Particle Swarm Optimization: Classic or State of the Art?

Samantha Ball
1603701
School of Computer Science and Applied Mathematics
University of the Witwatersrand

***Abstract*—Stemming from the emergence of particle swarm optimization (PSO) as an effective, population-based optimization technique, numerous variants have been proposed, each offering specific benefits over the original algorithm. In particular, the latest Standard PSO (SPSO-2011) algorithm offers a two-fold advantage of ensuring rotational invariance and incorporating an adaptive topology. However, a fundamental drawback of many of the more complex variants is the associated lack of stability criteria. This is a significant limitation in terms of the practical use of the PSO variant as stability has a great impact on the overall performance of the PSO algorithm.**

In this paper we investigate the proposed benefits of the SPSO-2011 variant in comparison to the classic inertia weight approach (IPSO). We compare the performance of the two algorithms on nine well-known objective functions to benchmark their performance on a variety of problems. Several shifted and rotated functions are included in the analysis to determine the effects of the rotational invariance offered by the SPSO-2011 algorithm. For each algorithm, comprehensive hyperparameter tuning is implemented to ensure order-1 and order-2 stability as well as optimal performance.

Through a Mann-Whitney U analysis, it was determined that the IPSO algorithm produced statistically better results for seven out of the nine objective functions. Furthermore, the IPSO algorithm continued to perform better even in the case of the shifted and rotated variants of the objective functions. This unexpected result indicates that novel techniques do not always offer substantial improvement over the more traditional approach.

***Index Terms*—particle swarm optimization, inertia weight, spso2011, benchmark functions**

I. Introduction

Particle swarm optimization (PSO) is a popular, stochastic search algorithm that relies on collective decision-making to inform the direction of the search [1]. Particle swarm optimization has been shown to perform effectively for a variety of continuous optimization tasks. The core mechanism of particle swarm optimization relies on the use of both cognitive and social information within the population. The classic technique relies on a weighted summation of the directions towards the best position obtained by the individual particle and towards the best position obtained by the swarm as a whole.

Since the initial formulation of particle swarm optimization, several variants of PSO have been proposed. Variants such as Fully Informed PSO (FIPS) which incorporates a statistical summary from all the neighbours in the neighbourhood of a particle rather than only the best-performing neighbour, and unified PSO (UPSO), which introduces a unification factor to combine the benefits of the GBest and LBest approaches, have associated stability criteria to guide the appropriate selection of control parameters.

It is known that the performance of particle swarm optimization is highly dependent on an appropriate selection of control parameters. In particular, it is necessary to ensure both order-1 and order-2 stability in order for particle swarm optimization to provide an effective solution. However, a central challenge to the practical use of PSO is the availability of appropriate stability criteria used to guide control parameter choice. For novel variants such as Standard Particle Swarm Optimization (SPSO2011), this criteria is often unavailable as it constitutes a yet unsolved area of research. Thus there exists a trade-off between the benefits offered by novel approaches and the lack of corresponding stability criteria.

In this paper, we compare the classic inertia weight particle swarm optimization (IPSO) algorithm with the state-of-the-art SPSO2011 variant on nine well-known benchmarking functions. For each function, hyperparameter tuning is implemented in order to determine the most suitable control parameter values, while ensuring both order-1 and order-2 stability. The performance of each algorithm is then compared using a Mann-Whitney U test with 30 samples for each objective function.

The paper is structured as follows. In Section II, we outline the relevant techniques and core differences between the two PSO variants. In Section III, we provide an in-depth discussion of the algorithms, their specific design features, hyperparameter tuning and the objective functions selected for benchmarking. Section IV presents our findings and a comprehensive comparison of the two algorithms. Lastly, Section V summarizes our central findings.

II. Background

The core mechanisms and benefits of the two PSO variants are explored further below, along with the characteristics of the relevant objective functions.

A. Inertia Weight Particle Swarm Optimization (IPSO)

Inertia weight PSO is often considered to be the classic or base implementation of the particle swarm optimization metaheuristic. The inertia weight approach allows velocity to be limited without the introduction of additional hyperparameters, such as in the case of velocity clamping techniques.

The inertia weight particle swarm optimization algorithm, assuming minimization, is summarized in the process below.

Algorithm 1 Inertia Weight PSO

```
Let p_i denote the best position attained by particle i
Let n̂_i denote the best position attained in the neighbourhood of particle i
Let g denote the best position attained by the entire population
while stopping condition false do
  for i in population do
    if f(x_i) < f(p_i) then
      p_i = x_i
    end if
    for î with particle i in the neighbourhood do
      if f(p_i) < f(n̂_i) then
        n̂_i = p_i
        if f(n̂_i) < f(g) then
          g = n̂_i
        end if
      end if
    end for
  end for
  for i in population do
    Perform velocity update for particle i
    Perform position update for particle i
  end for
end while
return g
```

Overall, the PSO algorithm is guided by the stochastically weighted sum of the particle velocity, cognitive information and social information. The current position of a particle in the swarm is updated through the addition of a movement vector, as indicated by the following position update equation,

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

where $x_i(t+1)$ denotes the updated position for particle i, $x_i(t)$ is the current position and $v_i(t+1)$ is velocity of particle i.

The velocity term $v_i(t+1)$ is comprised of cognitive information based on the particle's own optimal position and social information regarding the best position obtained by the neighbourhood of the particle or the swarm in its entirety. These terms are then stochastically weighted through the uniformly sampled vectors $\mathbf{r}_1$ and $\mathbf{r}_2$. In addition to the stochasticity introduced by $\mathbf{r}_1$ and $\mathbf{r}_2$, the cognitive and social components are weighted using selected coefficients c_1 and c_2. The optimal values of c_1 and c_2 are problem dependent and require hyperparameter tuning. The velocity update equation for the inertia weight approach is denoted as follows,

$$\mathbf{v}_i(t+1) = w\mathbf{v}_i(t) + c_1\mathbf{r}_1 \otimes (\mathbf{p}_i(t) - \mathbf{x}_i(t)) + c_2\mathbf{r}_2 \otimes (\hat{\mathbf{n}}(t) - \mathbf{x}_i(t)) \tag{2}$$

where w is the inertia weight, $p_i(t)$ is the best position attained by particle i, $\mathbf{r}_1, \mathbf{r}_2 \sim U(0,1)^d$ and $\hat{n}(t)$ is the best position obtained by either the neighbourhood of the particle or the swarm as a whole, depending on the chosen topology. It should be noted that $\mathbf{r}_1$ and $\mathbf{r}_2$ are d-dimensional vectors in order to ensure that the entire search space can be reached during optimization.

The neighbourhood N_i of each particle i is dependent on the chosen topology which dictates the communication structure within the swarm. The particle i derives its social information from the neighbourhood N_i of particles that it is connected to. The implications of various topological structures are discussed further in the implementation.

B. Standard Particle Swarm Optimization (SPSO-2011)

The SPSO-2011 variant [4] is the latest of the standard PSO algorithms, following SPSO-2006 and SPSO-2007. The SPSO-2011 variant proposes modifications to the position and velocity update equations, as well as a novel adaptive topology. The core benefits proposed in the SPSO2011 variant include rotational invariance and a dynamic network structure.

B.1 Rotational Invariance

In order to allow for rotational invariance, the modified velocity update equation is defined as follows,

$$\mathbf{v}_i(t+1) = w\mathbf{v}_i(t) + H_i(\mathbf{g}_i(t), ||\mathbf{g}_i - \mathbf{x}_i||_2) - \mathbf{x}_i(t) \tag{3}$$

where w is the inertia weight, $\mathbf{g}_i$ is the centre of gravity and H_i denotes the hypersphere distribution centred at $\mathbf{g}_i$ with radius $||\mathbf{g}_i - \mathbf{x}_i||_2$ being the distance between the centre of gravity and the current position $\mathbf{x}_i$.

The centre of gravity represents... and can be calculated as follows,

$$\mathbf{g}_i(t) = \frac{\mathbf{x}_i(t) + \alpha_i(t) + \beta_i(t)}{3} \tag{4}$$

where,

$$\begin{aligned} \alpha_i(t) &= \mathbf{x}_i(t)) + c_1\mathbf{r}_1 \otimes (\mathbf{p}_i(t) - \mathbf{x}_i(t)) \\ \beta_i(t) &= \mathbf{x}_i(t)) + c_2\mathbf{r}_2 \otimes (\mathbf{n}_i(t) - \mathbf{x}_i(t)) \end{aligned} \tag{5}$$

Finally, the position update equation remains the same,

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

The resulting rotational invariance stems from the fact that points are sampled from a hypersphere rather than a rectangular space. The associated hypersphere can be depicted as in Figure 1 below, indicating the relation between the centre of gravity $\mathbf{g}_i(t)$, personal best $\mathbf{p}_i(t)$ and neighbourhood best $\mathbf{n}_i(t)$ positions.

While inertia weight PSO offers scale and translation invariance, it lacks the rotational invariance offered by

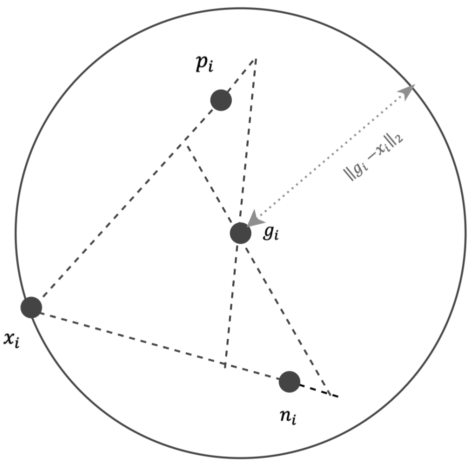


Fig. 1: Hypersphere with centre $\mathbf{g}_i$ and radius $||\mathbf{g}_i - \mathbf{x}_i||_2$.

SPSO2011. Thus the role of rotational invariance can be explored through the comparative performance of the IPSO and SPSO2011 approaches.

B.2 Adaptive Random Topology

The topology proposed by SPSO-2011 forms a specific case of the stochastic star topology [1]. Each particle is initialised with a neighbourhood of k particles which are randomly selected from the swarm, inclusive of the particle itself. The number of particles in the neighbourhood is often chosen as $k = 3$. In the case of an unsuccessful iteration, the neighbourhoods of each particle are reconstructed, resulting in an adaptive communication structure.

In this study, an unsuccessful iteration is defined as an iteration during which the global best did not improve and therefore no new position was found that improved the best position found by the whole swarm.

C. CEC2014 Benchmark Problems

The relevant objective functions used in this investigation, together with their search boundaries and defining characteristic are outlined below. In some cases, the global minimum is only known for a certain dimensionality. It should be noted that all functions are minimization problems.

Spherical:

$$f(x) = -\sum_{j=1}^{n_x} x_j^2 \tag{7}$$

with each $x_j \in [-5.12, 5.12]$. The spherical function is a continuous, convex and unimodal function. The global minimum of the spherical function with no applied rotations or shifts is $f^*(x) = 0$.

Ackley:

$$f(x) = -20e^{-0.2\sqrt{\frac{1}{n}\sum_{j=1}^{n_x} x_j^2}} - e^{\frac{1}{n}\sum_{j=1}^{n_x} cos(2\pi x_j)} + 20 + e \tag{8}$$

with each $x_j \in [-32.768, 32.768]$. The ackley function is a non-convex function containing many local minima. The global minimum of the ackley function with no shift or rotation is found at $f^*(x) = 0$.

Michalewicz:

$$f(x) = -\sum_{j=1}^{n_x} sin(x_j)\left(sin\left(\frac{jx_j^2}{\pi}\right)\right)^{2m} \tag{9}$$

where $m = 10$ and each $x_j \in [0, \pi]$. The michalewicz function is a non-convex, multimodal optimization landscape containing many steep ridges and drops. The global minimum of the michalewicz function with no shift or rotation is found at $f^*(x) = -19.6370136$ for $n_x = 2$.

Katsuura:

$$f(x) = -20e^{-0.2\sqrt{\frac{1}{n}\sum_{j=1}^{n_x} x_j^2}} - e^{\frac{1}{n}\sum_{j=1}^{n_x} cos(2\pi x_j)} + 20 + e \tag{10}$$

with each $x_j \in [0, 100]$. The Katsuura function is a multimodal function with a global minimum at $f^*(x) = 0$.

Shubert:

$$f(x) = \prod_{j=1}^{n_x}\left(\sum_{i=1}^{5}(i\cos((i+1)x_j + i)\right) \tag{11}$$

with each $x_j \in [-10, 10]$. The shubert function has many local minima and is both non-convex and multimodal. The global minimum of the shubert function with no shift or rotation is found at $f^*(x) = -186.7309$ for $n_x = 2$.

C.1 Shifts and Rotations

Shifts and rotations are applied to the benchmark functions to ensure that the search algorithm performs well regardless of where the function is located in the search space. Thus the optimum point is shifted to a new shifted optimum o together with a rotation, applied through a rotation matrix M. Further details regarding the implementation of the shifts and rotations is provided in Section III.

D. Particle Convergence and Stability

In the context of particle swarm optimization, convergence is critical to ensure useful results. Convergence in this context is defined as the velocity of the particle tending to zero rather than a guarantee that the swarm will converge to a local optima. Specifically, both order-1 and order-2 stability needs to be fulfilled in order for convergence to be achieved.

D.1 Order-1 Stability

Order-1 stability is defined as a convergence in the expectation of the particle position to a constant α as follows,

$$lim\ E[x_t] \longrightarrow \alpha \tag{12}$$

However, in isolation order-1 stability is insufficient as we must also ensure that the variance of the particle position converges to a constant. This is given by order-2 stability.

D.2 Order-2 Stability

Order-2 stability requires that the variance of the particle position x_i converges to a constant as follows,

$$lim\ V[x_t] \longrightarrow \beta \tag{13}$$

The significance of order-1 and order-2 stability has been shown as particle swarm optimization can exhibit worse than random performance in the absence of suitable control parameters.

III. Implementation

In order to compare the efficacy of IPSO and SPSO2011 algorithms, the approaches were assessed on nine objective functions, with 20 dimensions and a fixed swarm size of 20.

A. Inertia Weight Particle Swarm Optimization (IPSO)

A.1. Topology Choice

The choice of topology has a significant impact on PSO performance as it dictates the number of particles from which a certain particle derives its social information. Common fixed social topologies can be separated into two broader categories, namely GBest and LBest PSO.

a) GBest

In GBest PSO, all particles are connected to every other particle in the swarm. Consequently, there is no concept of a 'neighbourhood' as all particles can communicate with one another. The star topology is utilised for GBest PSO.

b) LBest

In contrast, LBest PSO or Local Best PSO defines a set of neighbourhood particles for each particle i. The social component of the velocity update equation is then determined by the social informer $\mathbf{n}_i$, known as the neighbourhood best or local best. A variety of topologies can be chosen to determine the neighbourhood of each particle, including the ring, wheel, n-cluster, pyramid and Von-Neumann communication structures. In this study, the ring topology was chosen where each particle is connected to one neighbour on either side, as depicted in Fig. 2, resulting in neighbourhoods of three particles.

For the purposes of this study, both GBest and LBest IPSO were implemented as topology can be viewed as a hyperparameter for the inertia weight PSO algorithm. The optimal performance between the GBest and LBest implementations was then chosen as the final IPSO configuration when comparing to the SPSO2011 variant, for each of the optimization functions.

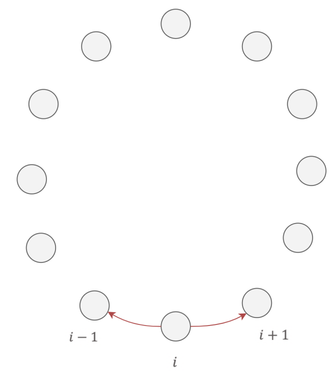


Fig. 2: Swarm communication using the ring topology. The neighbourhood of particle i is given by the two adjacent particles in the ring.

A.2. Initialisation

Each particle in the swarm was initialized randomly by sampling the value of each dimension uniformly from within the search space bounds for the given function. The velocity was initialised to zero ($v_i = 0$) in order to minimize potential particle roaming outside of the search space.

A.3. Hyperparameter Tuning

The choice of hyperparameters was guided by the most recent stability criteria for the inertia weight PSO algorithm. Specifically it has been shown that particle convergence can be achieved if the coefficients satisfy the following relation,

$$c_1 + c_2 < \frac{24(1 - w^2)}{7 - 5w} \tag{14}$$

Combinations of values were then chosen along the perimeter of the curve defined by the equation above in order to find the optimal values for the search.

a) Inertia weight w

The inertia weight determines the proportion of the previous velocity that is retained and should fulfil the following condition,

$$|w| < 1 \tag{15}$$

in order for the stability criteria outlined above to hold true.

b) Cognitive and social coefficients c_1 and c_2

The cognitive and social coefficients play a key role in the level of exploration and exploitation that occurs during the search. The cognitive coefficient c_1 determines to what extent the particle relies on its own knowledge of the best position obtained while the social coefficient c_2 determines the amount of reliance on the collective intelligence from the rest of the swarm. Overall lower values of the coefficients c_1 and c_2 result in smoother particle trajectories while higher values can result in more abrupt movements. Conventionally, the two coefficients are set to the same value $c_1 = c_2 > 0$ such that particles are attracted to the average of the personal best p_i and global or local best n_i. For the purposes of this investigation, the coefficients are always assumed to take on the same value in line with this convention.

B. Standard Particle Swarm Optimization (SPSO-2011)

B.1. Sampling from the Hypersphere Distribution

In order to correctly generate a uniformly sampled random position from the hypersphere, n_x scalars where generated from the normal distribution in order to ensure that the uniformity of the sampling was maintained during normalization. The vector formed from these scalars was the normalized and scaled by scalar s and shifted to the correct position in the search space.

B.2. Hyperparameter Tuning

Although no standard stability criteria exist for the SPSO2011 variant, this investigation utilised the convergence analysis found in [1] to guide the parameter search. The combinations of coefficient and inertia weight values were thus chosen in line with the stability region outlined in [1]. In order to ensure order-1 and order-2 stability, the movement of a particle in the swarm was plotted in order to ensure that both the expectation and variance of the particle position converged to a constant value. This is depicted in Fig. 3.

C. CEC2014 Benchmark Functions

C.1. Shifts and Rotations

In order to implement the shifted and rotated variants of the objective functions, the input position x was transformed as follows,

$$z = M(x - o) \tag{16}$$

where M represents an $n \times n$ rotation matrix and o is a vector representing the applied shift and modified optimal position. The shift vector o is generated randomly, ensuring each dimension is within the given search space bounds for the particular objective function. In addition a constant γ was added to the objective function as follows,

$$f_{ShR} = f(z) + \gamma \tag{17}$$

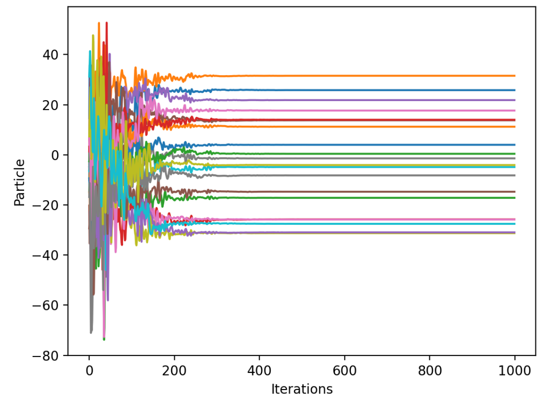


Fig. 3: Movement of a particle in the swarm over the search duration, indicating order-1 and order-2 stability for all 20 dimensions. Each line represents the movement in value for one of the twenty dimensions.

where γ represents the shift in the global minimum. The new global minimum is then given by $f^*_{new}(x) = f^*_{old}(x) + \gamma$ A γ value of 500 was chosen for all functions, resulting in a new global minimum of $f^*_{new}(x) = f^*_{old}(x) + 500$ for each function. These transformations were performed in line with the CEC2014 benchmark function definitions [3].

D. Evaluation

In order to provide a statistically sound experimental design, the IPSO and SPSO2011 algorithms were compared using the Mann-Whitney test with 30 simulations. For each experiment, a fixed swarm size of 20, dimensionality of 20 and 1000 iterations were used. Each PSO algorithm was run with the final c_1, c_2 and w values obtained through the hyperparameter tuning process, as tabulated in Table I. The testing procedure is summarized as follows:

- Each PSO algorithm was run using the final configurations and control parameter values discussed above.
- Thirty simulations were run for the IPSO and SPSO2011 algorithms respectively to obtain a statistically sufficient number of samples.
- This was then repeated for each of the nine benchmarking functions.

All algorithms were implemented in Python 3 and all experiments were run in a Google Colab environment.

IV. Results

The performance results for each objective funtion obtained for the IPSO and SPSO2011 algorithm respectively are shown in Table I. It can be observed that the global best IPSO algorithm performs the best on the majority of the objective functions. However, in most cases the performance of the variants displays high similarity. Several notable observations include the improved performance displayed by the SPSO2011

TABLE I: Mean accuracy indicating minimum value obtained by each PSO variant, averaged over 30 samples

Function	Algorithm	Inertia weight w	Cognitive coefficient c_1	Social coefficient c_2	*lower is better* Mean Accuracy
Spherical	*IPSO G-Best*	0.5	1.9	1.9	$\mathbf{1.0767 \times 10^{-24}}$
	IPSO L-Best	0.6	1.75	1.75	1.0673×10^{-16}
	SPSO2011	0.729844	1.49619	1.49619	1.9737×10^{-22}
Ackley	*IPSO G-Best*	0.8	1.5	1.5	2.810×10^{-1}
	IPSO L-Best	0.6	1.75	1.75	$\mathbf{4.745 \times 10^{-2}}$
	SPSO2011	0.85	1.2	1.2	9.598×10^{-1}
Michalewicz	*IPSO G-Best*	0.5	1.9	1.9	$\mathbf{-1.581 \times 10^{1}}$
	IPSO L-Best	0.5	1.9	1.9	-1.544×10^{1}
	SPSO2011	0.8	1.5	1.5	-1.314×10^{1}
Katsuura	*IPSO G-Best*	0.6	1.75	1.75	$\mathbf{2.4514 \times 10^{-2}}$
	IPSO L-Best	0.5	1.9	1.9	1.213×10^{-1}
	SPSO2011	0.729844	1.49619	1.49619	5.852×10^{-1}
Shubert	*IPSO G-Best*	0.729844	1.49618	1.49618	$\mathbf{-1.1225 \times 10^{22}}$
	IPSO L-Best	0.5	1.9	1.9	-3.0366×10^{21}
	SPSO2011	0.8	1.5	1.5	-1.2006×10^{19}
Rotated and Shifted Ackley	*IPSO G-Best*	0.5	1.9	1.9	5.0833×10^{2}
	IPSO L-Best	0.729844	1.49619	1.49619	5.0197×10^{2}
	SPSO2011	0.8	1.5	1.5	$\mathbf{5.0167 \times 10^{2}}$
Rotated and Shifted Michalewicz	*IPSO G-Best*	0.6	1.75	1.75	$\mathbf{4.905 \times 10^{2}}$
	IPSO L-Best	0.5	1.9	1.9	4.916×10^{2}
	SPSO2011	0.8	1.5	1.5	4.928×10^{2}
Rotated and Shifted Katsuura	*IPSO G-Best*	0.729844	1.49619	1.49619	$\mathbf{5.001 \times 10^{2}}$
	IPSO L-Best	0.5	1.9	1.9	5.002×10^{2}
	SPSO2011	0.8	1.5	1.5	5.006×10^{2}
Rotated and Shifted Shubert	*IPSO G-Best*	0.6	1.75	1.75	$\mathbf{-3.1921 \times 10^{21}}$
	IPSO L-Best	0.5	1.9	1.9	-1.6441×10^{20}
	SPSO2011	0.729844	1.49618	1.49618	-8.4571×10^{18}

variant for the rotated and shifted Ackley function in comparison with both IPSO variants, and the improved results obtained by the LBest IPSO on the Ackley ans shifted Ackley functions only.

As noted by Engelbrecht [2], intuition expecting the local best variant to perform better in terms of avoiding local minima is not always the case in practice. Similarly to the study by Engelbrecht [2], the choice between GBest and LBest IPSO is seen to be function dependent. However it can be noted that the more novel SPSO2011 variant did not significantly outperform the classic inertia weight approach, even in the case of the rotated and shifted objective functions.

In order to further confirm this observation and determine whether the results given by the IPSO and SPSO2011 variants are statistically different, a Mann-Whitney test was employed for each benchmark function as summarised in Table II. The Mann-Whitney U Test is a null hypothesis test, used to detect differences between two independent data sets. Since our performance data is continuous, and we make no assumption about the distribution of our data, the Mann-Whitney test is appropriate for our context. We consider the null hypothesis $\mathbf{h_0}$ which states that the distributions of test accuracies obtained from the IPSO and SPSO2011 algorithms respectively are the same.

A. One-tailed Test

Based on our summary statistics indicating which variant produces superior results on average for each objective function, we first consider a one-tailed test in which the alternative hypothesis $\mathbf{h_a}$ states that the the samples from the more performant distribution tend to be smaller than the samples from the less performant distribution. We calculate the U statistic and p-value for each objective function as indicated in Table II. The critical value for a one-sided Mann-Whitney test at a significance level of $\alpha = 0.025$ and samples sizes $n_1 = n_2 = 30$ is given by 317.0. If our U statistic is less than the critical value and our p-value is less than 0.025, we can conclude that the minimum values obtained from one variant are lower than the other for a particular objective function, at a significance level of 0.025.

B. Two-tailed Test

We next consider the more rigorous two-tailed test which makes no prior assumptions regarding direction. In this case, the alternative hypothesis $\mathbf{h_a}$ is defined such that the two distributions are statistically different. We find the critical value for U such that if the observed value of U is less than or equal to the critical value, we reject $\mathbf{h_0}$ in favour of $\mathbf{h_a}$. The critical value can be obtained from the table according to the sample size $n_1 = n_2 = 30$ and our level of significance $\alpha = 0.05$ and is found to be 317.0. Our U statistic and p-value for the two-tailed test is calculated for each objective function as shown in Table II. If $U_{statistic} < U_{critical}$ and $p < 0.05$, we can conclude that the two distributions are not the same.

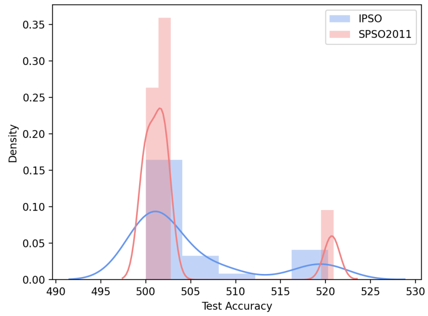
According to the conditions of both the one-tailed and two-tailed tests, we conclude that the distributions of performance results for the IPSO and SPSO2011 are different for all objective functions excluding the Ackley and Shifted and Rotated Ackley functions, and that the IPSO algorithm produces better results for these seven functions. For both

TABLE II: One-Sided and Two-Sided Mann Whitney Test, with 30 samples for each objective function

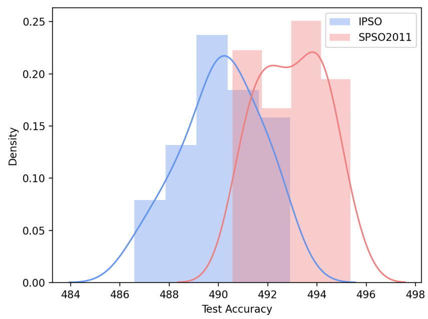
Function	Test	Simulation Parameter	Final Value	Reject h_0
Spherical	*One-Sided Mann-Whitney*	U Statistic p-value Sample size n	16.0 7.322×10^{-11} 30	Yes
	Two-Sided Mann-Whitney	U Statistic p-value Sample size n	16.0 1.464×10^{-10} 30	Yes
Ackley	*One-Sided Mann-Whitney*	U Statistic p-value Sample size n	364.0 0.1031 30	No
	Two-Sided Mann-Whitney	U Statistic p-value Sample size n	364.0 0.2062 30	No
Michalewicz	*One-Sided Mann-Whitney*	U Statistic p-value Sample size n	91.0 5.783×10^{-8} 30	Yes
	Two-Sided Mann-Whitney	U Statistic p-value Sample size n	91.0 1.157×10^{-7} 30	Yes
Katsuura	*One-Sided Mann-Whitney*	U Statistic p-value Sample size n	0.0 1.510×10^{-11} 30	Yes
	Two-Sided Mann-Whitney	U Statistic p-value Sample size n	0.0 3.020×10^{-11} 30	Yes
Shubert	*One-Sided Mann-Whitney*	U Statistic p-value Sample size n	0.0 1.510×10^{-11} 30	Yes
	Two-Sided Mann-Whitney	U Statistic p-value Sample size n	0.0 3.020×10^{-11} 30	Yes
Rotated and Shifted Ackley	*One-Sided Mann-Whitney*	U Statistic p-value Sample size n	380.0 0.1520 30	No
	Two-Sided Mann-Whitney	U Statistic p-value Sample size n	380.0 0.3042 30	No
Rotated and Shifted Michalewicz	*One-Sided Mann-Whitney*	U Statistic p-value Sample size n	81.0 2.546×10^{-8} 30	Yes
	Two-Sided Mann-Whitney	U Statistic p-value Sample size n	81.0 5.092×10^{-8} 30	Yes
Rotated and Shifted Katsuura	*One-Sided Mann-Whitney*	U Statistic p-value Sample size n	7.0 3.033×10^{-11} 30	Yes
	Two-Sided Mann-Whitney	U Statistic p-value Sample size n	7.0 6.065×10^{-11} 30	Yes
Rotated and Shifted Shubert	*One-Sided Mann-Whitney*	U Statistic p-value Sample size n	20.0 1.077×10^{-10} 30	Yes
	Two-Sided Mann-Whitney	U Statistic p-value Sample size n	20.0 2.154×10^{-10} 30	Yes

the one-tailed and two-tailed tests, a very small p-value is observed. Since a lower p-value indicates stronger evidence against the null hypothesis, we have strong confidence in our result that the values obtained from the IPSO algorithm are lower than those of the SPSO2011 algorithm for seven of the functions. These findings can be further understood by visualizing the respective distributions. It can be observed in Fig. 4, that the distributions of the accuracies obtained from the IPSO and SPSO2011 algorithms are very similar in the case of the Ackley function, therefore the two distributions are not found to be statistically different and $\mathbf{h_0}$ is not rejected. It can be seen from the shape of the distributions that both algorithms have a tendency to get stuck in the local minima located at $f(x) = 520$, but are also able to move towards the global minima located at $f^*(x) = 500$. This is further visualised in Fig. 5, showing the movement of the global best over the search duration in both cases.

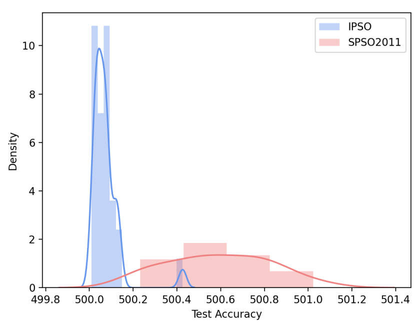
In contrast, it can be observed from Fig. 4 that the IPSO distribution provides better results for the other objective functions, as confirmed by the results of the Mann Whitney test, indicating that a sample from the IPSO distribution is on average lower than a sample from the SPSO2011 distribution.



(a) Ackley (ShR)



(b) Michalewicz (ShR)

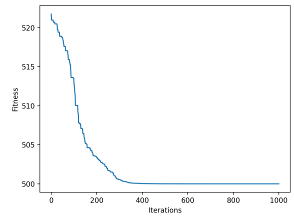


(c) Katsuura (ShR)

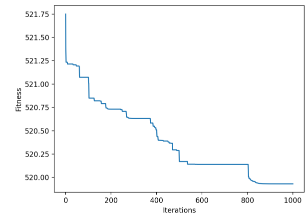
Fig. 4: Distribution of performance results for IPSO and SPSO2011 respectively for the rotated and shifted Ackley, Michalewicz and Katsuura functions. It can be observed that the distributions match very closely for the Ackley functions while the IPSO algorithm clearly exhibits better performance for the Michalewicz and Katsuura objective functions.

Therefore we find that IPSO performs better on all benchmark functions with the exception of the Ackley function.

We expect that this discrepancy stems from the increased difficulty of tuning the control parameters for the SPSO2011 algorithm. While inertia weight particle swarm optimization has clear stability criteria and theory supporting the optimal combination of parameters, the SPSO2011 variant has no corresponding guidelines. Therefore even though the SPSO2011 variant offers rotational invariance, this benefit is not realised as improved performance on the shifted and rotated functions is not observed.



(a) Global minimum



(b) Local minimum

Fig. 5: Movement of the global best for the shifted and rotated variant of the Ackley function with global optimum $f^* = 500$

It can also be noted that the only cases in which the distributions of performance results were deemed to be statistically the same was in the case of local best IPSO (LBest) as local best IPSO provided optimal results over global best IPSO for the Ackley and Shifted and Rotated Ackley functions only.

V. Conclusion

Particle swarm optimization provides a means of employing collective intelligence to optimize complex continuous functions. In this study, we have investigated the classic inertia weight particle swarm optimization in comparison to the more recent SPSO2011 variant with respect to nine benchmark functions, including shifts and rotations. Through a statistical analysis using the Mann Whitney U test, it was determined that the classic IPSO algorithm produced better results on seven out of the nine benchmark functions. Although the SPSO2011 variant offers the benefit of rotational invariance, no marked improvement over the IPSO algorithm was observed on the shifted and rotated objective functions. This unexpected result underlines the significance of careful control parameter selection as a core difference between the IPSO and SPSO2011 approaches is the respective presence and absence of corresponding stability criteria, and therefore insight into the parameter values which lead to more optimal results. Overall this investigation indicates that the choice of PSO algorithm is function dependent and that novel methods do not always provide clear benefits over the classic approach.

References

[1] Christopher W Cleghorn and Andries P Engelbrecht. Particle swarm variants: standardized convergence analysis. *Swarm Intelligence*, 9(2):177–203, 2015.

[2] Andries Petrus Engelbrecht. Particle swarm optimization: Global best or local best? In *2013 BRICS congress on computational intelligence and 11th Brazilian congress on computational intelligence*, pages 124–135. IEEE, 2013.

[3] Jing J Liang, Bo Y Qu, and Ponnuthurai N Suganthan. Problem definitions and evaluation criteria for the cec 2014 special session and competition on single objective real-parameter numerical optimization. *Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou China*, 635:490, 2013.

[4] Mauricio Zambrano-Bigiarini, Maurice Clerc, and Rodrigo Rojas. Standard particle swarm optimisation 2011 at cec-2013: A baseline for future pso improvements. In *2013 IEEE Congress on Evolutionary Computation*, pages 2337–2344. IEEE, 2013.