Tournament Topology Particle Swarm Optimization

Jason Kuo
Engineering for Professionals
Johns Hopkins University
Baltimore, Maryland 21218
Email: jkuo14@jhu.edu

John W. Sheppard
Gianforte School of Computing
Montana State University
Bozeman, MT 59717
Email: john.sheppard@montana.edu

***Abstract*—Particle swarm optimization (PSO) has become a popular algorithm for performing global numerical optimization; however, it is known that the topology of PSO has a large influence on its performance. Topologies with high connectivity can have fast convergence, but they are also susceptible to convergence to local minima. Topologies with low connectivity may avoid converging to local minima and achieve high quality solutions, but they tend to have slow convergence. In this paper, we propose a novel PSO topology based on a single-elimination tournament. In the proposed tournament topology, particles move up a tree structure through a fitness-based tournament. PSO updates then propagate information about the global best position from the top of the tree to the bottom. Experimental results on eleven benchmark functions show that the proposed topology can achieve both the high quality solutions of low-connectivity topologies and the fast convergence of high-connectivity topologies.**

I. Introduction

Particle swarm optimization (PSO) is a population-based meta-heuristic for optimizing continuous nonlinear functions [1]. The population (swarm) consists of individuals (particles), each of which represents a candidate solution through its position in the search space of the objective function [2]. The particles communicate with some subset (neighborhood) of particles by sharing information about the best positions that they have found. Particles are attracted to move towards both their personal best solutions and the best known solution among the particles in a defined neighborhood [3].

More specifically, suppose there is a swarm of m particles in an n-dimensional search space. The overall goal is for the swarm to find the global optimum of some objective function f. Let $\mathbf{x}_i$, $\mathbf{v}_i$, and $\mathbf{p}_i$ be the n-dimensional vectors respectively representing the current position of particle i, the current velocity of particle i, and the personal best position ever achieved by particle i. Let $\mathcal{N}_i \subseteq \mathcal{P}$ denote the neighbors of $\mathbf{x}_i$ drawn from population $\mathcal{P}$ according to some defined topology. Let particle b_i be the particle with the best known solution in $\mathcal{N}_i$, and let $\mathbf{p}_{b_i}$ denote the personal best position ever achieved by particle b_i. Finally, let $r_1, r_2 \sim \mathcal{U}(0,1)$, and ω, c_1, and c_2 be hyperparameters for inertia, personal best update, and neighborhood best update respectively. Then the velocity of each particle is updated as

$$\mathbf{v}_i = \omega\mathbf{v}_i + c_1 r_1(\mathbf{p}_i - \mathbf{x}_i) + c_2 r_2(\mathbf{p}_{b_i} - \mathbf{x}_i) \qquad (1)$$

and the position is updated as

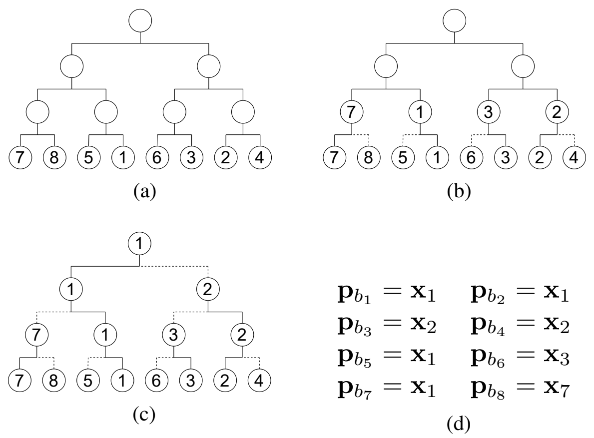
$$\mathbf{x}_i = \mathbf{x}_i + \mathbf{v}_i. \qquad (2)$$

An important issue of the PSO algorithm is the balance between exploration and exploitation [4]. Exploration finds good regions of the search space, and exploitation finds the best point in a good region [5]. However, high exploitation is susceptible quick convergence to non-optimal solutions, and high exploration may require a large number of iterations to find the global optimum [6]. Some approaches to address this problem include a linear decreasing inertia weight [7], a constriction coefficient [8], and heterogeneous search behaviors [9]. Some PSO variants use other methods to choose particles from which to learn. In the fully-informed PSO, a particle learns from all of its neighbors [10]. In chaotic heterogeneous comprehensive learning PSO, particles use tournament selection to select for each dimension a particle from which to learn [11].

Several studies have looked at less-connected neighborhood topologies. The original "star" topology uses the entire swarm as the subset, but it tends to favor exploitation [12]. Structured topologies can help reduce stagnation and premature convergence by slowing down the spread of good solutions [13]. In the "ring" topology, which favors exploration, each particle communicates with its two adjacent particles in index order [14]. In the "von Neumann" topology, each particle communicates with its four neighbors on a two-dimensional lattice. Ultimately, finding both quality and efficiency depends on the chosen topology [15].

Other PSO variants, such as the dynamic neighbor PSO [16] and the fitness distance ratio PSO [17], used dynamic topologies based on the most similar particles in each iteration. Similarly, Suganthan proposed a model where each particle communicates with a fixed number of its nearest neighbors in each iteration, where the number of neighbors increased as iterations increased [18].

The hierarchical version of PSO (H-PSO) proposed by Janson and Middendorf uses a tree structure and a hierarchy to define a topology [19]. The shape of the tree is defined by the branching factor d, the maximum out-degree of the interior nodes. The tree contains a node for each particle in the swarm, and each particle is influenced by itself and the particle that is directly above it in the hierarchy. In each iteration, the hierarchy is updated with $O(m)$ time complexity in a top-down manner: if the best solution of a particle in an inner



$$\mathbf{p}_{b_1} = \mathbf{x}_1 \quad \mathbf{p}_{b_2} = \mathbf{x}_1$$
$$\mathbf{p}_{b_3} = \mathbf{x}_2 \quad \mathbf{p}_{b_4} = \mathbf{x}_2$$
$$\mathbf{p}_{b_5} = \mathbf{x}_1 \quad \mathbf{p}_{b_6} = \mathbf{x}_3$$
$$\mathbf{p}_{b_7} = \mathbf{x}_1 \quad \mathbf{p}_{b_8} = \mathbf{x}_7$$

Fig. 1. Example tournament topology, where $i < j$ indicates particle i's best position has better fitness than particle j's. (a) Particles placed in leaf nodes. (b) Particle with the better fitness placed in parent node. (c) Fully populated tournament. (d) Neighborhood best for particle at b_i.

node is worse than the best solution found by the best particle in the child nodes, the two particles swap places.

Motivated by the need to balance exploration and exploitation, we propose a tournament topology for particle swarm optimization. Similar to H-PSO, the tournament topology defines a topology based on a tree structure; however, our proposed topology is dynamic in that it enters the particles in a single-elimination tournament and defines each particle's neighborhood as the subset of particles that it competed against in the tournament.

To present this new topology, rest of the paper is organized as follows. Section II introduces the proposed tournament topology for PSO. The experimental setup is discussed in Section III with the results presented in Section IV. Finally, we present our conclusions and future work (Section VI).

II. A Tournament Neighborhood Topology

In our proposed tournament topology, we represent a tournament bracket as a tree structure such that all leaf nodes are in the same level. All particles are placed in leaf nodes at the lowest level of the tree (Fig. 1a). In each iteration, a tournament is conducted to place particles into the internal nodes in a bottom-up manner. Particles in sibling nodes "compete," and the particle with the best known position is then placed in the corresponding parent node (Fig. 1b, 1c). The resulting neighborhood bests are then shown in Fig. 1d.

Updates are calculated using Equations (1) and (2). In the context of the tournament, a particle's neighborhood consists of the particles that it competed against. Then, with the exception of the particle at the root node, each particle is attracted to the particle in its parent node (Fig. 1d). The particle at the root node always has the global best position and is attracted to itself. We also note that the tournament topology with only two levels is equivalent to the star topology.

Using the tournament topology, information about the global best is propagated slowly through the levels of the tree from the top down. Unlike H-PSO, where each particle appears exactly once in the tree, each interior node of the tournament tree has the same particle in one of its child nodes. In this way, better particles influence more of the swarm. For example, if the tree has branching factor d and the swarm has m particles, at most $(d-1)\log_d m + 1$ particles are attracted to the global best in each iteration. In H-PSO, d particles are attracted to the global best in each iteration.

A major factor that may affect the performance of the tournament topology is the shape of the tree, i.e., the number of levels in the tree and the number of nodes at each level. Intuitively, information about the global best is propagated more quickly in trees with fewer levels. As in H-PSO, the shape of the tree is controlled by the branching factor $d \geq 2$, equal to the maximum out-degree of the interior nodes. Since the tree is constructed bottom-up, each interior node contains d child nodes, except at most one interior node per level may have fewer than d child nodes due to non-even divisibility.

Another factor that may affect the behavior of the tournament is the order of the particles in the last level of the tree. Initially, the particles may be placed randomly. However, since each particle always has its first match against the same group of other particles, this may result in a lack of diversity. To attempt to mitigate this, we also include a parameter $q \in [0,1]$ such that, at each iteration, the algorithm may reshuffle the order of the particles in the leaf nodes with probability q. Intuitively, a higher value of q would increase the chance that particles previously not in a path to the global best can appear in such a path, thus enabling them to receive information about the global best more quickly.

For a swarm with m particles, a tournament tree with branching factor d has $O(m)$ nodes. Each interior node requires finding its best child node, so each child node is checked once. Thus, the complete tournament has $O(m)$ time complexity. It may be computationally expensive and unnecessary to hold a tournament in every iteration. In some iterations, holding a tournament might make no changes to the best neighbor of each particle.

To investigate this observation, we include a parameter $k \in [0,1]$ corresponding to the probability of holding a tournament in a given iteration. If a tournament is not held, each particle's best neighbor corresponds to its best neighbor from the previous generation. A value of $k = 1$ results in holding a tournament after every generation, whereas a value of $k = 0$ corresponds to holding a tournament in only the first generation. Reducing the tournament frequency reduces the time complexity, at the risk of degraded performance. However, it may be possible to find a minimal value of k that does not lead to significantly poorer performance.

III. Experimental Approach

We conducted three experiments to test the performance of the tournament topology, measuring performance in terms of both solution quality and number of generations to get within a specified error threshold relative to the global optimum. In the first experiment, we studied the effects of the branching factor d and probability of reshuffling the leaf nodes q on the performance of the tournament topology. In the second experiment, we compared the performance of PSO with the tournament topology to the performance of PSO with four other topologies: star, ring, von Neumann, and H-PSO. In the third experiment, we studied the effect of the tournament frequency k on performance using the tournament topology.

A. Configurations

As a measure of solution quality, we measured the best function value returned by the algorithm after a fixed number of iterations. For the two experiments limited to our topology, we reported the best function value after 1,000 iterations, taking the mean across all trials. For the experiment comparing topologies, we also reported the best function value after 10,000 iterations, again taking the mean across all trials.

Since a good function value might also correspond to a local optimum, we also measured the number of iterations required to reach a specified error threshold relative to the global optimum. The thresholds are considered to indicate that the swarm's best particle is in the region of the global optimum [12], so the number of generations would give an indication of convergence time. We used the same thresholds as those used in the studies of Trelea [6] and Zhan *et al.* [20].

As in the work of Eberhart and Shi [21], the maximum number of iterations was set to 10,000. If the threshold was not met within the maximum number of iterations, the number of iterations was considered to be infinite and reported as 10,001. Thus, the mean across all trials is not appropriate for this measure, so we reported the median across all trials, the mean across all trials that met the threshold, and the percentage of trials that met the threshold. If no trials met the threshold, the mean reduced to 10,001.

B. Benchmark Functions

This paper used ten benchmark test functions: Ackley, Griewank, Rastrigin, Rosenbrock, Schaffer's f6, Schwefel's 1.2, Schwefel's 2.22, Schwefel's 2.26, Sphere, and Step [22], [23]. Schaffer's f6 was optimized in a 2-dimensional space, and the other nine functions were optimized in a 30-dimensional space. The Griewank function was also optimized in a 10-dimensional space (Griewank (10-D)), which is considered to be more difficult than the 30-dimensional variant (Griewank (30-D)) [12].

Boundary constraints $[l, u]$ were also included for eight of the ten benchmark functions such that, for each particle $\mathbf{x} = (x_1, ..., x_n)$ in an n-dimensional search space, the constraints required that $l \leq x_j \leq u$ for all $1 \leq j \leq n$. To handle particles that would exceed the bounds, we employed the periodic boundary handling method [24]. That is, the bounded search space was treated as periodic, so that exceeding one of the bounds "wraps around" to the other bound. In that case, the new value for x_j would be calculated as

$$x_j = \begin{cases} u - ((l - x_j) \bmod (u - l)) & \text{if } x_j < l \\ l + ((x_j - u) \bmod (u - l)) & \text{if } x_j > u \\ x_j & \text{otherwise} \end{cases}$$

All problems were treated as minimization problems with a global minimum function value of $f(\mathbf{x}^*) = 0$. The dimensions, bounds, and error threshold for each function are given in Table I.

C. PSO Parameters

All experiments were carried out with a swarm of 50 particles. For all five topologies compared, we tested combinations of the hyperparameters parameters in the ranges of $c_1 \in [0.1, 2.5]$, $c_2 \in [0.1, 2.5]$, and $\omega \in [0.05, 0.95]$. The combinations used in both experiments were chosen based on both solution quality and time to reach the global optimum and are given in Table II.

IV. Results

A. Branching Factor and Reshuffling

In our first experiment, we studied the effect of the branching factor d and reshuffling probability q on the performance of the tournament topology. Values of d were chosen in the set of $\{2, ..., 9\}$, and values of q were chosen in the range of $[0, 1]$. In total, 50 trials were performed for each combination of (d, q). For this experiment, we fixed $k = 1$.

We hypothesized that smaller values of d and q would yield better fitness values, take more generations to reach the global optimum, and have a higher success rate of reaching the error threshold. This was motivated by the fact that smaller values of d and q would result in a topology with lower connectivity.

Representative results of this experiment are shown as heatmaps for the best fitness after 1,000 iterations (Fig. 2), the median number of iterations to reach the threshold (Fig. 3), the mean number of iterations to reach the threshold (Fig. 4), and the success rate of reaching the threshold (Fig. 5). With respect to the reshuffling probability q, larger values tended to give both better fitness and fewer generations for the Ackley, Griewank (30-D), Schaffer's f6, Schwefel's 2.22, Sphere, and Step functions. Smaller values led to both better fitness and fewer generations for the Griewank (10-D), Rastrigin, and Schwefel's 1.2 functions. For the Rosenbrock function, larger values led to better fitness, but smaller values led to fewer generations. For Schwefel's 2.26 function, smaller values led to better fitness, but no values reached the threshold. Larger values also gave higher success rates for the Ackley and Step functions, whereas smaller values gave higher success rates for the Griewank (10-D) function.

Larger values of the branching factor tended to give both better fitness and fewer generations for the Griewank (10-D), Griewank (30-D), Rastrigin, Rosenbrock, Schwefel's 1.2, Schwefel's 2.22, and Sphere functions. Smaller values led to both better fitness and fewer generations for the Schaffer's f6 and Step functions. For the Rosenbrock function, smaller values led to better fitness, fewer median generations, and higher success rates; but larger values led to fewer mean generations. For Schwefel's 2.26 function, larger values led to better fitness. Larger values also gave higher success rates for the Griewank (10-D) function, whereas smaller values gave higher success rates for the Step function.

TABLE I
BENCHMARK FUNCTIONS WITH DIMENSIONS, BOUNDS, AND ERROR THRESHOLDS.

Function Name	**Dimensions**	**Search Space**	**Error Threshold**	**Function**
Ackley	30	$[-32, 32]$	0.01	$f_1 = -20\exp(-0.2\sqrt{1/D\sum_{i=1}^{D} x_i^2}) - \exp(1/D\sum_{i=1}^{D}\cos(2\pi x_i)) + 20 + e$
Griewank	10	$[-600, 600]$	0.1	$f_2 = 1/4000\sum_{i=1}^{D} x_i^2 - \prod_{i=1}^{D}\cos(x_i/\sqrt{i}) + 1$
Griewank	30	$[-600, 600]$	0.1	$f_3 = 1/4000\sum_{i=1}^{D} x_i^2 - \prod_{i=1}^{D}\cos(x_i/\sqrt{i}) + 1$
Rastrigin	30	$[-5.12, 5.12]$	100	$f_4 = \sum_{i=1}^{D}[x_i^2 - 10\cos(2\pi x_i) + 10]$
Rosenbrock	30	$(-\infty, \infty)$	100	$f_5 = \sum_{i=1}^{D-1}[100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2]$
Schaffer's f6	2	$[-100, 100]$	1×10^{-5}	$f_6 = 0.5 + \frac{\sin^2(\sqrt{x_1^2+x_2^2}) - 0.5}{[1+0.001(x_1^2+x_2^2)]^2}$
Schwefel's 1.2	30	$[-100, 100]$	100	$f_7 = \sum_{i=1}^{D}(\sum_{j=1}^{i} x_j)^2$
Schwefel's 2.22	30	$[-10, 10]$	0.01	$f_8 = \sum_{i=1}^{D}\lvert x_i\rvert + \prod_{i=1}^{D}\lvert x_i\rvert$
Schwefel's 2.26	30	$[-500, 500]$	2570	$f_9 = 418.9829D - \sum_{i=1}^{D} x_i \sin(\sqrt{\lvert x_i\rvert})$
Sphere	30	$(-\infty, \infty)$	0.01	$f_{10} = \sum_{i=1}^{D} x_i^2$
Step-2	30	$[-100, 100]$	0.1	$f_{11} = \sum_{i=1}^{D}(\lfloor x_i + 0.5\rfloor)^2$

TABLE II
TUNED VALUES OF c_1, c_2, AND ω FOR EACH TOPOLOGY.

Topology	**Parameters**		
	c_1	c_2	ω
Tournament	1.5	1	0.7
Star	1.5	1.5	0.7
Ring	1.25	1.25	0.7
von Neumann	1	1	0.7
H-PSO	1.75	1.5	0.5

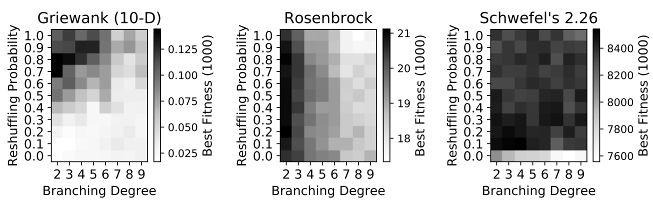


Fig. 2. Best fitness after 1,000 generations. Lighter regions represent better fitness.

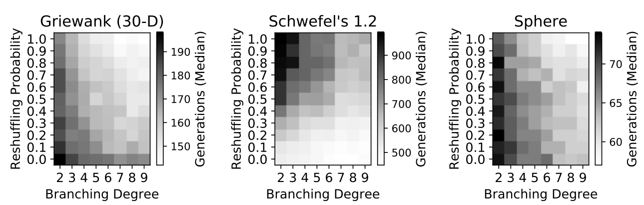


Fig. 3. Median number of iterations to reach error threshold. Lighter regions represent fewer generations.

For statistical analysis, we performed a Friedman test on each of the dependent variables. We found that over all combinations, there was a significant difference in: the best fitness after 1,000 generations [$F(87, 880) = 835.45$, $p < 0.001$]; the median number of generations to reach the threshold [$F(87, 880) = 1083.90$, $p < 0.001$]; the mean number of generations to reach the threshold [$F(87, 880) = 1103.73$, $p < 0.001$]; and the success rate [$F(87, 440) = 890.37$, $p < 0.001$].

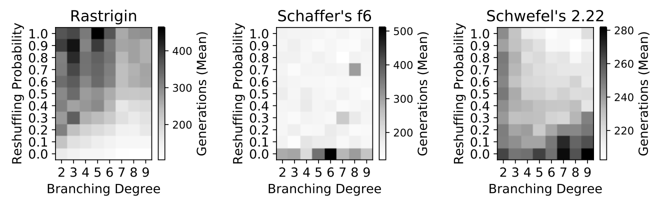


Fig. 4. Mean number of generations to reach error threshold. Lighter regions represent fewer generations.

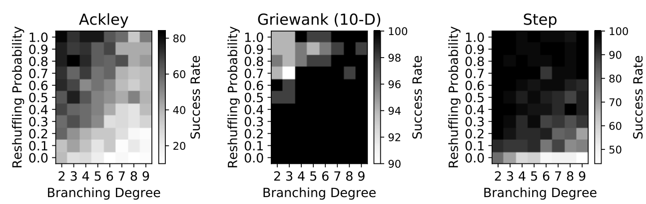


Fig. 5. Success rate of reaching error threshold. Darker regions represent higher success rates.

B. Topology Performance

In our second experiment, we compared the performance of our tournament topology to four other topologies: star, ring, von Neumann, and H-PSO. A total of 50 trials was performed for each topology. For this experiment, k was fixed at $k = 1$. For the tournament topology, values of the reshuffling probability and the branching factor were set for each function using the best results of the previous experiment. Values of the branching factor were tuned for H-PSO in a similar way. The chosen values are given in Table III.

We hypothesized that the tournament topology would reach a better fitness value than each of the other four topologies; that the tournament topology would reach the threshold more quickly than the ring, von Neumann, and H-PSO topologies, but more slowly than the star topology; and that the tournament topology would have a higher success rate of reaching the error threshold than the star topology; but the same success rate as the ring, von Neumann, and H-PSO topologies. These

TABLE III
TUNED VALUES OF d AND q.

f	Tournament Topology		H-PSO
	Branching Factor (d)	Reshuffling Probability (q)	Branching Factor (d)
f_1	3	0.8	3
f_2	2	0.1	3
f_3	6	0.6	4
f_4	2	0.1	5
f_5	9	0.9	9
f_6	7	0.4	6
f_7	7	0.1	8
f_8	6	0.9	9
f_9	9	0.1	7
f_{10}	9	1.0	9
f_{11}	6	0.9	2

hypotheses were motivated by the tournament topology having a higher connectivity than the ring, von Neumann, and H-PSO topologies; but a lower connectivity than the star topology, and the star topology likely converging to local optima.

The mean best fitness value at each iteration is plotted in Fig. 6. On three of the eleven functions, the tournament topology attained the best fitness after 1000 generations. On seven functions, H-PSO attained the best fitness after 1,000 generations, with the tournament topology in second. On Schwefel's 2.26 function, the tournament topology attained the worst fitness after both 1,000 and 10,000 generations. However, on seven functions, the tournament topology attained the best fitness after 10,000 generations. On three functions, H-PSO attained the best fitness after 1,000 generations, with the tournament topology in second. Across these experiments, the tournament topology and H-PSO generally performed similarly. Generally, the tournament topology seemed to have similar convergence times to the star, von Neumann, and H-PSO topologies, and a faster convergence than the ring topology.

The median generations to reach the error threshold, the mean iterations to reach the threshold in successful runs, and the percentage of runs that successfully reached the threshold (% Success) are listed in Table IV. The tournament topology reached the threshold in the fewest generations in terms of the median on seven functions, and in terms of the mean on seven functions. In terms of the median, the tournament topology ranked second on one function and fourth on two functions; the von Neumann topology ranked first on these three functions. In terms of the mean, the tournament topology ranked second on one function (the von Neumann topology ranked first), third on one function (H-PSO ranked first), and fourth on one function (the von Neumann topology ranked first).

TABLE IV
GENERATIONS TO REACH THE ERROR THRESHOLD. THE BEST RESULTS FOR EACH FUNCTION ARE IN BOLD.

Topology	f	Generations		% Success
		Median	Mean	
Tournament	f_1	**230.5**	$\mathbf{217.9 \pm 14}$	64
	f_2	179.5	$\mathbf{196.0 \pm 99}$	**100**
	f_3	**156**	$\mathbf{157.8 \pm 15}$	**100**
	f_4	202.5	227.8 ± 90	**100**
	f_5	21	21.8 ± 6	**100**
	f_6	**110.5**	$\mathbf{193.4 \pm 252}$	**100**
	f_7	**472**	$\mathbf{484.8 \pm 97}$	**100**
	f_8	**206**	$\mathbf{209.8 \pm 21}$	**100**
	f_9	**10001**	$\mathbf{10001 \pm 0}$	**0**
	f_{10}	**58**	$\mathbf{58.2 \pm 4}$	**100**
	f_{11}	**136**	325.0 ± 600	**98**
Star	f_1	10001	346.5 ± 6.4	4
	f_2	167.5	225.5 ± 230	82
	f_3	256	258.7 ± 29	94
	f_4	142	143.2 ± 30	**100**
	f_5	35	36.4 ± 12	**100**
	f_6	260.5	382.0 ± 600	82
	f_7	634	654.9 ± 109	**100**
	f_8	440	557.7 ± 282	**100**
	f_9	**10001**	$\mathbf{10001 \pm 0}$	**0**
	f_{10}	93	95.6 ± 12	**100**
	f_{11}	10001	2656.0 ± 3174	46
Ring	f_1	10001	525.9 ± 28	20
	f_2	212.5	584.8 ± 1589	84
	f_3	559.5	744.7 ± 611	94
	f_4	216.5	237.0 ± 104	**100**
	f_5	43	46.9 ± 12	**100**
	f_6	10001	2200.0 ± 2374	40
	f_7	1630	2010.0 ± 1074	**100**
	f_8	507	758.4 ± 715	98
	f_9	**10001**	$\mathbf{10001 \pm 0}$	**0**
	f_{10}	78	181.7 ± 77	**100**
	f_{11}	10001	317.7 ± 27	12
von Neumann	f_1	10001	274.8 ± 28	28
	f_2	**103**	283.5 ± 834	**100**
	f_3	211	210.6 ± 14	**100**
	f_4	**100**	$\mathbf{105.8 \pm 33}$	**100**
	f_5	**18**	$\mathbf{18.6 \pm 4.0}$	**100**
	f_6	344.5	751.6 ± 1125	94
	f_7	618.5	614.1 ± 110	**100**
	f_8	263	258.3 ± 21	**100**
	f_9	**10001**	$\mathbf{10001 \pm 0}$	**0**
	f_{10}	78	77.3 ± 5	**100**
	f_{11}	10001	10001 ± 0	0
H-PSO	f_1	266	269.8 ± 18	**98**
	f_2	150.5	213.1 ± 278	**100**
	f_3	196	197.0 ± 12	**100**
	f_4	160	164.1 ± 52	**100**
	f_5	36	38.4 ± 11	**100**
	f_6	156.5	482.1 ± 1280	98
	f_7	526	528.8 ± 77	**100**
	f_8	207	210.8 ± 14	**100**
	f_9	**10001**	$\mathbf{10001 \pm 0}$	**0**
	f_{10}	73.5	74.5 ± 6	**100**
	f_{11}	207.5	$\mathbf{213.4 \pm 32}$	**98**

For statistical analysis, we performed pairwise Wilcoxon signed-rank tests relative to the tournament topology for each of the star, ring, von Neumann, and H-PSO topologies. We found that the tournament topology reached a significantly better fitness value than the star topology after 1,000 generations [$W(11) = 11$, $p < 0.05$], and 10,000 generations [$W(11) = 11$, $p < 0.05$]; than the ring topology after 1,000 generations [$W(11) = 10$, $p < 0.025$], and 10,000 generations, [$W(11) = 11$, $p < 0.05$]; and than the von Neumann topology after 1,000 generations [$W(11) = 11$, $p < 0.05$], and 10,000 generations [$W(11) = 11$, $p < 0.05$]. Compared to H-PSO, there was no significant difference in fitness value after either 1,000 generations [$W(11) = 15$, $p > 0.1$], or 10,000 generations [$W(9) = 17$, $p > 0.2$].

Overall, the tournament topology reached the threshold in significantly fewer generations than the star topology in

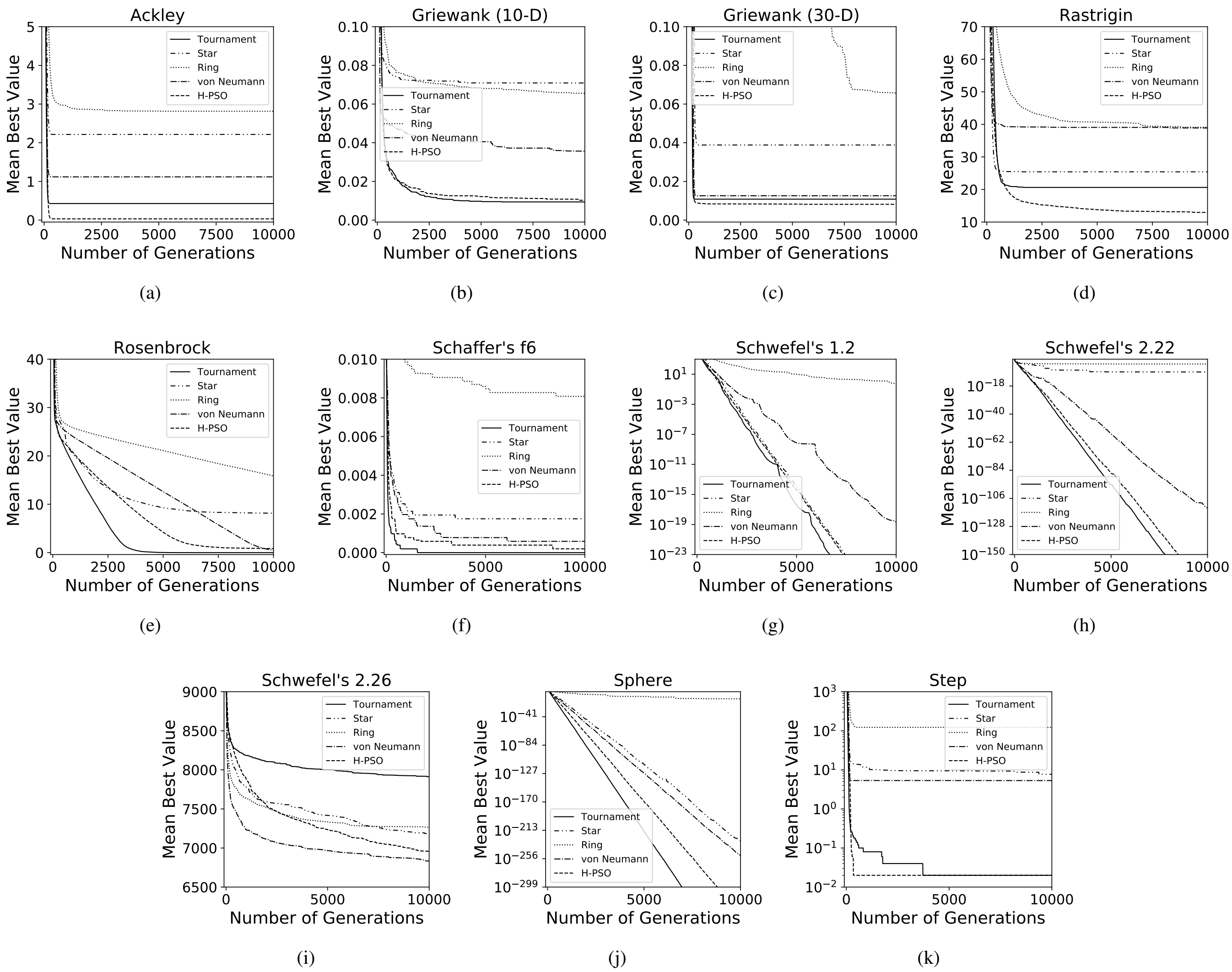


Fig. 6. Mean best fitness with the five topologies. Lower fitness values represent better results. Subfigures (g), (h), (j), and (k) are plotted on a log scale.

terms of both median [$W(10) = 5$, $p = 0.0125$] and mean [$W(10) = 4$, $p < 0.0125$]; and than the ring topology in terms of both median [$W(10) = 0$, $p < 0.001$] and mean [$W(10) = 1$, $p = 0.001$]. Compared to the von Neumann topology, there was no significant difference in terms of median [$W(10) = 12$, $p > 0.1$], but the tournament topology reached the threshold in significantly fewer iterations in terms of mean [$W(10) = 8$, $p = 0.025$]. However, compared to H-PSO, there was no significant difference in generations to reach the threshold in terms of either median [$W(10) = 11$, $p > 0.1$] or mean [$W(10) = 17$, $p > 0.2$].

None of the five topologies ever successfully reached the threshold on Schwefel's 2.26 function. On all functions, the tournament topology had the same or better success rate as the star, ring, and von Neumann topologies. The tournament topology and H-PSO had the same success rate on nine functions, having a better success rate on one function each.

We then performed two-sample t-tests on individual functions to compare the tournament topology and H-PSO (Table V). Compared to H-PSO, the tournament topology found significantly better solutions after 1,000 generations on one function, and after 10,000 generations on two functions. It found significantly worse solutions after 1,000 generations on three functions, and after 10,000 generations on two functions. The tournament topology required significantly fewer generations to reach the threshold on five functions, and significantly more on one function. These results suggest that the tournament topology may have faster convergence, whereas H-PSO may be more robust.

C. Tournament Frequency Effects

In our third experiment, we studied the effect of the tournament frequency k, on the performance of the tournament topology. Values of k were chosen in the set of $\{0.0, 0.1, 0.25, 0.5, 0.75, 0.85, 0.9, 0.95, 1.0\}$. A total of 50 trials was performed for each value of k. As in the previ-

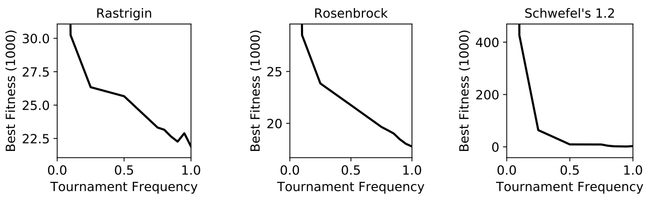


Fig. 7. Mean best fitness after 1,000 generations. Lower values represent better results.

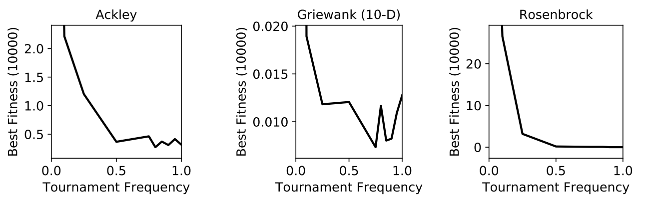


Fig. 8. Mean best fitness after 10,000 iterations. Lower values represent better results.

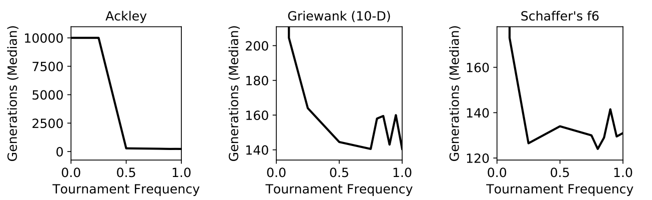


Fig. 9. Median generations to reach error threshold.

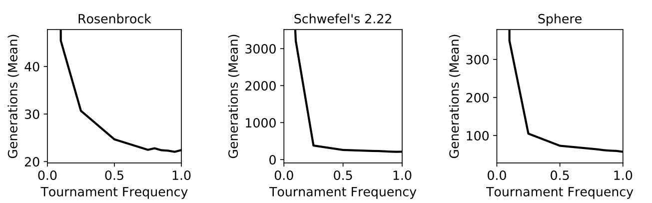


Fig. 10. Mean generations to reach error threshold.

ous experiment, values of the reshuffling probability and the branching factor were set to those in Table III.

We hypothesized that smaller values of k would yield worse fitness, take more generations to reach the error threshold, and have a lower success rate. This was motivated due to smaller values of k corresponding to more generations being required for a new global best particle to become the root of the tree. Thus, it may take more generations for other particles to receive knowledge about the global best.

Representative results of this experiment are shown in the graphs for the best fitness after 1,000 generations (Fig. 7), the best fitness after 10,000 generations (Fig. 8), the median generations to reach the error threshold (Fig. 9), the mean generations to reach the error threshold (Fig. 10), and the success rate of reaching the threshold (Fig. 11).

In general, smaller values of the tournament frequency

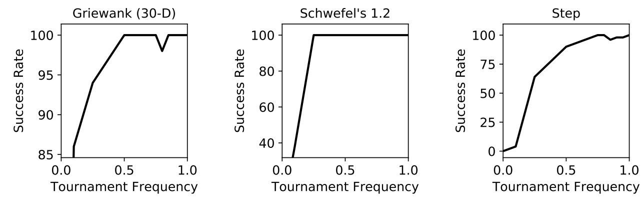


Fig. 11. Success rate of reaching error threshold.

TABLE V
TOURNAMENT (T) VS. H-PSO (H). BOLD INDICATES STATISTICAL SIGNIFICANCE AT LEVELS INDICATED IN THE TEXT.

f	1000	10000	Median	Mean	Success
f_1	**H**	**H**	T	**T**	H
f_2	H	T	H	T	
f_3	H	H	T	**T**	
f_4	**H**	**H**	H	**H**	
f_5	**T**	**T**	T	**T**	
f_6	T	T	T	T	T
f_7	H	**T**	T	**T**	
f_8	H		T	T	
f_9	**H**	H			
f_{10}	T		T	**T**	
f_{11}	H		T	H	

resulted in worse fitness, more generations, and lower success rates than higher values. For most of the functions, as k decreased from 1, performance appeared to be similar to $k = 1$ before decreasing, such as Fig. 8 (Rosenbrock) and Fig. 10.

Next, we performed a Friedman test on each of the dependent variables. We found that over all values of k, there was a significant difference in the best fitness after 1,000 generations $[F(9, 100) = 47.00,\ p < 0.001]$ and after 10,000 generations $[F(9, 100) = 41.15,\ p < 0.001]$; in the median number of generations to reach the threshold $[F(9, 100) = 55.16,\ p < 0.001]$; in the mean number of generations to reach the threshold $[F(9, 1000) = 51.93,\ p < 0.001]$; and in the success rate $[F(9, 100) = 31.31,\ p < 0.001]$.

We then performed pairwise Wilcoxon signed-rank tests relative to a tournament frequency of $k = 1$ for each of the nine other values. Tournament frequencies up to 0.5, with the exception of 0.25, led to significantly worse solutions after 1,000 generations; and only a frequency of 0.0 led to significantly worse solutions after 10,000 generations. Tournament frequencies up to 0.95 took significantly more generations to reach the threshold in terms of the median, and tournament frequencies up to 0.25 took significantly more generations in terms of the mean. Tournament frequencies up to 0.1 had significantly lower success rates. These results suggest that the tournament frequency could be reduced to certain values without leading to significantly worse performance.

V. DISCUSSION

The tournament topology appeared to generally perform better than H-PSO on functions with certain properties (Table V): non-separable functions (f_1, f_2, f_3, f_5, f_6, f_7), unimodal functions (f_7, f_8, f_{10}), and multimodal functions with few local minima (f_5, f_{11}). In particular, the tournament topology performed equally or better than H-PSO on all five measures on f_5 and f_6. On the other hand, H-PSO performed equally or better than the tournament topology on all five measures on f_4 and f_9, the functions that are both multimodal and separable.

In comparison, the star topology has generally been preferred for unimodal and separable functions, and the ring topology has been preferred for multimodal and non-separable functions [5]. Since unimodal functions and separable functions are considered to be simpler optimization problems, the tournament topology shows promise for optimization of non-separable functions, including the more difficult optimization of multimodal, non-separable functions.

VI. Conclusions and Future Work

In this paper, we presented a tournament topology for particle swarm optimization. The proposed topology was tested under different parameter settings for the branching factor and the probability of reshuffling the leaf nodes. The results showed that optimal parameter values could significantly improve performance in terms of both solution quality and time to reach the global optimum. The proposed topology was also tested under different parameter settings for the tournament frequency. The results showed that reducing the tournament frequency could reduce time complexity without significantly reducing performance in terms of either solution quality and time to reach the global optimum.

The proposed topology was also tested on eleven benchmark functions against the star, ring, von Neumann, and H-PSO topologies. Compared to H-PSO, it was able to find solutions of similar fitness and reached the global optimum in similar, if not less, time. Compared to the star, ring, and von Neumann topology, it was able to find solutions of better fitness and reached the global optimum in less time. Based on these results, we found that the tournament topology was able to achieve the fast convergence of high-connectivity topologies while also achieving the high quality solutions of low-connectivity topologies, maintaining balance between exploration and exploitation.

The topology presented in this paper used a rudimentary method for constructing the tournament tree. In particular, further research should be done to investigate the impact of more balanced tree designs into the topology.

In addition, similar to the adaptive variant of H-PSO, further research will investigate an adaptive version of the tournament topology, such as changing the value of the branching factor over the course of the run. In the adaptive variant of H-PSO, adapting an existing tree to the new branching factor would be expensive. Since the tree in the tournament topology is filled in a bottom-up manner for each tournament, however, changing the branching factor would require no additional time.

References

[1] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of the International Conference on Neural Networks (IJCNN)*, vol. 4, 1995, pp. 1942–1948.

[2] I. Boussaïd, J. Lepagnot, and P. Siarry, "A survey on optimization metaheuristics," *Information Sciences*, vol. 237, pp. 82—-117, 11 2013.

[3] D. Bratton and J. Kennedy, "Defining a standard for particle swarm optimization," in *Proceedings of the IEEE Swarm Intelligence Symposium (SIS)*, 2007, pp. 120–127.

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

[5] T. Blackwell and J. Kennedy, "Impact of communication topology in particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, vol. 23, no. 4, pp. 689–702, 2019.

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

[7] Y. Shi and R. Eberhart, "A modified particle swarm optimizer," in *1998 IEEE International Conference on Evolutionary Computation Proceedings. IEEE World Congress on Computational Intelligence*, 1998, pp. 69–73.

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

[9] A. P. Engelbrecht, "Heterogeneous particle swarm optimization," in *Proceedings of the 7th International Conference on Swarm Intelligence*, 2010, pp. 191–202.

[10] R. Mendes, J. Kennedy, and J. Neves, "The fully informed particle swarm: simpler, maybe better," *IEEE Transactions on Evolutionary Computation*, vol. 8, no. 3, pp. 204–210, 2004.

[11] D. Yousri, D. Allam, M. Eteiba, and P. Suganthan, "Chaotic heterogeneous comprehensive learning particle swarm optimizer variants for permanent magnet synchronous motor models parameters estimation," *Iranian Journal of Science and Technology, Transactions of Electrical Engineering*, vol. 44, pp. 1299–1318, 2020.

[12] J. Kennedy and R. Mendes, "Neighborhood topologies in fully informed and best-of-neighborhood particle swarms," *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)*, vol. 36, no. 4, pp. 515–519, 2006.

[13] N. Lynn, M. Z. Ali, and P. N. Suganthan, "Population topologies for particle swarm optimization and differential evolution," *Swarm and Evolutionary Computation*, vol. 39, pp. 24–35, 2018.

[14] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proceedings of the Sixth International Symposium on Micro Machine and Human Science (MHS)*, 1995, pp. 39–43.

[15] J. Kennedy and R. Mendes, "Population structure and particle swarm performance," in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, vol. 2, 2002, pp. 1671–1676.

[16] X. Hu and R. Eberhart, "Multiobjective optimization using dynamic neighborhood particle swarm optimization," in *Proceedings of the IEEE Congress on Evolutionary Computation. (CEC)*, vol. 2, 2002, pp. 1677–1681.

[17] T. Peram, K. Veeramachaneni, and C. K. Mohan, "Fitness-distance-ratio based particle swarm optimization," in *Proceedings of the IEEE Swarm Intelligence Symposium (SIS)*, 2003, pp. 174–181.

[18] P. N. Suganthan, "Particle swarm optimiser with neighbourhood operator," in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, vol. 3, 1999, pp. 1958–1962.

[19] S. Janson and M. Middendorf, "A hierarchical particle swarm optimizer and its adaptive variant," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 35, no. 6, pp. 1272–1282, 2005.

[20] Z. Zhan, J. Zhang, Y. Li, and H. S. Chung, "Adaptive particle swarm optimization," *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, vol. 39, no. 6, pp. 1362–1381, 2009.

[21] R. C. Eberhart and Y. Shi, "Comparing inertia weights and constriction factors in particle swarm optimization," in *Proceedings of the IEEE Congress on Evolutionary Computation. (CEC)*, vol. 1, 2000, pp. 84–88.

[22] M. Jamil and X. Yang, "A literature survey of benchmark functions for global optimization problems," *Int. Journal of Mathematical Modelling and Numerical Optimisation*, vol. 4, no. 2, pp. 150–194, 2013.

[23] J. M. Dieterich and B. Hartke, "Empirical review of standard benchmark functions using evolutionary global optimization," *Applied Mathematics*, vol. 03, no. 10A, pp. 1552–1564, 2012.

[24] W. J. Zhang, X. F. Xie, and D. C. Bi, "Handling boundary constraints for numerical optimization by particle swarm flying in periodic search space," in *Proceedings of the IEEE Congress on Evolutionary Computation (CEC)*, vol. 2, 2004, pp. 2307–2311.