

# CPSO with Spatially Meaningful Neighbors

Peter Wilson, *Member, IEEE*, Ombuki Berman, *Fellow, OSA*,

**Abstract**—Particle Swarm Optimization (PSO) is a population-based computational intelligence algorithm to optimize continuous functions [1] through communication among its particles. This communication often works via naive, static communication channels. One improvement would be to utilize the dynamic distances of particles to dictate communication with particles that are positionally close. However, such a technique would require intensive calculations to determine these distances and is therefore not viable for larger dimension problems. This paper proposes and tests an alternative version, showing that using cooperative swarms to divide up higher dimension problems makes this new form of communication possible for more complex problems.

## I. INTRODUCTION

Particle Swarm Optimization is a population-based computational intelligence algorithm based on the concept of birds flocking patterns [1]. The power of population allows it to find optimal solutions to a problem over multiple generations. By randomizing potential solutions (particles) and continuously changing them based on information already determined, the swarm is able to converge on solutions without an exhaustive search. As such, communication between particles is key. The standard form of communication (follow the leader) has the overall best solution telling the rest of the particles to follow. As the PSO algorithm matured, different communication methods have been introduced [2].

One method, called spatially meaningful neighbors, is to have particles communicate only if they are searching the same subspace [3] (pockets of search space that have their own local optimal value), allowing for more cooperation and a more thorough search of this space. To determine communication, distances between each particle would be calculated. This quickly becomes a bottleneck as dimensions increase past 3-dimensions since calculating the distance for each iteration can become incredibly time consuming [4].

A way of solving this dimension problem would be to split the larger problem into a number of smaller problems. A variation on the standard PSO algorithm is called Cooperative PSO [5] does exactly that, breaking dimensions down into smaller problems that collaborate to solve the main problem. By using CPSO in conjunction with the spatially meaningful neighbors, this thesis aims to bring that technique to higher dimension problems while still maintaining the subspace searching benefits it offers.

## II. BACKGROUND

### A. Particle Swarm Optimization

1) *Overview:* Particle Swarm Optimization (PSO) is an iterative population-based computational intelligence algorithm which searches for optimal solutions to non-linear continuous problems. Computational intelligence is a domain of study in which the patterns and behaviours of intelligent agents (such as humans or animals) are used to help solve other types of problems [6]. PSO specifically draws from the flight patterns of birds in their search of food, as well as the schooling pattern of fish. In both cases, the animals tend to follow each other from a distance which allows each to get a different vantage point [1]. When one of the fish/birds find the best source of food (or the best place to land) they all converge toward that point.

2) *Particle:* PSO builds off this metaphor by representing the birds or fish as particles. A particle has a position, velocity and personal best value, updating them with each iteration. The position of the particle maps logically to the location of where the bird is on a map. The x,y,z,... coordinates of the position also correlate to the input values in the problem which we are trying to optimize. The greater number of input values (variables) the problem contains, the higher dimensions are needed to represent the problem. The initial position of the particle is randomly generated to allow the swarm to effectively test the entire search space. The velocity represents the direction and speed of the particle, dictating the new position after each iteration. The velocity is also updated at the end of each iteration, and is influenced by: the current velocity, the personal best value (the best value previously found by the particle) and the global/neighborhood best (the best value found by all the particles neighbors). The velocity and position are updated using the functions below.

3) *Position Update Formula:*

$$x_i = x_i + v_i$$

Updating the position is as simple as taking the current position ( $x_i$ ) and adding the current velocity ( $v_i$ ) to it. For multi-dimensional problems where the particles have multiple dimensions, this addition will need to be done for each dimension. In this case, the  $i$  represents the index of the position.

4) *Velocity Update Formula:*

$$v_i = \alpha v_i + \omega C_1(y_i - x_i) + \lambda C_2(\hat{y}_i - x_i)$$

The velocity update function has a little more to it. The new velocity is determined by a combination of its current velocity, the distance it is from the particles personal best and the distance it is from the global or neighborhood best [1]. In this case,  $y_i$  represents the personal best for that index, and

P. Wilson is with the Computer Science Department, Brock University, St. Catharines, ON, Canada.

Professor Ombuki Berman is also with Brock University.

Manuscript received April 19, 2005; revised January 11, 2007.

$\hat{y}_i$  represents the global best. The influence that each has is determined by the inertial weight ( $\alpha$ ), the cognitive weight ( $\omega$ ) and social weight ( $\lambda$ ) respectively. These values typically add up to 1 and determine the percentage of influence they have on the new velocity. Additionally, the formula also makes use of a random component. The values  $C_1$  and  $C_2$  are randomly generated values in the range [0,1] that add more randomness and aid in keeping the swarm from converging too quickly.

5) *Swarm*: The swarm is the combination of a number of particles working in unison to solve the same problem. The particles are connected as a topology which allows them to share information allowing them to work together to converge on an optimal solution. With each iteration, the swarm tries a number of different potential solutions and gains information to help it decide the overall best solution. The algorithm to accomplish this is in Table I.

TABLE I. PSO PSEUDO CODE

```

Initialize particles randomly
while current iteration < max iterations do
  for each particle  $j \in \text{swarm}$  do
    //Update the personal and global bests
    if  $f(j.\vec{x}) > f(j.\vec{y})$  then
       $j.(y) \leftarrow j.(x)$ 
    end if
    if  $f(j.\vec{y}) > f(\vec{y})$  then
       $(y) \leftarrow j.(y)$ 
    end if
  end for
  for all particles  $\in \text{swarm}$  do
    Update the Velocity
    Update the Position
  end for
end while

```

Every iteration, the swarm will update the personal best for each particle (represented by  $\vec{y}$ ) and update the global best if a new best value is found [1]. The value is calculated through passing the solution set into the fitness function (see below). Each particle is then iterated through to update velocity and position. It is important to complete all global/personal best updates prior to updating the velocities since those values influence the new velocity and therefore should be correct for those calculations.

6) *Fitness Function*: Finding an optimal solution to a problem first requires a way to evaluate individual solutions in relation to the optimization problem. The fitness function takes the list of input values which make up the solution and determines how well it fits the problem. This will return a single value that helps the swarm determine the success of the solution, ultimately helping it converge on the best value. The fitness function itself is the representation of the problem that we are looking to optimize and is how we tailor the PSO algorithm to solve different problems.

7) *Parameters*: Parameters can be altered to affect the success of the algorithm [1]. Changing these values can lead to different results depending on the problem at hand. Experimentation is often required in order to get the best performance out of the algorithm [7].

**Max Iterations**: How many iterations the swarm will go through in to search for the optimal solution. Greater Iterations provides the algorithm with more time to converge to an

optimal solution, though better algorithms should require less iterations to hit this target.

**Swarm Size**: The number of particles the swarm employs to search for the optimal solution. Increasing this gives the swarm more information for each iteration and allows for a higher likely hood of finding the overall best solution.

**Inertial Weight ( $\alpha$ )**: Determines the influence the current velocity has on future velocities. A large inertial weight will stubbornly lead to the particle on its current direction allowing for greater exploration. A low inertial weight will give more influence to the communication, allowing for greater exploitation.

**Cognitive Weight ( $\omega$ )**: The effect the particles personal best solution has on its future velocity. Increasing this will keep the particle from venturing too far from this solution and only explore the neighboring space unless it leads to better results.

**Social Weight ( $\lambda$ )**: Determines the effect neighborhood particles (global best) have on the current particle. Different neighborhood layouts/topologies can change the effect of the social communication (See II-C). A low value will have the particles mostly work on their own, whereas a large social weight will increase exploitation.

## B. Co-operative Particle Swarm Optimization

Although PSO very successful for solving continuous problems with reasonable dimensions, it often struggles as those dimensions increase. For a problem with a large enough number of inputs, it can be difficult for the swarm to hone in on an optimal solution given the breadth of the search space size, often getting stuck in local optima. This issue is referred to as the curse of dimensionality and affects most other stochastic optimization algorithms [5].

This problem spawns from all the dimensions updating at the same time prior to calculating the fitness. This can result in a situation where a new solution is objectively worse for most dimensions but leads to a better fitness value overall due to a better value for a single dimension [8]. For instance, if the intention of the algorithm is to minimize the sum of the values (represented as  $\sum_{n=0}^d x_n$ ) the position [5,5,5] will give a fitness value of 15, while the position [6,1,7] returns a better value of 14 even though the first and last dimension values are moving away from the optimal solution [8]. This can influence the velocity, leading to the particles heading in the wrong direction.

CPSO attempts to solve the curse of dimensionality by instead dividing the problem into multiple sub-problems [5]. This allows a subset of the dimensions to be solved separately by partitioning the dimensions and solved with their own swarm. Each iteration will test its solution along with static random values in place of the other inputs. Different variants of the algorithm introduce alternative ways of partitioning the larger problem. An issue which arises from this partitioning where it can remove potential input dependencies from the original problem. If two dimensions are required to work together to improve the fitness, they will be no longer able to do so directly if partitioned to different swarms.

1) *CPSO-S*: The original version of CPSO is called CPSO-S. Developed by Van den bergh and Engelbrecht [5], it tackles the problem by dividing an n-dimensional problem into n 1-dimensional swarms. The algorithm will iterate through each single dimension swarm, operating a standard PSO for each partition (as seen in Table II).

TABLE II. CPSO-S PSEUDO CODE

```

Create and initialize n one-dimensional PSO Swarms
Randomly initialize solution vector
while current iteration < max iterations do
  for each swarm  $i \in \text{swarms}$  do
    for each particle  $j \in i$  do
      //Update the personal and global bests
      if  $f(j.\vec{x}) > f(j.\vec{y})$  then
         $j.\vec{y} \leftarrow j.\vec{x}$ 
      end if
      if  $f(j.\vec{y}) > f(\vec{y})$  then
         $\vec{y} \leftarrow j.\vec{y}$ 
      end if
      Update the Velocity
      Update the Position
    end for
  end for
end while

```

Since 1 dimensional problems are trivial to solve through PSO, separating the problem into 1-dimensional problems significantly simplifies the end goal. However, since each swarm operates almost in a vacuum, it removes all variable dependencies from the original problem, which can make it difficult to find the optimum solution.

2) *Variant: CPSO-S<sub>k</sub>*: The first variant proposed by Van den bergh and Engelbrecht is CPSO-S<sub>k</sub>[5]. It evenly divides the problem into k swarms (or as close to even as possible). This attempts to combine dependent variables to keep them intact. The rest of the algorithm is identical to CPSO-S (See Table II). Though this does give the potential to combine dependent terms, it still does so blindly, therefore there is no guarantee.

3) *Variant: CPSO-R<sub>k</sub>*: Another potential solution to the partitioning problem would be to randomly partition into a set number of blocks, which is implemented in CPSO-R<sub>k</sub> (also known as Random CPSO). This can lead to a larger likelihood of combining dependant variables depending on block size (k). If it does successfully make those connections it helps the chances for a successful run. Once the partitions are established, the CPSO-S algorithm is run to solve the PSOs (See Table II).

### C. Neighborhood Topologies

The neighborhood topology of a PSO dictates how the swarms individual particles are allowed to communicate with each other. This communication is the real strength of the PSO as it allows its particles to work together in solving the problem. This particle communication is achieved through the use of global best values. For each particle, when it is updating its velocity for the next iteration, it will consult the global best in its neighborhood to aid in determining its next move. In the standard PSO (Section II-A5), all particles are able to communicate with each other. This means that all particles in the swarm are led by the overall best particle

in the entire swarm. This behaviour can sometimes lead to a naive swarm that quickly converges on an optimal solution without fully exploring the entire search space of the problem. To combat this, a local PSO [9] was proposed, which utilizes different, more constrained, neighborhood topologies in hopes of creating a better PSO. A visual representation of these topologies can be found on Figure 1.

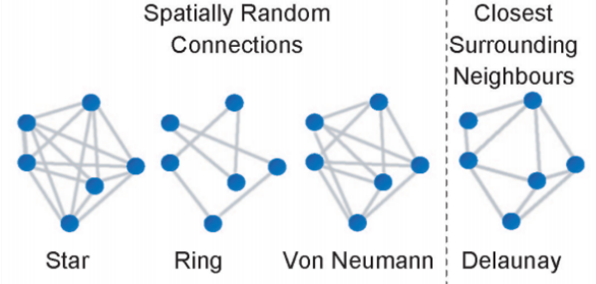


Fig. 1. Visual Representation of the Neighborhoods[3]

1) *Common Spatially Random Topologies*: There are a number of neighborhood topologies that are currently being used and tested in the PSO community. Many of these topologies have their own pros and cons and can lead to different or better results depending on the problem in question. These topologies are known as spatially random since the location of the particles at any point in the calculation has no bearing on what particles it is connected to [3].

#### 3emOpt Star

The star topology is the most common since it is used in the standard PSO. As mentioned, star topology connects every particle with each other. That means the neighborhood best for every particle is the same as the global best position in the swarm. The star topology uses this information to quickly converge onto the optimal best without a thorough exploration of the search space.

#### Ring

The ring topology changes the neighborhood topology by reducing the connections per particle to just two. All the particles are ultimately still connected to each other, but through multiple intermediary particles. This topology slows the transfer of information among the particles, decreasing the impact the global best has on the swarm itself[9]. This helps to slow down the convergence of the swarm, allowing for a more thorough investigation of the search space. However, as a result, it requires a much larger number of iterations to find an optimal solution. Additionally, the connections are often determined randomly during the initialization of the particles. As a result, the effect of the topology can be dependant on the success of that initial setup.

#### Von Neumann

The Von Neumann topology expands on the ring topology by increasing the number of connections each particle has while still maintaining the slow information transfer rate within

the network[10]. This helps to speed up convergence while allowing for exploration. The particles are connected in a cubic-lattice arrangement [3] where particles are connected to particles on the left, right, above and below. This is not determined by actual location of the particle, but often based on the index value that particle is in the swarm array, where above and below are determined by some sort of offset based on the swarm size (See Figure 2). This means it does not take advantage of positioning information and the connections are not updated throughout the completion of the run.

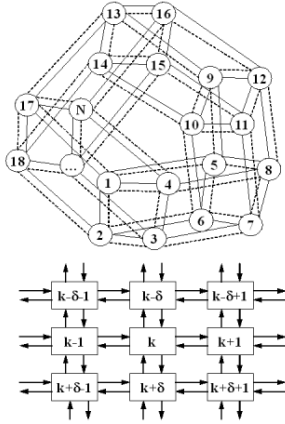


Fig. 2. Von Neumann[9]

2) *Spatially Meaningful Neighbors* : Where the updates to the spatially random neighborhood topologies allowed for more time in searching the search space, the actual searching ultimately ends up being mostly random. This is due to the fact that these connections are both randomly selected and only determined once during the life cycle of the swarm. If the connections were dynamically chosen based on information within the swarm, that information could be used to better direct the collaboration between particles. Lane, Engelbrecht and Gain proposed [3] using the distance between particles to determine which particle is connected to which. It was theorized that particles that are close together have a high probability to be searching the same subspace (a pocket of space that has its own local max/min). The proposed new neighborhood topology called spatially meaningful neighbors uses the Delaunay triangulation algorithm to connect all particles with the neighbors that are closest to them. Those connections then represent which particles they can talk to and draw information from. Decisions are then made to avoid or proceed in searching a space based on how the particles are moving (See Section II-D). This led to a more guided search and a higher success rate in their tests [3].

#### D. Dynamic Connections

Just having the list of the closest neighbors for each particle is not in itself inherently helpful for the PSO. Of this list, the particles will now want to determine which of these particles

are working together so we can have the particles cooperate to find a local optima [3]. Setting up connections between these particles will allow them to communicate and share information, allowing them to more quickly converge on that optima. An example of particles working together can be found in Figure 3. A particle is considered to be working with another if:

- Particle  $P_1$  is following behind particle  $P_2$ . i.e. both are heading in the same direction. In this case, a directed connection will be established so the particle in front can send information to the trailing particle[3]
- Both particle  $P_1$  and particle  $P_2$  are heading towards each other but not passing each other. Here we would want to set up an undirected connection so that both particles can share information [3].

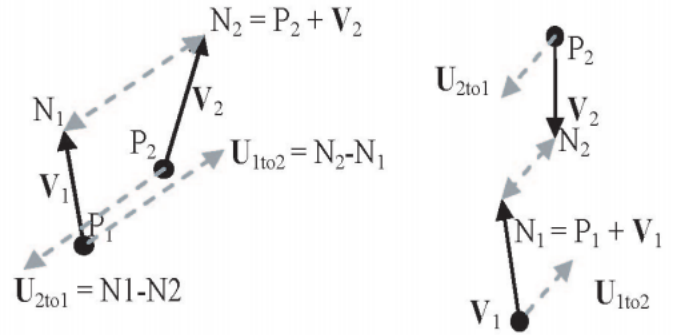


Fig. 3. How to know if the particles are working together[3]

Both cases can be tested with a simple equation:

$$V_1 \cdot U_{1to2} > 0$$

In the function above, the  $V_1$  represents the current velocity of  $P_1$  and  $U_{1to2}$  is the vector representing the distance from  $P_1$  to  $P_2$  after the two particles' next position update. Using the dot product on both these vectors allows us to determine the angle between them. If the dot product is greater than 0, we know the angle between the two vectors is less than 90 degrees. This outcome can only occur if it satisfies one of the two.

A caveat to this would be that we still wish to encourage local exploration. Delaunay triangulation does not guarantee that the particles are close together. If two neighboring particles are working together, but are still reasonably far away, having them communicate will inhibit local exploration. Therefore we need to set a threshold in which we want the particles to search within before they start communicating. Table III details the process of finding the best neighbor to communicate with. In this algorithm  $P_f$  represents the personal best position of the fittest connected neighbor.  $P_c$  represents the personal best position of the closest fitter connected neighbor. The idea behind the code below is that if the overall fittest neighbor is too far to be considered as searching the same subspace (in this case, is further than our defined *localExploitationRatio*),

then the algorithm will instead choose the closer of the neighbors as long as the fitness value is better than its own.

TABLE III. CHOOSING THE BEST NEIGHBOR

```

input:  $N_i$  particle  $i$ s closest neighbors
procedure CHOOSEBESTNEIGHBOR( $N_i$ )
  hasConnectedNeighbors = false
   $P_f \leftarrow \min(N_k)$ 
  for  $k \leftarrow 1$  to neighborsetSize do
    if workingtogether( $P_i, P_k$ ) and  $\text{dist}(X_i - P_c) < \text{dist}(X_i - P_k)$  and
     $f(P_k) < f(X_i)$  then
       $P_c \leftarrow P_k$ 
      hasConnectedNeighbors  $\leftarrow$  true
    end if
  end for
  localExploitationRatio  $\leftarrow$  swarm.diameter/200
  if hasConnectedNeighbors and
   $\text{distance}(X_i - P - c)/\text{distance}(X_i - P - f) > \text{localExploitationRatio}$ 
  and
   $V_i < 2 * \text{distance}(X_i - P_x)$  then
     $P_n \leftarrow P_c$ 
  else
     $P_n \leftarrow P_f$ 
  end if
  return  $P_n$ 
end procedure

```

#### E. Delaunay triangulation

1) *Triangulation*: Triangulation in computational geometry is the decomposition of a polygon or set of points into a set of pairwise non-intersecting triangles in which all points in the set or edge in the polygon are an edge in at least one of newly the formed triangles [3]. In  $n$ -dimensions, the triangles are instead represented as simplices (the  $n$ -dimensional equivalent)[11]. Triangulation is mainly used in computer graphics to simplify complex shapes into easily drawn triangles, though the information gathered from this process can be utilized in a number of different problems. An optimal triangulation is one that minimizes the total length of all the edges, otherwise known as having the minimum weight. This means that an optimal triangulation will, on average, connect all the points to their nearest neighbors. Definitively finding the optimal triangulation of a set of points is considered to be an NP-Hard problem [?]. However, there are a number of triangulation algorithms that do a very good job of coming close to optimal.

2) *Delaunay triangulation*: Delaunay triangulation (DT) is an efficient triangulation method in computational geometry that attempts avoid sliver triangles (or thin triangles not often found in minimum weight triangulations) by maximizing the minimum angle of all the triangles in the calculated in the triangulation[?]. This triangulation methodology creates a set of triangles such that the circumcircle of each triangle contains no other points (See Figure 4). What this means is that if one were to draw a circle around every triangle where each point of the triangle lies on the circumference of that circle, no points will fall inside another triangles circle [3].

The issue with Delaunay triangulation is one of dimensionality. In 2-dimensions, Delaunay triangulation can be done with a worst-case time complexity of just  $O(n \log n)$  which is one of the quickest methods for pseudo-optimal triangulation [3]. However, it becomes increasingly difficult to compute as dimensions increase, making it a less viable solution to all

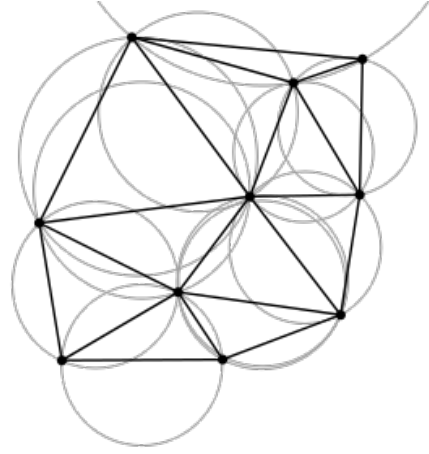


Fig. 4. Circumcircles drawn amongst the Triangles[12]

versions of problem. The majority of algorithms that compute Delaunay triangulation are typically limited to just 2 and 3 dimensions. One way of computing Delaunay triangulation in larger dimensions would be to increase the dimension by 1 (giving the value of the new dimension as  $|p|^2$ ) and then computing the convex hull on those sets of points [?]. However, as the dimensions increase even that becomes increasingly complex to compute. In 4-dimensions the best case algorithm will have a worst-case  $O(n^3)$  run time. In  $d$ -dimensions, it takes  $O(n^{\lceil d/2 \rceil + 1})$  time.

### III. CPSO USING DELAUNAY TRIANGULATION

Building off of the previous work, it is clear the major downside to the use of Delaunay triangulation spawns from its issues in higher dimensions. In the past, such issues were solved by utilizing CPSOs ability to break the problem into a number of smaller sub problems, as seen with PSOs curse of dimensionality [8]. Knowing that it has been known to lead to better results, it would be worth trying Delaunay triangulation in conjunction with CPSO algorithms to see if they can both benefit from each others skills to better solve problems in  $n$ -dimensions. It was stated that the tipping point for this benefit is for problems greater than 3 dimensions. As such, all CPSO variations have been initialized such that no one swarm is handling more than 3 dimensions.

#### A. Experimental Setup

To test this theory, a number of known CPSO algorithms were implemented with the added functionality of the communication via spatially significant neighbors using the Delaunay triangulation approach. The following algorithms were tested for this experiment:

- **PSO**: For comparisons sake, all CPSO variants will be compared to the standard PSO algorithm to ensure either of the changes lead to a meaningful benefit. All the values will be identical to their CPSO counterparts.
- **CPSO-S**: This variant restricts each swarm to just one dimension. It solves all dimensions dependently

of each other. In problems of a single dimension, the distances are easily calculated directly instead of using the Delaunay triangulation algorithm. This is due to Delaunay triangulation not working for problems under 2 dimensions.

- **CPSO-S<sub>k</sub>**: Due to Delaunay triangulation being limited to just 2 and 3 dimensions, a  $k$  needed to be selected in which the swarm is divided into 2 or 3 dimension chunks. For the purposes of this test, swarms were divided as such that each swarm solves 2 dimensions. ( $k = \frac{\text{dimensions}}{2}$ )
- **CPSO-R<sub>k</sub>**: The algorithm was altered to ensure all the random swarms are selected to only be between 1 and 3 dimensions to ensure the spatially significant neighbors algorithm works. The number of swarms is set to half that of the overall dimensions to allow for an even distribution of 1 dimension and 3 dimensional swarms.

### B. CPSO configuration

The CPSO and PSO parameters were selected to match those tested in both the Delaunay triangulation and the CPSO paper in attempt to replicate the results achieved in those experiments. As per the other experiments, the number of iterations were capped at 10,000[5]. To keep the playing field even, the particles allotted to each algorithm will be distributed evenly amongst each swarm. This means that each iteration for the swarm will contain the exact same number of fitness calls, allowing direct comparison of the results. Each experiment was performed 50 times; the reported results are the average of the global bests after all of the 50 runs. The experiments were also repeated for each type of swarm using 15, 20, and 25 particles for 6 dimension problems and 50,75,100 for 20 dimension problem with and without the use of spatially significant neighbors. All variants of CPSO used the following input values:

- **Social Weight** ( $\lambda$ ): 1.49
- **Cognitive Weight** ( $\omega$ ): 1.49
- **Inertial Weight** ( $\alpha$ ): starts at 1 and linearly decreases with each iteration, ultimately hitting 0 on the last iteration
- **Max Velocity** ( $V_{max}$ ): clamped to the domain of the function
- **Max Position** ( $P_{max}$ ): also clamped to the domain of the function
- **Particle Count**: particles are evenly distributed among the swarms in the algorithm. for example, if the particle count is set to 100 and the CPSO has 10 different swarms, each swarm will only have 10 particles to solve its particular sub problem.

### C. Benchmark functions

Again, in an attempt to replicate the results of the previous experiments, for comparisons sake, the same minimization benchmark functions were selected here as were in the two previous papers[3] [5]. The functions which were selected are below.

Function	Domain	Criterion
Rastrigin	[-5.12;5.12]	0.01
Rosenbrock	[-30;30]	100
Griewank	[-600;600]	0.05
Ackley	[-32;32]	0.01

TABLE IV. FUNCTIONS, STOP CRITERIA AND DOMAINS

The Success Rate evaluation indicates the percentage of runs that reached the end criterion and the Iterations to Criterion value returns the average number of iterations it took to reach that stopping point.

Below are the benchmark functions that are being tested as well as a description of the attributes of the function which makes each worth testing in this experiment.

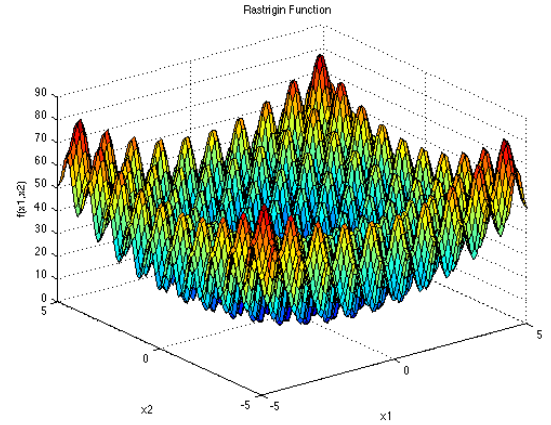


Fig. 5. The Graphical representation of the Rastrigin function in 2 dimensions [13]

#### 1) Rastrigin:

$$\sum_{i=1}^n x_i^2 + 10 - 10\cos(2\pi x_i)$$

Figure 5 gives a visual representation of the fitness values outputted from the Rastrigin function based on a 2 dimensional set of inputs. As can be noted, the function has a large number of local optima throughout the search space intended to trap optimization functions in incorrectly converging in a pseudo-optimal value. The overall minimum value of the function is when all inputs are at zero (the coordinate of the origin), in which the final fitness is also 0. The graph shows the Rastrigin function then continues to increase as inputs venture away from the origin. The further away from the origin, the larger the fitness values from both the peaks and valleys of the subspace pockets. Rastrigin is a good benchmark function as it helps determine how well the optimization algorithm is able to avoid getting stuck at pseudo-optimal solutions.

#### 2) Rosenbrock:

$$\sum_{i=1}^{n-1} 100(x_{i+1} - x_i^2)^2 + (x_i - 1)^2$$



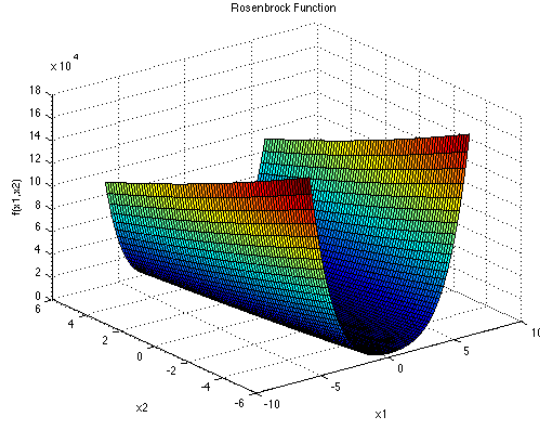


Fig. 6. The Graphical representation of the Rosenbrock function in 2 dimensions [14]

The Rosenbrock function takes on the shape of a valley (as can be seen in Figure 6). This initially seems to be easy to optimize but it is deceptively tricky. Though locating the valley within the function is generally not too difficult for most optimization problems, the majority of the problem comes in finding the optimal value within it. The global minimum is inside a long, narrow, parabolic shaped valley and resides at the coordinate when all inputs are 1. This will return the global minimum value of 0.

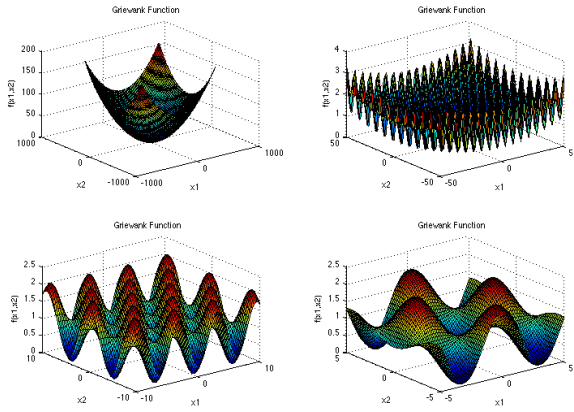


Fig. 7. The Graphical representation of the Griewank function in 2 dimensions [15]

3) *Griewank:*

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

The Griewank function follows a similar pattern as the Rastrigin function above. Figure 7 shows that, similar to Rastrigin, it utilizes pockets of local optima to try to get the algorithm to converge on a pseudo-optima. However, the

difference between the functions can be seen when looking at the scales for the graphs. Where the Rastrigin function features steep/narrow peaks, the Griewank function is much more spread out. The subspaces are larger and require the algorithm to explore more to escape the local optima. Similar to Rastrigin, the optimal solution is when the inputs are all 0 (at the origin) which returns the fitness value of 0.

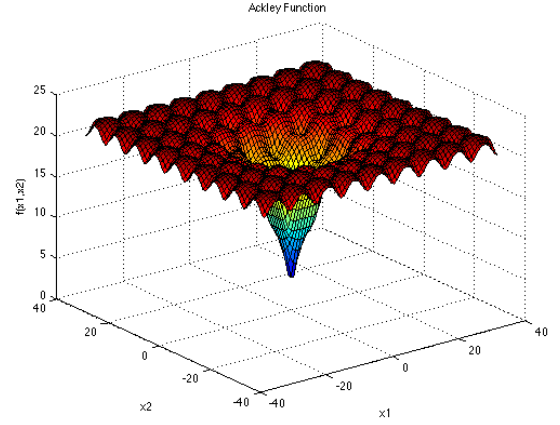


Fig. 8. The Graphical representation of the Ackley function in 2 dimensions [16]

4) *Ackley:*

$$20 + e - 20e^{-0.2\sqrt{\frac{\sum_{i=1}^n x_i^2}{n}}} - e \frac{\sqrt{\sum_{i=1}^n \cos(2\pi x_i)}}{n}$$

The Ackley function has a similar set of local optimal subspaces to both the Griewank and Rastrigin functions. However, Figure 8 shows that the pockets remain flat as the coordinate move away from the origin, making it more difficult to surmise which direction to head. Unlike the other functions, Ackley has a large hole in the center of the graph that leads straight to the global best value. If the optimization function is able to find this hole, it should have no problem converging on the optimal solution, though finding the hole is not as trivial as it may seem. The optimal solution for the Ackley function is when all inputs are at 0, this will return the minimum fitness value of 0.

#### D. Experiment 1: Robustness Test on Lower Dimensions

The mark of a good PSO algorithm is one that can hit or get very close to the optimal result on a consistent basis. As such, it is important that the robustness of the algorithm is tested against other similar algorithms to see how it compares. In this case, robustness refers to the percentage of successful runs and the number of iterations required to reach that success over the course of 50 runs. A successful run is one that reaches the threshold for the optimal solution before the maximum number of iterations is reached. The thresholds vary per function (see III-C) but the max number of iterations is set to a constant 10,000 iterations for all tested algorithms.

To make the iterations equivalent for all the various types of PSO algorithms, the total number of particles per algorithm type is divided evenly amongst all available swarms. This means each iteration evaluates the same number of particles regardless of swarm count. The resulting tables are structured as such that the algorithm being tested is in the first column, followed by the number of particles used in that calculation. Next is the percentage of successful runs with the average iterations for tests using the star topology followed by tests using the spatially meaningful neighbors topology. The best runs are those that have the highest success percentage while also taking the least number of iterations to achieve it. It is to be noted that if the run was not successful, its iteration count does not factor into the average iterations for the test. The spatially meaningful neighbors topology was not tested for PSO since it has already been determined that the method should not be used for swarms larger than 3 dimensions. For this first experiment, tests were run against problems of 6-dimensions to see if the Delaunay triangulation calculations led to noticeable improvements.

Algorithm	s	Standard		Spatially Meaningful Neighbors	
		Succeeded	Iterations	Succeeded	Iterations
PSO	15	54%	1291		
	20	63%	1264		
	25	60%	1185		
CPSO-S	15	100%	837	100%	772
	20	100%	597	100%	624
	25	100%	462	100%	465
CPSO-S <sub>2</sub>	15	90%	1091	92%	1191
	20	98%	1033	100%	1165
	25	98%	969	100%	1145
CPSO-R <sub>3</sub>	15	58%	1173	74%	1295
	20	80%	1110	80%	1312
	25	88%	1032	84%	1314

TABLE V. RASTRIGIN ROBUSTNESS ANALYSIS (N=6)

1) *Results:* Table V details the results of the Rastrigin function against all tested algorithms over the course of 50 runs. Of all 4 algorithms tested, the standard PSO has the lowest success rate as well as the highest average iteration count for the runs that were successful. The CPSO-S algorithm was able to converge on the success threshold 100% of the time regardless of the number of particles used. In this case, the new topology on average led to fewer average iterations, though that is the only case that holds true for in the Rastrigin runs. For the CPSO-S<sub>k</sub> tests, the spatially meaningful neighbors topology consistently led to better results having one more successful run on each test, though requiring a higher average number of iterations to do it. The CPSO-R<sub>k</sub> run shows the first time that the new topology has led to a worse result, however on average, there is still a net benefit to the new topology due to the 16% increase in success percentage for the 15 particle run.

The results of the Rosenbrock tests are represented in Table VI. The table shows that, again, PSO has one of the higher average iteration counts but is much more successful than with the Rastrigin function. It was successful for almost all of their runs which is also true for all of the other algorithms tested. For both the CPSO-S and CPSO-S<sub>k</sub> algorithms, the spatially

Algorithm	s	Standard		Spatially Meaningful Neighbors	
		Succeeded	Iterations	Succeeded	Iterations
PSO	15	98%	346		
	20	100%	308		
	25	100%	350		
CPSO-S	15	100%	287	100%	356
	20	100%	128	100%	116
	25	100%	106	100%	96
CPSO-S <sub>3</sub>	15	100%	284	100%	316
	20	100%	153	100%	156
	25	100%	90	100%	85
CPSO-R <sub>3</sub>	15	94%	301	98%	386
	20	98%	392	100%	415
	25	100%	260	100%	275

TABLE VI. ROSENBROCK ROBUSTNESS ANALYSIS (N=6)

meaningful neighbors topology leads to a higher iteration count for the 15 particle test, however leads to similar or fewer iterations when using a greater number of particles. In this test, the CPSO-R algorithm performs the worst, however is improved upon with the new topology. With the use of spatially meaningful neighbors, CPSO-R<sub>k</sub> is able to match the success rate of the PSO algorithm, however, in all cases it requires more iterations than the standard CPSO-R<sub>k</sub> and on average more iterations than PSO.

Algorithm	s	Standard		Spatially Meaningful Neighbors	
		Succeeded	Iterations	Succeeded	Iterations
PSO	15	40%	1323		
	20	28%	1269		
	25	22%	1222		
CPSO-S	15	98%	980	100%	976
	20	100%	724	96%	759
	25	98%	647	100%	733
CPSO-S <sub>3</sub>	15	82%	1220	98%	1465
	20	84%	1208	100%	1440
	25	94%	1104	100%	1396
CPSO-R <sub>3</sub>	15	62%	1301	82%	1531
	20	66%	1247	86%	1581
	25	80%	1230	98%	1700

TABLE VII. GRIEWANK ROBUSTNESS ANALYSIS (N=6)

The Griewank result in Table VII show the largest boost from the spatially meaningful neighbors topology yet. Once again, the PSO algorithm is the least successful, unable to get above 50% success rate while still requiring over 1200 iterations. The CPSO-S runs show both an increase and a decrease in success rate for different particle tests. On average, however, the success rate is the same for both topologies. The CPSO-S<sub>k</sub> and CPSO-R<sub>k</sub> algorithms see the largest benefit to the new topology here. Each test receives around a 20% boost in success rate at the cost of much higher iteration counts (an increase of around 200 each test).

The Ackley test, however, sees the least change between topologies (as seen in Table VIII). Throughout the table, the algorithms are almost identical. Both success rates and average iteration counts see little to no change with the exception of the CPSO-R<sub>k</sub> algorithm which displays a small decrease in success rate when using the spatially meaningful neighbors topology. In the graph at Figure ?? the 2 versions of the algorithms tend to mirror each other and stay close to their topology counterpart.



Algorithm	s	Standard		Spatially Meaningful Neighbors	
		Succeeded	Iterations	Succeeded	Iterations
PSO	50	100%	1375		
	75	100%	1287		
	100	100%	1262		
CPSO-S	50	94%	1223	94%	1272
	75	100%	1115	100%	1137
	100	100%	1014	100%	1013
CPSO-S <sub>3</sub>	50	100%	1261	100%	1263
	75	100%	1202	100%	1237
	100	100%	1148	100%	1189
CPSO-R <sub>3</sub>	50	96%	1351	94%	1365
	75	100%	1309	100%	1299
	100	100%	1211	100%	1267

TABLE VIII. ACKLEY ROBUSTNESS ANALYSIS (N=6)

2) *Discussion*: The robustness metric shows how successful the swarms are at finding optimal solutions and, in comparison, the effect the spatially meaningful neighborhood Topology has on this success value. In most cases, we can see a positive effect of utilizing spatial information in determining the movement of the particles. The trade off to this is the spatially meaningful neighbor tests tend to take more iterations before they converge on the threshold. This is an expected side effect since it opts for more exploration. This trade off can range from a single extra iteration to hundreds. The benefits can be seen in a number of cases, especially for the Griewank (Table VII) and Rastrigin functions (Table V), there is a large boost of 10%-20% coming from the change in the neighborhood topologies. This shows that the new topology is better able to explore sub-spaces as originally theorised. Since the particles are able to work together to investigate these spaces, they are able to more easily find if that space is worth further investigation. In the table results, there are a number of instances where the new topology leads to around 2% fewer successes (or one extra failure). These discrepancies can often potentially be explained by the random nature of the initial particle generation.

#### E. Experiment 2: Robustness Test on Higher Dimensions

The issue with the Delaunay triangulation method for calculating spatially meaningful neighbors, as originally pitched, was that the algorithm did not scale well to higher dimensions [3]. When the dimensions reached 4 or more, the process to calculate this triangulation becomes significantly more complex and time consuming. This excess calculation negates the benefits that come from the new topology since it no longer becomes a viable option. By limiting the Delaunay triangulation calculations to just 3-dimensions, the number of calculations required to create the triangulation drops to just  $O(n^2)$ , which is significantly more manageable than the dimension dependant exponential problem it otherwise would be. In the following test, the Delaunay triangulation method is used in conjunction with CPSO to attempt to still gain the benefit of the topology while solving problems of dimension 20. The number of particles and swarms were uniformly increased with the dimension to ensure that all CPSO variants have enough particles to adequately search the search space. CPSO-S, for instance, would then need to be divided into 20 swarms so ensuring each swarm had at the very least, 2

particles was key. If the method can scale to larger dimensions, it may be a viable solution for a greater number of problem.

Algorithm	s	Standard		Spatially Meaningful Neighbors	
		Succeeded	Iterations	Succeeded	Iterations
PSO	50	28%	1434		
	75	44%	1351		
	100	62%	1236		
CPSO-S	50	94%	986	94%	1012
	75	100%	800	100%	827
	100	100%	639	100%	624
CPSO-S <sub>3</sub>	50	76%	1258	86%	1321
	75	86%	1158	90%	1287
	100	100%	1047	100%	1238
CPSO-R <sub>3</sub>	50	26%	1458	34%	1324
	75	56%	1244	56%	1548
	100	68%	1485	74%	1152

TABLE IX. RASTRIGIN ROBUSTNESS ANALYSIS (N=20)

1) *Results*: By increasing the number of dimensions, as seen in Table IX, PSO is shown to have a harder time converging onto the threshold without a large number of particles. For all but the 100 particle test the PSO performs worse than its lower dimension counterpart despite the same particles to dimensions ratio. Of the tested algorithms, the standard PSO performed the worst. The CPSO-S algorithm performed very similarly to the smaller dimension counterpart though it did not see any benefit from the use of the new topology. CPSO-S<sub>k</sub> and CPSO-R<sub>k</sub>, however, both saw consistent improvements in success rate when using the spatially meaningful neighbors topology. This led to around a 10% increase on each test and was able to bring the CPSO-S<sub>k</sub> algorithm performing around the same success rate as the lower dimensions. The CPSO-R<sub>k</sub> algorithm also struggled in higher dimensions, however, and without the new topology had similar results to the standard PSO.

Algorithm	s	Standard		Spatially Meaningful Neighbors	
		Succeeded	Iterations	Succeeded	Iterations
PSO	50	100%	366		
	75	100%	300		
	100	100%	250		
CPSO-S	50	100%	653	94%	605
	75	100%	382	100%	393
	100	100%	259	100%	187
CPSO-S <sub>3</sub>	50	100%	622	96%	596
	75	98%	517	100%	545
	100	100%	465	100%	454
CPSO-R <sub>3</sub>	50	100%	625	100%	704
	75	100%	542	100%	708
	100	100%	419	100%	611

TABLE X. ROSENBROCK ROBUSTNESS ANALYSIS (N=20)

Table X details the results from the high-dimension Rosenbrock tests. In this case, the results are very similar to the lower dimension tests. All algorithms are able to consistently hit 100% success rate for both topologies. In this case, the standard PSO algorithm actually performs the best, requiring on average the least number of iterations to achieve the desired success rate. In this case, the new topology had either no effect or in some cases a negative effect. On 2 separate occasions, the new topology lowered the success rate for the test runs, while also taking similar or more iterations to achieve that.

The table for the Griewank results (Table XI) in the higher

Algorithm	s	Standard		Spatially Meaningful Neighbors	
		Succeeded	Iterations	Succeeded	Iterations
PSO	50	92%	1240		
	75	100%	1118		
	100	98%	1064		
CPSO-S	50	88%	1016	89 %	1052
	75	76%	939	76%	962
	100	74 %	705	80%	714
CPSO-S <sub>3</sub>	50	78%	1256	78%	1415
	75	78%	1206	94%	1487
	100	84%	1136	94%	1486
CPSO-R <sub>3</sub>	50	42%	1588	44%	1588
	75	50%	1286	52%	1570
	100	40%	1240	64%	1689

TABLE XI. GRIEWANK ROBUSTNESS ANALYSIS (N=20)

dimension test shows the largest difference compared to the 6-dimension tests. In this case, the standard PSO went from the least successful to the most, being the only algorithm to achieve a 100% success rate. All other algorithms had a drop in success rating though the benefit gained from the new topology remained consistent here. All test either had similar or better results from using the spatially meaningful neighbors topology though that still was not enough to match the results from the previous experiment. It demonstrates very few of the algorithms even getting close to the optimal solution, converging early and having trouble improving.

Algorithm	s	Standard		Spatially Meaningful Neighbors	
		Succeeded	Iterations	Succeeded	Iterations
PSO	50	96%	1507		
	75	100%	1387		
	100	100%	1295		
CPSO-S	50	100%	1254	100%	1262
	75	100%	1168	100%	1147
	100	100%	997	100%	992
CPSO-S <sub>3</sub>	50	98%	1327	100%	1311
	75	100%	1236	100%	1226
	100	100%	1127	100%	1178
CPSO-R <sub>3</sub>	50	94%	1426	100%	1426
	75	100%	1290	100%	1318
	100	100%	1217	100%	1246

TABLE XII. ACKLEY ROBUSTNESS ANALYSIS (N=20)

The final test was to optimize the Ackley function and the results are in Table XII. This test remains pretty consistent to the previous experiment. Similar to the lower dimension test, the majority of the runs we successful. This time, there was a more consistent increase for the spatially meaningful neighbor topology as it was more successful in hitting the threshold. It was also able to do so in very similar iterations as its counterpart. Figure ?? shows just how successful the new topology was in this case. The CPSO-S<sub>k</sub> algorithm without Delaunay triangulation performed significantly worse than version with that calculation. The other algorithms converge quickly to the threshold and are able to consistently be successful.

2) *Discussion:* After promising results in lower dimensions, tests were performed to see how well the algorithm scales to higher dimension problems. The problem set was increased from 6-dimensions to 20-dimensions to see the effect this has on the algorithms robustness. The results show a slight decrease in the benefit from the topology though still leading

to on average better results. Part of the reason for this is the increase in particles, which led to all the variants having an adequate number to optimize the solution. The cases where it was not 100% saw a small increase in effectiveness but not a large enough boost to definitively conclude that it was solely due to the use of spatially meaningful neighbors. The CPSO algorithms also, in general, performed better or the same as the standard PSO algorithm. The only case where this is not true is the Griewank algorithm. This behaviour is expected, however, since it is known the function becomes easier to solve in higher dimensions [17]. This means the dimension separation in CPSO negatively affects the ability to converge on the optimal solution. It must also be noted that the run-time of the Delaunay triangulation version of the algorithms also ran exceedingly slow. This is mostly due to the number of times we are calling that Delaunay triangulation (DT) algorithm. Since the swarm is divided in  $n/2$  or  $n/3$  sub swarms, each of which needs to calculate the DT with each iteration. This leads to an extra  $O(n * n^2)$  algorithm being run every iteration of the CPSO. Though it is not as time consuming as  $O(n^{\lceil d/2 \rceil + 1})$ , especially on significantly large dimensions, it is also not an insignificant use of resources.

This paper examined the work brought fourth by Lane, et al in search of improving the PSO algorithm by utilizing positional data to intelligently decide the movement of the particle, [3] and attempts to expand the work's reach by expanding it to new problems. In removing the initial dimensional limitation through the incorporation of cooperative swarms, this paper hopes this technique could be brought to larger more complicated problems.

The results of the study proved the benefits of this technique still hold up under the use of cooperative particle swarm optimization as the particles we able to better work together to exploit sub-spaces. By separating higher dimension problem into multiple smaller problems, the Delaunay triangulation calculations required for defining the distance between particles becomes significantly more manageable and actually feasible in appropriately large problems. However, although this change significantly decreases the overall number of calculations needed, there is still a large overhead required for this calculation. Since it needs to be calculated for every swarm on every iteration, this can dramatically slow down the run time of the application. Additionally, the benefit gained from this technique is highly dependent on the problem at hand. While functions that required the optimization algorithm to navigate difficult sub-spaces experienced noticeable boosts in performance, other problems experienced little to no gain, making the use of the topology (and large overhead required) hard to justify. Ultimately, if the success of the algorithm is of higher priority than the time it takes to complete, the spatially meaningful neighbor topology is certainly worth a look.

#### IV. FUTURE WORK

This paper only touches on the potential of new neighborhood topologies that can come from this general idea. Having particles be more aware of their position in relation to the rest of the swarm is ripe for improvements. Testing out a

number of different knowledge transfer techniques along with the dynamic nature of this neighborhood topology can lead to a more intelligent particle. Potentially having particles more directly communicate with the swarm as a whole would be invaluable.

Further work could also be invested in discovering a quicker, less time consuming way of determining the neighbors. The Delaunay triangulation algorithm adds a large amount of overhead, especially for a calculation that is being repeated as often as it is. Even in lower dimension problems, the Delaunay triangulation algorithm is the bottleneck calculation and significantly slows down execution of the PSO.

Finally, work could also be invested in adding this technique to different variants of the CPSO algorithm. For the purposes of the preliminary tests of this paper a number of the CPSO variants were not tested here. A potential next step could be to try it to more and see if there is a difference in the effect it has on other variants, such as the mentioned CPSO- $H_k$  algorithm.

## APPENDIX A

### PROOF OF THE FIRST ZONKLAR EQUATION

Some text for the appendix.

## REFERENCES

- [1] R. Poli, J. Kennedy, and T. Blackwell, "Particle swarm optimization," *Swarm Intelligence*, vol. 1, no. 1, pp. 33–57, 2007. [Online]. Available: <http://dx.doi.org/10.1007/s11721-007-0002-0>
- [2] A. J. R. Medina, G. T. Pulido, and J. G. Ramirez-Torres, "A comparative study of neighborhood topologies for particle swarm optimizers," in *IJCCI*, A. D. Correia, A. C. Rosa, and K. Madani, Eds. INSTICC Press, 2009, pp. 152–159. [Online]. Available: <http://dblp.uni-trier.de/db/conf/ijcci/ijcci2009.html#MedinaPR09>
- [3] J. Lane, A. Engelbrecht, and J. Gain, "Particle swarm optimization with spatially meaningful neighbours," in *Proceedings of the 2008 IEEE Swarm Intelligence Symposium. Piscataway, NJ: IEEE*. Press, 2008, pp. 1–8.
- [4] D. T. Lee and B. J. Schachter, "Two algorithms for constructing a Delaunay triangulation," *International Journal of Computer & Information Sciences*, vol. 9, no. 3, pp. 219–242, 1980. [Online]. Available: <http://dx.doi.org/10.1007/BF00977785>
- [5] F. V. den Bergh and A. Engelbrecht, "A cooperative approach to particle swarm optimization," *IEEE Transactions on Evolutionary Computation*, p. 225239, June 2004.
- [6] D. Poole, A. Mackworth, and R. Goebel, *Computational Intelligence: A Logical Approach*. Oxford, UK: Oxford University Press, 1997.
- [7] M. E. H. Pedersen, "Good parameters for particle swarm optimization," *Hvass Lab., Copenhagen, Denmark, Tech. Rep. HLI001*, 2010.
- [8] J. Maltese, B. Ombuki-Berman, and A. Engelbrecht, "Co-operative vector-evaluated particle swarm optimization for multi-objective optimization," in *Computational Intelligence, 2015 IEEE Symposium Series on*. IEEE, 2015, pp. 1294–1301.
- [9] A. E. Muñoz Zavala, *A Comparison Study of PSO Neighborhoods*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2013, pp. 251–265. [Online]. Available: [http://dx.doi.org/10.1007/978-3-642-31519-0\\_16](http://dx.doi.org/10.1007/978-3-642-31519-0_16)
- [10] W. Zou, Y. Zhu, H. Chen, and T. Ku, "Clustering approach based on Von Neumann topology artificial bee colony algorithm," in *2011 international conference on data mining (DMIN11)*, 2011.
- [11] M. d. Berg, O. Cheong, M. v. Kreveld, and M. Overmars, *Computational Geometry: Algorithms and Applications*, 3rd ed. Santa Clara, CA, USA: Springer-Verlag TELOS, 2008.

Gjacquetot, "Delaunay circumcircles vectorial," [https://upload.wikimedia.org/wikipedia/commons/d/db/Delaunay\\_circumcircles\\_vectorial\\_May-2016](https://upload.wikimedia.org/wikipedia/commons/d/db/Delaunay_circumcircles_vectorial_May-2016).

S. F. U. Sonja Surjanovic Darek Bingham, "Rastrigin function," <http://www.sfu.ca/ssurjano/rastr.png> [Online; accessed 14-September-2016].

—, "Rosenbrock function," <http://www.sfu.ca/ssurjano/rosen.png> [Online; accessed 14-September-2016].

—, "Griewank function," <http://www.sfu.ca/ssurjano/griewank.png> [Online; accessed 14-September-2016].

—, "Ackley function," <http://www.sfu.ca/ssurjano/ackley.png> [Online; accessed 14-September-2016].

M. Locatelli, "A note on the griewank test function," *Journal of Global Optimization*, vol. 25, no. 2, pp. 169–174, 2003. [Online]. Available: <http://dx.doi.org/10.1023/A:1021956306041>

**John Doe** Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam lobortis facilisis sem. Nullam nec mi et neque pharetra sollicitudin. Praesent imperdiet mi nec ante. Donec ullamcorper, felis non sodales commodo, lectus velit ultrices augue, a dignissim nibh lectus placerat pede. Vivamus nunc nunc, molestie ut, ultricies vel, semper in, velit. Ut portitor. Praesent in sapien. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Duis fringilla tristique neque. Sed interdum libero ut metus. Pellentesque placerat. Nam rutrum augue a leo. Morbi sed elit sit amet ante lobortis sollicitudin. Praesent blandit blandit mauris. Praesent lectus tellus, aliquet aliquam, luctus a, egestas a, turpis. Mauris lacinia lorem sit amet ipsum. Nunc quis urna dictum turpis accumsan semper.