Population Diversity Based Study on Search Information Propagation in Particle Swarm Optimization

Shi Cheng*†, Yuhui Shi‡, Quande Qin‡
*Department of Electrical Engineering and Electronics, University of Liverpool, Liverpool, UK
† Department of Electrical & Electronic Engineering, Xi’an Jiaotong-Liverpool University, Suzhou, China
‡ College of Management, Shenzhen University, Shenzhen, China

Abstract—Premature convergence happens in Particle Swarm Optimization (PSO) partially due to improper search information propagation. Fast propagation of search information will lead particles get clustered together quickly. Determining a proper search information propagation mechanism is important in optimization algorithms to balance between exploration and exploitation. In this paper, we attempt to figure out the relationship between search information propagation and the population diversity change. Firstly, we analyze the different characteristics of search information propagation in PSO with four kinds of topologies: star, ring, four clusters, and Von Neumann. Secondly, population diversities of PSO, which include position diversity, velocity diversity, and cognitive diversity, are utilized to monitor particles’ search during optimization process. Position diversity, velocity diversity, and cognitive diversity, represent distributions of current solutions, particles’ “moving potential”, and particles’ “moving target”, respectively. From the observation of population diversities, the effect of search information propagation on PSO’s optimization performance is discussed at last.

Index Terms—Particle Swarm Optimization, Population Diversity, Search Information Propagation, Exploration, Exploitation.

I. INTRODUCTION

Particle Swarm Optimization (PSO) was introduced by Eberhart and Kennedy in 1995 [1], [2]. It is a population-based stochastic algorithm modeled on social behaviors observed in flocking birds. A particle flies through the search space with a velocity that is dynamically adjusted according to its own and its companion’s historical behaviors. Each particle’s position represents a solution to the problem. Particles tend to fly toward better and better search areas over the course of the search process [3].

Optimization, in general, is concerned with finding the “best available” solution(s) for a given problem, and the problem may have several or numerous optimum solutions, of which many are local optimal solutions. Evolutionary optimization algorithms are generally difficult to find the global optimum solutions for multimodal problems due to the possible occurrence of the premature convergence.

Particles fly in the search space. If particles can easily get clustered together in a short time, particles will lose their “search potential.” Premature population convergence around a local optimum is a common problem for population-based algorithms. It is a result of individuals hastily congregating within a small region of the search space. An algorithm’s search ability of exploration is decreased when premature convergence occurs, particles will have a low possibility to explore new search areas. Normally, diversity, which is lost due to particles getting clustered together, is not easy to be recovered. An algorithm may lose its search efficacy due to premature. As a population becomes more converged, an algorithm will spend most of iterations to search in a small region.

Premature convergence and low search efficacy happen partially due to the reason that population-based algorithm has an improper search information propagation strategy. For example, to solve a difficult problem, star topology in classical PSO or fully informed PSO may cause a fast propagation of search information. On contrast, to solve a simple problem, ring topology in classical PSO has a slow propagation in search information, which may lead to a slow convergence. The algorithm may need many iterations to find the “good enough” solution.

Although many methods were reported to be designed to avoid premature convergence, these methods did not incorporate an effective way to measure the degree of premature convergence, in other words, the measurement of particles' exploration/exploitation is still needed to be investigated. Shi and Eberhart gave several definitions on diversity measurement based on particles’ positions, velocity, and historical best positions [4], [5]. Through diversity measurements, useful exploration and/or exploitation search information can be obtained.

This paper is organized as follows. Section II reviews the basic PSO algorithm and fully informed PSO algorithm. Section III introduces the three kinds of dimension-wise population diversities, which is based on $L_1$ norm. Section IV describes the search information propagation during the search process. Experiments are conducted in Section V followed by analysis and discussion on the population diversity changing curves of...
PSOs in Section VI. Finally, Section VII concludes with some remarks and future research directions.

II. PARTICLE SWARM OPTIMIZATION

For the purpose of generality and clarity, \( m \) represents the number of particles and \( n \) the number of dimensions. Each particle is represented as \( x_{ij} \), \( i \) represents the \( i \)th particle, \( i = 1, \cdots, m \), and \( j \) is the \( j \)th dimension, \( j = 1, \cdots, n \).

A. Classical Particle Swarm

The basic equations of the original PSO algorithm are as follow [6], [7]:

\[
\begin{align*}
v_i &\leftarrow w v_i + c_1 \text{rand}() (p_i - x_i) + c_2 \text{Rand}() (p_g - x_i) \\
x_i &\leftarrow x_i + v_i
\end{align*}
\]

where \( w \) denotes the inertia weight and usually is less than 1, \( c_1 \) and \( c_2 \) are two positive acceleration constants, \( \text{rand}() \) and \( \text{Rand}() \) are two random functions to generate uniformly distributed random numbers in the range \([0, 1]\), \( x_i \) represents the \( i \)th particle’s position, \( v_i \) represents the \( i \)th particle’s velocity, \( p_i \) is termed as personal best, which refers to the best position found by the \( i \)th particle, and \( p_g \) is termed as local best, which refers to the position found by the members in the \( i \)th particle’s neighborhood that has the best fitness value so far.

B. Fully Informed Particle Swarm

Fully informed PSO (FIPS) does not share the concept of “global/local best”. A particle in FIPS does not follow the leader in its neighborhood, but follow all other particles in its neighborhood. The basic equations of the FIPS algorithm are as follow [8], [9]:

\[
\begin{align*}
v_i &\leftarrow \gamma \left( v_i + \frac{1}{N_i} \sum_{k=1}^{N_i} U(0, \varphi) (p_{b(k)} - x_i) \right) \\
x_i &\leftarrow x_i + v_i
\end{align*}
\]

where \( \gamma \) denotes the acceleration coefficient, \( U(0, \varphi) \) is a random function to generate random numbers in the range \([0, \varphi]\), \( N_i \) represents the neighborhood size of the \( i \)th particle, and \( p_{b(k)} \) represents the \( k \)th particle’s personal best position.

III. POPULATION DIVERSITY

The most important factor affecting an optimization algorithm’s performance is its ability of “exploration” and “exploitation.” Exploration means the ability of a search algorithm to explore different areas of the search space in order to have high probability to find good promising solutions. Exploitation, on the other hand, means the ability to concentrate the search around a promising region in order to refine a candidate solution. A good optimization algorithm should optimally balance the two conflicted objectives.

Population diversity of PSO is useful for measuring and dynamically adjusting algorithm’s ability of exploration or exploitation accordingly. Shi and Eberhart gave three definitions on population diversity, which are position diversity, velocity diversity, and cognitive diversity [4], [5]. Position, velocity, and cognitive diversity are used to measure the distribution of particles’ current positions, current velocities, and \( p\text{bests} \) (the best position found so far for each particles), respectively. Cheng and Shi introduced the modified definitions of the three diversity measures based on \( L_1 \) norm [10]–[13].

From diversity measurements, the useful information can be obtained. The detailed definitions of PSO population diversities are as follow [10]–[13]:

A. Position Diversity

Position diversity measures distribution of particles’ current positions, therefore, can reflect particles’ dynamics. Position diversity gives the current position distribution information of particles. Definition of position diversity, which is based on the \( L_1 \) norm, is as follows

\[
x = \frac{1}{m} \sum_{i=1}^{m} x_{ij} \quad D^p = \frac{1}{m} \sum_{i=1}^{m} |x_{ij} - \bar{x}_j| \quad D^p = \sum_{j=1}^{n} w_j D^p_j
\]

where \( \bar{x} = [\bar{x}_1, \cdots, \bar{x}_j, \cdots, \bar{x}_n] \), \( \bar{x} \) represents the mean of particles’ current positions on each dimension. \( D^p = [D^p_1, \cdots, D^p_j, \cdots, D^p_n] \), which measures particles’ position diversity based on \( L_1 \) norm for each dimension. \( D^p \) measures the whole swarm’s population diversity.

B. Velocity Diversity

Velocity diversity, which represents diversity of particles’ “moving potential”, measures the distribution of particles’ current velocities. In other words, velocity diversity measures the “activity” information of particles. Based on the measurement of velocity diversity, particle’s tendency of expansion or convergence could be revealed. The velocity diversity based on \( L_1 \) norm is defined as follows

\[
\bar{v} = \frac{1}{m} \sum_{i=1}^{m} v_{ij} \quad D^v = \frac{1}{m} \sum_{i=1}^{m} |v_{ij} - \bar{v}_j| \quad D^v = \sum_{j=1}^{n} w_j D^v_j
\]

where \( \bar{v} = [\bar{v}_1, \cdots, \bar{v}_j, \cdots, \bar{v}_n] \), \( \bar{v} \) represents the mean of particles’ current velocities on each dimension; and \( D^v = [D^v_1, \cdots, D^v_j, \cdots, D^v_n] \), \( D^v \) measures velocity diversity of all particles on each dimension. \( D^v \) represents the whole swarm’s velocity diversity.

C. Cognitive Diversity

Cognitive diversity, which represents distribution of particles’ “moving target”, measures the distribution of historical best positions for all particles. The measurement definition of cognitive diversity is the same as that of the position diversity except that it utilizes each particle’s current personal best position instead of current position. The definition of PSO cognitive diversity is as follows

\[
\bar{p} = \frac{1}{m} \sum_{i=1}^{m} p_{ij} \quad D^c = \frac{1}{m} \sum_{i=1}^{m} |p_{ij} - \bar{p}_j| \quad D^c = \sum_{j=1}^{n} w_j D^c_j
\]

where \( \bar{p} = [\bar{p}_1, \cdots, \bar{p}_j, \cdots, \bar{p}_n] \) and \( \bar{p} \) represents the average of all particles’ personal best position in history (\( p\text{best} \) on
each dimension; \(D^C = [D^p_1, \ldots, D^p_n, \ldots, D^p_n]\), which
represents the particles’ cognitive diversity for each dimension
based on \(L_1\) norm. \(D^C\) measures the whole swarm’s cognitive
diversity.

Without loss of generality, every dimension is considered
equally in this paper. Setting all weights \(w_j = \frac{1}{n}\), then the
position diversity, velocity diversity, and cognitive diversity of
the whole swarm can be rewritten as:

\[
D^p = \sum_{j=1}^{n} \frac{1}{n} D^p_j = \frac{1}{n} \sum_{j=1}^{n} D^p_j
\]

\[
D^v = \sum_{j=1}^{n} \frac{1}{n} D^v_j = \frac{1}{n} \sum_{j=1}^{n} D^v_j
\]

\[
D^c = \sum_{j=1}^{n} \frac{1}{n} D^c_j = \frac{1}{n} \sum_{j=1}^{n} D^c_j
\]

### IV. Search Information Propagation

A particle updates its position in the search space at each
iteration. The velocity update equation consists of three parts,
which are previous velocity, cognitive part, and social part.
The cognitive part means that a particle learns from its own
searching experience, and correspondingly, the social part
means that a particle can learn from other particles, or learn
from the best in its neighbors in particular. Topology defines
the neighborhood of a particle.

Particle swarm optimization algorithm has different kinds
of topology structures, e.g., star, ring, four clusters, or Von
Neumann structure. A particle in a PSO with a different
structure has different number of particles in its neighborhood
with a different scope. Learning from a different neighbor means
that a particle follows different neighborhood (or local) best,
in other words, topology structure determines the connections
among particles, and the strategy of search information propa-
gation. Although it does not relate to the particle’s cognitive
part directly, topology can affect the algorithm’s convergence
speed and the ability of avoiding premature convergence, i.e.,
the PSO algorithm’s ability of exploration and exploitation.

#### A. Topology Structure

In this paper, four most commonly used topology structures are
considered [14], [15]. They are star, ring, four clusters,
and Von Neumann structure, which are shown in Figure 1.

- **The star topology** is shown in Fig. 1 (a). This topology
  is frequently termed as global or all topology. Because all
  particles or nodes are connected, search information is
  shared in a global scope.

- **The ring topology** is shown in Fig. 1 (b). A particle is
  connected with two neighbors in this topology.

- **The four clusters topology** is shown in Fig. 1 (c). As
  the name indicated, the whole swarm are divided into four
  subgroups. Each subgroup, which is a small star topology,
  has three link particles links to other three groups.

- **The Von Neumann topology** is shown in Fig. 1 (d). This
  topology is also named as Square or NWES neighborhood
  (for each particle has four neighbors in the North, East,
  West, and South).

Table 1 gives different properties of topologies, where
“Neighbor” indicates the number of neighbors that a particle
has, “Diameter” indicates the diameter of a topology, which
is the largest step that search information propagation from
a particle to another, and “Average distance” indicates the
average step of search information propagation in the swarm.
The “Average distance” of a topology structure will always
less than or equal to the “Diameter.”

![Table 1](image-url)

<table>
<thead>
<tr>
<th>Topology</th>
<th>Neighbors</th>
<th>Diameter</th>
<th>Average Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Star</td>
<td>(m - 1)</td>
<td>(\frac{m}{2}) or (\frac{m-1}{2})</td>
<td>(\frac{m^2}{4(m-1)}) or (\frac{m+1}{4})</td>
</tr>
<tr>
<td>Ring</td>
<td>2</td>
<td>(\frac{m}{2}) + (\frac{m-1}{2})</td>
<td>(\frac{7m^2-11d}{2(m-1)}), or (\frac{7d^2-9d-4}{2(m-1)})</td>
</tr>
<tr>
<td>Four Clusters</td>
<td>(\frac{m}{4} - 1) ((\frac{m}{4}))</td>
<td>3</td>
<td>(\frac{5m-14}{2(m-1)})</td>
</tr>
<tr>
<td>Von Neumann</td>
<td>4</td>
<td>2 + ([m/8])</td>
<td>(\frac{m^2}{4(m-1)})</td>
</tr>
</tbody>
</table>

Star structure has the most neighbors and the smallest
diameter. It only needs one iteration to propagate “good”
search information over whole swarm.

Ring structure has the longest diameter among the four
topologies. The calculation of diameter and average distance
are different when the population size is even or odd. When
population size \(m\) is an even number, the search information
needs \(\frac{m}{2}\) to propagate over whole swarm, therefore, the
diameter is \(\frac{m-1}{2}\). A particle needs only one step to two nearby
neighbors, and another step to another two particles, and so
on. The average distance is calculated as follows:

\[
d_{ring} = \frac{2 \times (1 + 2 + \cdots + \frac{m-2}{2}) + \frac{m}{2}}{m-1} = \frac{m^2}{4(m-1)}
\]

On the contrary, when population size \(m\) is an odd number,
the search information needs \(\frac{m-1}{2}\) to propagate over whole
swarm. The average distance is calculated as follows:

\[
d_{ring} = \frac{2 \times (1 + 2 + \cdots + \frac{m-1}{2})}{m-1} = \frac{m+1}{4}
\]

The population size in four clusters structure should be a
multiple of 4, and large than 16. A particle in the cluster has
\(n/4 - 1\) neighbors, and a link particle has \(n/4\) neighbors.
A swarm with four clusters topology has 12 link particles, and
\(m - 12\) inner particles. The average distance is calculated as
follows:

\[
d_{four} = \frac{(m-12)[(m/4 - 1) + 3 \times 2 + (3m/4 - 3) \times 3]}{m(m-1)} + \frac{12(m/4 + (2 + m/4 - 1) \times 2 + (m/2 - 2) \times 3)}{m(m-1)} = \frac{5m-14}{2(m-1)}
\]

The population size in Von Neumann structure also should
be a multiple of 4, and large than 16, normally. Diameter
d equals \(2 + \lfloor m/8\rfloor\) in this structure. The calculation of
average distance is different when the equation \(\mod (n,8)\)
Fig. 1. Topologies used in the paper are presented in the following order: (a) Star topology, where all particles or nodes share the search information in the whole swarm; (b) Ring topology, where every particle is connected to two neighbors; (c) Four clusters topology, where four fully connected subgroups are inter-connected among themselves by linking particles; (d) Von Neumann topology, which is a lattice and where every particle has four neighbors that are wrapped on four sides. Each swarm has 16 particles.

TABLE II

THE BENCHMARK FUNCTIONS USED IN OUR EXPERIMENTAL STUDY, WHERE $n$ IS THE DIMENSION OF EACH FUNCTION, $\mathbf{x} = (\mathbf{x} - \mathbf{a})$.

<table>
<thead>
<tr>
<th>Function</th>
<th>Test Function</th>
<th>$n$</th>
<th>$S$</th>
<th>$f_{\text{min}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Parabolic</td>
<td>$f_0(\mathbf{x}) = \sum_{i=1}^{n} z_i^2 + \text{bias}_0$</td>
<td>100</td>
<td>$[-100, 100]^n$</td>
<td>-450.0</td>
</tr>
<tr>
<td>Schwefel’s P2.22</td>
<td>$f_1(\mathbf{x}) = \sum_{i=1}^{n}</td>
<td>z_i</td>
<td>+ \prod_{i=1}^{n}</td>
<td>z_i</td>
</tr>
<tr>
<td>Schwefel’s P1.2</td>
<td>$f_2(\mathbf{x}) = \sum_{i=1}^{n} (\sum_{k=1}^{i} z_k)^2 + \text{bias}_2$</td>
<td>100</td>
<td>$[-100, 100]^n$</td>
<td>450.0</td>
</tr>
<tr>
<td>Step</td>
<td>$f_3(\mathbf{x}) = \sum_{i=1}^{n} ([z_i + 0.5])^2 + \text{bias}_3$</td>
<td>100</td>
<td>$[-100, 100]^n$</td>
<td>330.0</td>
</tr>
<tr>
<td>Quartic Noise</td>
<td>$f_4(\mathbf{x}) = \sum_{i=1}^{n} iz_i^2 + \text{random}(0, 1) + \text{bias}_4$</td>
<td>100</td>
<td>$[-1.28, 1.28]^n$</td>
<td>-450.0</td>
</tr>
<tr>
<td>Rosenbrock</td>
<td>$f_5(\mathbf{x}) = \sum_{i=1}^{n-1} [100(z_{i+1} - z_i^2)^2 + (z_i - 1)^2] + \text{bias}_5$</td>
<td>100</td>
<td>$[-10, 10]^n$</td>
<td>-330.0</td>
</tr>
<tr>
<td>Rastrigin</td>
<td>$f_6(\mathbf{x}) = \sum_{i=1}^{n} z_i^2 - 10 \cos(2\pi z_i) + 10 + \text{bias}_6$</td>
<td>100</td>
<td>$[-5.12, 5.12]^n$</td>
<td>450.0</td>
</tr>
<tr>
<td>Ackley</td>
<td>$f_7(\mathbf{x}) = -20 \exp \left( -0.2 \sqrt{\frac{1}{n} \sum_{i=1}^{n} z_i^2} \right)$</td>
<td>100</td>
<td>$[-32, 32]^n$</td>
<td>180.0</td>
</tr>
<tr>
<td>Griewank</td>
<td>$f_8(\mathbf{x}) = \frac{1}{10000} \sum_{i=1}^{n} z_i^2 - \prod_{i=1}^{n} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1 + \text{bias}_8$</td>
<td>100</td>
<td>$[-600, 600]^n$</td>
<td>120.0</td>
</tr>
<tr>
<td>Generalized</td>
<td>$f_9(\mathbf{x}) = \frac{x}{n} (10 \sin^2(\pi y_1) + \sum_{i=1}^{n} (y_i - 1)^2)$</td>
<td>100</td>
<td>$[-50, 50]^n$</td>
<td>330.0</td>
</tr>
<tr>
<td>Penalized</td>
<td>$\sum_{i=1}^{n} u(z_i, 10, 100, 4) + \text{bias}_9$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$y_i = 1 + \frac{1}{2} (z_i + 1)$</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>$u(z_i, a, k, m) = \begin{cases} k(z_i - a)^m &amp; z_i &gt; a, \ 0 &amp; -a &lt; z_i &lt; a, \ k(-z_i - a)^m &amp; z_i &lt; -a \end{cases}$</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

has different result. When $\mod(m, 8) = 0$, the average distance is

$$\text{dist}_{\text{von}} = \frac{4 + 7 \times (2 + \cdots + \lfloor m/8 \rfloor)}{m - 1} + \frac{4 \times (1 \lfloor m/8 \rfloor) + 1 \times (2 \lfloor m/8 \rfloor)}{m - 1} + \frac{7 \times (2 \cdots + \lfloor m/8 \rfloor) + 5 \times (2 + \lfloor m/8 \rfloor)}{m - 1} = \frac{7d^2 - 11d}{2(m - 1)}$$

When $\mod(m, 8) = 4$, the average distance is

$$\text{dist}_{\text{von}} = \frac{4 + 7 \times (2 + \cdots + \lfloor m/8 \rfloor)}{m - 1} + \frac{6 \times (1 \lfloor m/8 \rfloor) + 2 \times (2 + \lfloor m/8 \rfloor)}{m - 1} + \frac{7 \times (2 \cdots + \lfloor m/8 \rfloor) + 8 \times (2 + \lfloor m/8 \rfloor) - 2}{m - 1} = \frac{7d^2 - 5d - 4}{2(m - 1)}$$
Diameter and average distance is a fixed number in star topology. In four cluster topology, diameter is 3, the larger the population size $m$ is, the closer to 2.5 the average distance is. In ring and Von Neumann topology, the value of diameter and average distance increases with the population size, and they increase faster in ring topology than in Von Neumann topology.

V. EXPERIMENTAL STUDY

A. Benchmark Test Functions

The experiments have been conducted to test the benchmark functions listed in Table II. Without loss of generality, five standard unimodal and five multimodal test functions are selected [16], [17].

All functions are run 50 times to ensure a reasonable statistical result necessary to compare different approaches. Every tested function’s optimal point in solution space $S$ is shifted to a randomly generated point with different value in each dimension, and $S \subseteq \mathbb{R}^n$, $\mathbb{R}^n$ is a $n$-dimensional Euclidean space.

B. Parameter Setting

Clerc provided a reference setting of parameters in [18]. The setting of population size is calculated according to formula: $m = \lceil 10 + 2\sqrt{n} \rceil$, where $n$ is the dimension of problems. Due to reason that four cluster and Von Neumann topologies require that the population size should be a multiple of 4. In all experiments, problems have 100 dimensions, therefore, each PSO has $m = \lceil 10 + 2\sqrt{100} \rceil + 2 = 32$ particles. Other parameters are set as follow: $w = \frac{1}{2\ln(2)} \approx 0.721$, and $c_1 = c_2 = 0.5 + \ln(2) \approx 1.193$. Correspondingly, in fully informed PSO, $\chi$ is set as 0.72984, and $\varphi$ is 4.1 [9]. Each algorithm runs 50 times.

According to Table I, the parameters and criteria for the test functions are listed in Table III.

<table>
<thead>
<tr>
<th>Topology</th>
<th>Neighbors</th>
<th>Diameter</th>
<th>Average Distance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Star</td>
<td>31</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>Ring</td>
<td>2</td>
<td>16</td>
<td>8.25806</td>
</tr>
<tr>
<td>Four Clusters</td>
<td>7 (8)</td>
<td>3</td>
<td>2.35483</td>
</tr>
<tr>
<td>Von Neumann</td>
<td>4</td>
<td>6</td>
<td>3.0</td>
</tr>
</tbody>
</table>

C. Experimental Results

Table IV gives experimental results of classical PSO and fully informed PSO with star, ring, four clusters, and Von Neumann structure to solve benchmark functions. Mostly, fully informed PSO with ring structure has the best performance than other variants of PSO in this experiment.

VI. POPULATION DIVERSITY ANALYSIS AND DISCUSSION

Without loss of generality and for the purpose of simplicity and clarity, the results for one function from five unimodal benchmark functions and one function from five multimodal functions will be displayed because the results for the others in this experiment will be similar.

There are several definitions on the measurement of population diversities [4], [5], [11]. The dimension-wise population diversities based on the $L_1$ norm are utilized in this paper.

Figure 2 display the population diversity changing curves of classical PSO and fully informed PSO with four kinds of topologies solving unimodal function $f_0$ and multimodal function $f_5$, respectively. In Figure 2, (a), (e) is for classical PSO and fully informed PSO with star structure, (b), (f) is for classical PSO and fully informed PSO with ring structure, (c), (g) is for classical PSO and fully informed PSO with four clusters structure, and (d), (h) is for classical PSO and fully informed PSO with Von Neumann structure, respectively.

From the figures, we can see that classical PSO has a faster decrease of population diversity than fully informed PSO in general. For classical PSO, particles following “leader” in the neighborhood, this will cause a faster search information propagation, then the swarm get clustered in a small region quickly.

Fully informed PSO with a star structure can be seen as extremely examples. For fully informed PSO with a star structure, velocity diversity get clustered to a tiny value rapidly, position diversity and cognitive diversity kept during the whole search process. This is because too much search information propagation in the swarm, all particles quickly move to a small region, premature convergence happens in this case.

Figure 3 and 4 display the population diversity changing curves of classical PSO and fully informed PSO solving unimodal function $f_0$ and multimodal function $f_5$, respectively. In both figures, (a) is for position diversity, (b) is for velocity diversity, (c) is for cognitive diversity, respectively.

Fully informed PSO with star topology uses whole swarm’s search information to update each particle’s position. The particles get clustered together after few iterations, the velocity diversity falls to a tiny value. Velocity diversity represents the distribution of particles’ “moving potential”, from this case, we can concluded that we need to keep this search potential during the search process.

In general, classical PSO with ring structure or fully informed PSO with ring structure performs better than others. Cognitive diversity represents the distribution of particles’ “moving target.” Position diversity represents current solution’s distribution. Position diversity should tend to follow the cognitive diversity or vice versa. Dynamical changing position diversity to adapt cognitive diversity, or on the contrary, changing cognitive diversity to adapt position diversity may improve algorithm’s performance.

VII. CONCLUSIONS

An algorithm’s ability of exploration and exploitation is important in the optimization process. With good exploration
Fig. 2. Population diversity changing curves on PSO solving functions. Unimodal function $f_0$: (a) star, (b) ring, (c) four clusters, (d) von Neumann; multimodal function $f_5$: (e) star, (f) ring, (g) four clusters, (h) von Neumann.

Fig. 3. Comparison of Population diversities for PSO solving unimodal function $f_0$: (a) position, (b) velocity, (c) cognitive.
TABLE IV
RESULTS OF THE CLASSICAL PSO AND FULLY INFORMED PSO WITH DIFFERENT TOPOLOGIES. ALL ALGORITHMS ARE RUN 50 TIMES, WHERE "BEST", "MEDIAN", "MEAN", AND $\sigma$ INDICATE THE BEST, MIDDLE, AVERAGE, AND STANDARD DEVIATION OF THE BEST FITNESS VALUES OVER 50 RUNS, RESPECTIVELY. THE MAXIMUM ITERATION NUMBER IS 4 000.

<table>
<thead>
<tr>
<th>Result</th>
<th>Classical PSO</th>
<th></th>
<th></th>
<th>Fully informed PSO</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Best</td>
<td>Median</td>
<td>Mean</td>
<td>$\sigma$</td>
<td>Best</td>
<td>Median</td>
</tr>
<tr>
<td>f_0</td>
<td>-449.9449</td>
<td>-439.7927</td>
<td>-255.2338</td>
<td>560.8554</td>
<td>339613.56</td>
<td>447076.36</td>
</tr>
<tr>
<td></td>
<td>-449.9999</td>
<td>-449.9999</td>
<td>-449.9999</td>
<td>4.20E-07</td>
<td>-450</td>
<td>-450</td>
</tr>
<tr>
<td></td>
<td>-449.9999</td>
<td>-449.9999</td>
<td>-449.9999</td>
<td>7.99E-10</td>
<td>-449.9999</td>
<td>-449.9999</td>
</tr>
<tr>
<td>f_1</td>
<td>-328.4412</td>
<td>-291.4290</td>
<td>-289.4810</td>
<td>21.8558</td>
<td>7.96E+36</td>
<td>3.26E+43</td>
</tr>
<tr>
<td></td>
<td>-329.9999</td>
<td>-329.9992</td>
<td>-329.9992</td>
<td>0.01781</td>
<td>9.053966</td>
<td>4.80E+09</td>
</tr>
<tr>
<td></td>
<td>-329.9999</td>
<td>-329.9996</td>
<td>0.00129</td>
<td></td>
<td>-329.9979</td>
<td>-329.9939</td>
</tr>
<tr>
<td>f_2</td>
<td>116963.81</td>
<td>216822.59</td>
<td>232813.60</td>
<td>65798.359</td>
<td>806115.43</td>
<td>1555663.2</td>
</tr>
<tr>
<td></td>
<td>26076.057</td>
<td>54623.792</td>
<td>58720.959</td>
<td>16602.364</td>
<td>47111.608</td>
<td>56879.710</td>
</tr>
<tr>
<td></td>
<td>12338.844</td>
<td>67274.542</td>
<td>71598.946</td>
<td>40728.281</td>
<td>202707.49</td>
<td>249662.99</td>
</tr>
<tr>
<td></td>
<td>14159.235</td>
<td>45973.174</td>
<td>51052.310</td>
<td>72127.105</td>
<td>29281.216</td>
<td>84296.259</td>
</tr>
</tbody>
</table>

Fig. 4. Comparison of Population diversities for PSO solving multimodal function $f_5$: (a) position, (b) velocity, (c) cognitive.
ability, an algorithm can explore more areas in the search space, and find some potential regions where “good enough” solutions may exist. On the other hand, an algorithm with the ability of exploitation can finely search the potentially good regions, and find the optimum ultimately. An algorithm should have a good balance between exploration and exploitation during the search process.

In this paper, we have analyzed the different diameter and average distance of search information propagation in PSO with different topology. PSO with star topology has the smallest diameter and average distance, which means that search information has the fastest propagation in all topologies, and on the contrary, PSO with ring topology has the largest diameter and average distance.

Topology determines the search information propagation in swarm. With an improper search information propagation, premature or low efficacy search may happen. Algorithm’s exploration and exploitation can be monitored by population diversity changing during search process. In this paper, we monitored population diversities of PSO with different topologies. Normally, we can conclude that classical PSO with ring structure has faster propagation of search information than fully informed PSO with ring structure, and correspondingly, population diversities deceases fast in classical PSO at the same iteration. In this case, fully informed PSO can preserves population diversity for more iterations, the performance of fully informed PSO with ring structure is better than classical PSO with ring structure in most benchmark functions.

However, it does not mean that the slower the search information propagates the better. The fully informed PSO with star topology shares most of search information in the whole swarm. The velocity diversity decreases to a very tiny value after few iterations. The swarm will lose “search potential”, even the position diversity and cognitive diversity preserves during the search process.

Position diversity represents the distribution of current solutions. Cognitive diversity represents the distribution of particles’ “moving target.” Position diversity should tend to follow the cognitive diversity or vice versa. Velocity diversity represents the particles’ “moving potential”. This potential should be enlarged when particles are “stuck in” a small search region, and reduced when particles find an area that “good enough” solution may exist. According to the relationship between position diversity and cognitive diversity, it could improve the algorithm’s performance by dynamically adjusting particles’ velocities [5].

In this paper, population diversities are monitored on PSO solving single-objective problems. Multiobjective problems have different goal of optimization, it does not find one single solution, but many. The concept of convergence has different meanings between single and multiple objective problems [19]. Population diversity is also important when applying PSO to solve multi-objective problems [20]. Defining a population diversity for a multi-objective particle swarm optimization algorithm (MOPSO) and monitor its changing during PSO search process is our future research work.

ACKNOWLEDGMENT
The authors’ work is partially supported by National Natural Science Foundation of China under grant No.60975080.

REFERENCES