishers, 1987.

[2] M. Clerc and J. Kennedy. The Particle Swarm - Explosion, Stability, and Convergence in a Multidimensional Complex Space. *IEEE Transactions on Evolutionary Computation*, 6:58–73, 2002.

[3] A. O. Griewangk. Generalized Descent for Global Optimization. *Journal of Optimisation Theory and Applications*, 34:11–39, 1981.

[4] T. Hendtlass. Preserving Diversity in Particle Swarm Optimisation. In P. Chung, C. Hinde, and A. Moonis, editors, *Developments in Applied Artificial Intelligence*, volume 2718 of *Lecture Notes in Computer Science*, pages 155–199. Springer Berlin, 2003.

[5] T. Hendtlass. WoSP: A Multi-optima Particle Swarm Optimisation. *IEEE Congress on Evolutionary Computating*, 1:727–734, 2005.

[6] T. Hendtlass. Fitness Estimation and the Particle Swarm Optimisation Algorithm. *IEEE Congress on Evolutionary Computation*, pages 4266–4272, 2007.

[7] J. Kennedy. Small Worlds and Mega-minds: Effects of Neighborhood Topology on Particle Swarm Performance. *Proceedings of the 1999 Congress on Evolutionary Computation*, pages 1931–1938, 1999.

[8] J. Kennedy and R. C. Eberhart. Particle Swarm Optimisation. *Proceedings of IEEE International Conference on Neural Networks*, IV:1992–1998, 1995.

[9] J. Kennedy, R. C. Eberhart, and Y. Shi. *Swarm Intelligence*. Morgan Kaufmann, 2001.

[10] J. Kennedy and R. Mendes. Population Structure and Particle Swarm Performance. In *Proceedings of the Congress on Evolutionary Computation (CEC 2002)*, volume 2, pages 1671–1676. IEEE Press, 2002.

[11] R. Mendes, J. Kennedy, and J. Neves. The Fully Informed Particle Swarm: Simpler, Maybe Better. *IEEE Transactions on Evolutionary Computation*, 8:3:204–210, 2004.

[12] Y. Shi and R. Eberhart. Parameter Selection in Particle Swarm Optimisation. *Proceedings of Evolutionary Programming VII*, pages 591–600, 1998.

[13] Y. Shi and R. C. Eberhart. A Modified Particle Swarm Optimizer. *Proceedings of IEEE International Conference on Evolutionary Computation*, pages 69–73, 1998.