Convergence Time Analysis of Particle Swarm Optimization Based on Particle Interaction

Chao-Hong Chen
Ying-ping Chen

Natural Computing Laboratory (NCLab)
Department of Computer Science
National Chiao Tung University
329 Engineering Building C
1001 Ta Hsueh Road
HsinChu City 300, TAIWAN
http://nclab.tw/

Convergence Time Analysis of Particle Swarm Optimization Based on Particle Interaction

Chao-Hong Chen and Ying-ping Chen
Department of Computer Science
National Chiao Tung University
HsinChu City 300, Taiwan
chchen@nclab.tw, ypchen@nclab.tw

March 17, 2010

Abstract

In this paper, we analyze the convergence time of particle swarm optimization (PSO) on the facet of particle interaction. We firstly introduce a statistical interpretation of social-only PSO in order to capture the essence of particle interaction, which is one of the key mechanisms of PSO. We then use the statistical model to obtain theoretical results on the convergence time. Since the theoretical analysis is conducted on the social-only model of PSO, instead of on common models in practice, to verify the validity of our results, numerical experiments are executed on benchmark functions with a regular PSO program.

1 Introduction

Particle swarm optimizer (PSO), introduced by Kennedy and Eberhart in 1995 [1, 2], is a stochastic, population-based algorithm for solving continuous optimization problems. As shown in [3] and by lots of real-world applications, PSO is an efficient and effective optimization framework. Although PSO has been widely applied in many fields [4, 5, 6, 7], understanding of PSO from the theoretical point of view is still quite limited. Most of previous theoretical results [8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18] are derived under the system that assumes a fixed attractor or a swarm consisting of one single particle. In these studies, particle interaction, one of the key mechanisms of PSO, are obviously not taken into consideration.

Due to the lack of theoretical analysis on PSO particle interaction, in this paper, we will make an attempt to analyze the convergence time for PSO on the facet of particle interaction. In particular, we will firstly introduce a statistical interpretation of PSO, proposed in [19], to capture the essence of particle interaction. We will then analyze the convergence time based on the statistical model. Finally, numerical experiments will be conducted to confirm the validity of our theoretical results obtained on simplified PSO, the social-only model, in a normal PSO configuration.

In the next section, we will briefly introduce the algorithm of PSO and the statistical interpretation of social-only PSO. In section 3, we will analyze the convergence time of PSO based on the statistical model. The experimental results are presented in section 4, followed by section 5 which concludes this work.

```
procedure PSO(Objective function F : ℝⁿ → ℝ)
    Initialize m particles
    while the stopping criterion is not satisfied do
        for i = 1, 2, ..., m do
            if F(Xᵢ) < F(Pbᵢ) then
                Pbᵢ ← Xᵢ
                if F(Pbᵢ) < F(Nb) then
                    Nb ← Pbᵢ
                end if
            end if
        end for
        for i = 1, 2, ..., m do
            Vᵢ ← wVᵢ + Cₚ(Pbᵢ − Xᵢ) + Cₙ(Nb − Xᵢ)
            Xᵢ ← Xᵢ + Vᵢ
        end for
    end while
end procedure
```

Figure 1: Standard PSO

2 Particle swarm optimization and the statistical interpretation

The standard PSO can be described as pseudo code shown in Figure 1. In this paper, we will use boldface for vectors, e.g., $\mathbf{X_i}$, $\mathbf{V_i}$. Without loss of generality, we assume that the goal is to minimize the objective function.

According to Figure 1, in the beginning, m particles are initialized, where m is the swarm size, an algorithmic parameter of PSO. Each particle contains three types of information: its location ($\mathbf{X_i}$), velocity($\mathbf{V_i}$), and personal best position ($\mathbf{Pb_i}$). At each generation, each particle updates its personal best position ($\mathbf{Pb_i}$) and neighborhood best position ($\mathbf{Nb}$) according to its objective value. After updating the personal and neighborhood best positions, each particle updates the velocity according to $\mathbf{Pb_i}$ and $\mathbf{Nb}$. In the velocity update formula, w is the weight of inertia which is usually a constant. C_p and C_n are random values sampled from uniform distributions $U(0, c_p)$ and $U(0, c_n)$, where c_p and c_n are called acceleration coefficients. Finally, each particle updates its position according to the velocity and then go to next generation.

From the aforementioned brief description, we can already see that particle interaction is a crucial mechanism in the design of PSO. Although there have been previous studies on particle interaction and the PSO behavior, most of these studies were totally based on the assumption of fixed attractors, a false condition for PSO in action. In order to take particle interaction into consideration, we use an alternative view of PSO that regards the whole swarm as a unity. Instead of tracking the movement of each particle, we consider the overall swarm behavior and we conceptually transform the state of an entire swarm into a statistical abstraction. Furthermore, in order to concentrate on particle interaction, we adopt the social-only model of PSO [20] which does not consider personal best positions.

The statistical interpretation of PSO, proposed in [19], is summarized in Figure 2. In the statistical model, the exact particle locations are not traced but modeled as a distribution θ over $\mathbb{R}^n$. Velocities are viewed as random vectors $\mathcal{V} \in \mathbb{R}^n$. The swarm size m is considered as the number of samples from distribution θ, since the geographic knowledge is embodied in the distribution, the neighborhood attractor can be viewed as the best of the m samples.

procedure STATISTICAL INTERPRETATION OF PSO(Objective function $\mathcal{F}: \mathbb{R}^n \to \mathbb{R}$)

    Initialize:$\sigma \leftarrow \sigma_{\mathbf{0}}, \mu \leftarrow \mu_{\mathbf{0}}$

    while the stopping criterion is not satisfied **do**

        for $i = 1, 2, \ldots, m$ **do**

            for $j = 1, 2, \ldots, n$ **do**

                $\mathbf{P_i}_j \sim N(\mu_j, \sigma_j^2)$

            end for

        end for

        $\mathbf{P_a} = \min_{\mathbf{P_i}}\{\mathcal{F}(\mathbf{P_i})\}$

        for $i = 1, 2, \ldots, m$ **do**

            $\mathbf{P_i}' \leftarrow \mathbf{P_i} + C(\mathbf{P_a} - \mathbf{P_i})$

        end for

        $\mu \leftarrow (\sum_{i=1}^{m} \mathbf{P_i}')/m$

        $\sigma^2 \leftarrow \mathrm{MLE}(\mathbf{P_1'}, \mathbf{P_2'}, \ldots, \mathbf{P_m'})$

    end while

end procedure

Figure 2: Statistical model of PSO. Distribution θ is represented by $\mu = (\mu_1, \mu_2, \ldots, \mu_n)$ and $\sigma = (\sigma_1, \sigma_2, \ldots, \sigma_n)$. Acceleration coefficient $C \sim U(0, c)$.

Each particle $\mathbf{P_i}$ is considered as a random vector sampled from θ, and the velocity $\mathbf{V_i}$ is sampled from $\mathcal{V}$. The neighborhood attractor can then be defined as

$$\mathbf{P_a} := \min_{\mathbf{P_i}}\{\mathcal{F}(P_1), \mathcal{F}(P_2), \ldots, \mathcal{F}(P_m)\} \, .$$

At each generation, $\mathbf{P_i}$ is updated as $\mathbf{P_i} + w\mathbf{V_i} + C(\mathbf{P_a} - \mathbf{P_i})$. The next distribution is thus the statistical characterization denoted by functions of the observed values:

$$\theta \leftarrow \mathcal{T}_p(\mathbf{P_1}, \ldots, \mathbf{P_m}),$$
$$\mathcal{V} \leftarrow \mathcal{T}_v(\mathbf{P_1}, \ldots, \mathbf{P_m}, \mathbf{V_1}, \ldots, \mathbf{V_m}).$$

Since w is a constant, distribution $\mathcal{V}$ can be removed because given two random vectors $\mathbf{X} \sim \theta$ and $\mathbf{V} \sim \mathcal{V}$, we can simply let θ' be the distribution of $\mathbf{X}' := \mathbf{X} + w\mathbf{V}$.

For simplicity, in this paper, we consider the positions of each dimension of a particle is independently sampled from distribution θ_i. Consider the random variable $X \sim \theta_i$ and let $E[X] = \mu$. If we divide the support of θ_i into s disjoint regions $R_1, \ldots, R_s$ such that $Prob[X \in R_i] = 1/s$ for $i = 1, 2, \ldots, s$, and each region is associated with a random variable of velocity $V_i \sim \mathcal{V}_i$. By picking $x_i \in R_i$ for each region, when s is sufficiently large, the swarm can be characterized as

$$\sum_{i=1}^{s} \frac{1}{s}(x_i + V_i) = \sum_{i=1}^{s} \frac{x_i}{s} + \sum_{i=1}^{s} \frac{V_i}{s} \approx \mu + \sum_{i=1}^{s} \frac{V_i}{s} \, .$$

Each component of $\sum_{i=1}^{s} V_i/s$ can be approximated with a normal distribution by the central limit theorem. As a consequence, normal distributions are a reasonable choice for describing the behavior of the entire swarm. We let the distribution of i-th dimension, θ_i, be $N(\mu_i, \sigma_i^2)$, where $N(\mu_i, \sigma_i^2)$ is the normal distribution with mean μ_i and variance σ_i^2. The update of distribution becomes simply calculate the mean and the variance.

The mean can be calculated by taking the average of updated positions, and the variance is calculated by using the maximum likelihood estimation (MLE). Let $\sigma_{\mathbf{t}i}^2$ and $\mu_{\mathbf{t}i}^2$ be the variance and mean of the i-th dimension at the t-th generation. Let $y_j = \mathbf{P_j}_i$ and $y_j' = \mathbf{P_j'}_i$ for $j =$

$1, 2, \ldots, m$ and let $y_a = \mathbf{P}_{\mathbf{a}i}$, $\overline{y} = (1/m) \sum_{j=1}^{m} y'_j$. To estimate the variance of the i-th dimension at the $t+1$-th generation, we use the maximum likelihood estimation (MLE). The likelihood function $L(\sigma_{\mathbf{t}i}^2)$ is defined as the joint probability

$$L(\sigma_{\mathbf{t}i}^2) := \prod_{j=1}^{m} (\frac{1}{\sqrt{2\pi}\sigma_{\mathbf{t+1}i}^2}) \exp(\frac{-(y'_j - \overline{y})^2}{2\sigma_{\mathbf{t+1}i}^2}) = (\frac{1}{\sqrt{2\pi}\sigma_{\mathbf{t+1}i}^2})^m \exp(\frac{-\sum_{j=1}^{m}(y'_j - \overline{y})^2}{2\sigma_{\mathbf{t+1}i}^2}) \, .$$

To find $\sigma_{\mathbf{t+1}i}^2$ that maximizes $L(\sigma_{\mathbf{t}i}^2)$, we differentiate $L(\sigma_{\mathbf{t}i}^2)$ with respect to $\sigma_{\mathbf{t+1}i}^2$:

$$L'(\sigma_{\mathbf{t}i}^2) = -(\frac{m}{2})(\frac{1}{\sqrt{2\pi}})^m \sigma_{\mathbf{t+1}i}^{-m-2} \cdot \exp(\frac{-\sum_{j=1}^{m}(y'_j - \overline{y})^2}{2\sigma_{\mathbf{t+1}i}^2})$$
$$+ (\frac{1}{\sqrt{2\pi}})^m \cdot \frac{\sum_{j=1}^{m}(y'_j - \overline{y})^2}{2} \sigma_{\mathbf{t+1}i}^{-m-4} \cdot \exp(\frac{-\sum_{j=1}^{m}(y'_j - \overline{y})^2}{2\sigma_{\mathbf{t+1}i}^2}) \, ,$$

the value of $\sigma_{\mathbf{t+1}i}^2$ that maximizes $L(\sigma_{\mathbf{t}i}^2)$ is $\sum_{j=1}^{m}(y'_j - \overline{y})^2/m$. As a result, in our model of PSO, the results of MLE is $\sigma_{\mathbf{t+1}i}^2 = \sum_{j=1}^{m}(y'_j - \overline{y})^2/m$ for $i = 1, 2, \ldots, n$.

3 Convergence time analysis

In this section, we will analyze the PSO convergence time based on the aforementioned statistical interpretation of the social-only model. As the first step, we must define the state of convergence. Since in this work, we regard the entire swarm as a distribution, the state of convergence is then referred to the variance of the distribution. We define the state of convergence as the variance for every dimension is less than a given value $\epsilon > 0$. By using this definition, we can now start our analysis of PSO convergence time. To estimate the variance after distribution update, we need following lemma from [21]:

Lemma 1. *Let $X_1, X_2, \ldots, X_m \sim N(\mu, \sigma^2)$. Define $S = \sum_{i=1}^{m}(X_i - \overline{X})^2/(m-1)$, where $\overline{X} = \sum_{i=1}^{m} X_i/m$. We have $(m-1)S \sim \sigma^2 \chi_{m-1}^2$, where χ_{m-1}^2 is the chi-square distribution with $m-1$ degrees of freedom.*

With this lemma, we can obtain

Lemma 2. *Given the swarm size m, acceleration coefficient c, and variance of the i-th dimension at the t-th generation $\sigma_{\mathbf{t}i}^2$, we have $E[\sigma_{\mathbf{t+1}i}^2] = [\frac{1}{3}c^2 - c + 1][(m-1)/m]\sigma_{\mathbf{t}i}^2$.*

Proof. We know $\sigma_{\mathbf{t+1}i}^2 = \sum_{j=1}^{m}(y'_j - \overline{y})^2/m$. The expected value is

$$E[\frac{1}{m}\sum_{j=1}^{m}(y'_j - \frac{\sum_{k=1}^{m} y'_k}{m})^2] = \frac{1}{m}E[\sum_{j=1}^{m}(y_j + C(y_a - y_j) - \frac{\sum_{k=1}^{m} y_k + C(y_a - y_k)}{m})^2]$$
$$= \frac{1}{m}E[\sum_{j=1}^{m}(\frac{m(1-C)y_j - (1-C)\sum_{k=1}^{m} y_k}{m})^2]$$
$$= \frac{1}{m^3}E[(1-C)^2 \sum_{j=1}^{m}(my_j - \sum_{k=1}^{m} y_k)^2]$$
$$= \frac{1}{m}E[(1-C)^2]E[\sum_{j=1}^{m}(y_j - \frac{1}{m}\sum_{k=1}^{m} y_k)^2] \, .$$

Let $S = \sum_{j=1}^{m}(y_j - \frac{1}{m}\sum_{k=1}^{m} y_k)^2$. Since $y_j \sim N(\mu_{\mathbf{t+1}_i}, \sigma_{\mathbf{t+1}_i}^2)$ for $j = 1, 2, \ldots, m$ and $y_1, y_2, \ldots, y_m$ are i.i.d., by Lemma 1, $S \sim \sigma_{\mathbf{t}_i}^2 \chi_{m-1}^2$, and $E[S] = (m-1)\sigma_{\mathbf{t}_i}^2$. Then, we can obtain

$$
\begin{aligned}
E[\sigma_{\mathbf{t+1}_i}^2] &= \frac{1}{m}E[(1-C)^2]E[\sum_{j=1}^{m}(y_j - \frac{1}{m}\sum_{k=1}^{m} y_k)^2] \\
&= \frac{1}{m}E[1 - 2C + C^2]E[S] \\
&= \frac{1}{m}(\frac{1}{3}c^2 - c + 1)(m-1)\sigma_{\mathbf{t}_i}^2 \\
&= (\frac{1}{3}c^2 - c + 1)\frac{m-1}{m}\sigma_{\mathbf{t}_i}^2 .
\end{aligned}
$$

$\square$

Lemma 2 is derived under the condition that $\sigma_{\mathbf{t}_i}^2$ is given. The following lemma will derive the relationship between $E[\sigma_{\mathbf{t}_i}^2]$ and $E[\sigma_{\mathbf{t+1}_i}^2]$.

Lemma 3. $E[\sigma_{\mathbf{t+1}_i}^2] = (\frac{1}{3}c^2 - c + 1)\frac{m-1}{m}E[\sigma_{\mathbf{t}_i}^2]$.

Proof.

$$
\begin{aligned}
E[\sigma_{\mathbf{t}_i}^2] &= \int_{\sigma^2 \in \mathbb{R}^+} E[\sigma_{\mathbf{t+1}_i}^2 | \sigma_{\mathbf{t}_i}^2 = \sigma^2] Prob\{\sigma_{\mathbf{t}_i}^2 = \sigma^2\}d\sigma^2 \\
&= \int_{\sigma^2 \in \mathbb{R}^+} (\frac{1}{3}c^2 - c + 1)\frac{m-1}{m}\sigma^2 Prob\{\sigma_{\mathbf{t}_i}^2 = \sigma^2\}d\sigma^2 \text{(by lemma 2)} \\
&= (\frac{1}{3}c^2 - c + 1)\frac{m-1}{m}E[\sigma_{\mathbf{t}_i}^2] .
\end{aligned}
$$

$\square$

Now, we can obtain the relationship of convergence time and algorithmic parameters of PSO:

Theorem 4. *Given swarm size m, acceleration coefficient c, ϵ, and $\sigma_{\mathbf{0}}$. Let $h = \max_i\{\sigma_{\mathbf{0}_i}^2\}$. we have $E[\sigma_{\mathbf{t}_i}^2] < \epsilon$ for $i = 1, 2, \ldots, n$ when $[\frac{1}{3}c^2 - c + 1][(m-1)/m] < 1$ and $t > \log(\epsilon/\sigma_{0_h}^2)/\log([\frac{1}{3}c^2 - c + 1][(m-1)/m])$.*

Proof. From Lemma 3, we know

$$
E[\sigma_{\mathbf{t}_i}^2] = (\frac{1}{3}c^2 - c + 1)^t (\frac{m-1}{m})^t E[\sigma_{\mathbf{0}_i}^2] = (\frac{1}{3}c^2 - c + 1)^t (\frac{m-1}{m})^t \sigma_{\mathbf{0}_i}^2 .
$$

Since $[\frac{1}{3}c^2 - c + 1][(m-1)/m] < 1$, we have

$$
E[\sigma_{\mathbf{t}_i}^2] = (\frac{1}{3}c^2 - c + 1)^t (\frac{m-1}{m})^t \sigma_{\mathbf{0}_i}^2 < (\frac{1}{3}c^2 - c + 1)^t (\frac{m-1}{m})^t \sigma_{\mathbf{0}_h}^2 < \epsilon .
$$

The last inequality holds because

$$
\begin{aligned}
&\log\left((\frac{1}{3}c^2 - c + 1)^t (\frac{m-1}{m})^t \sigma_{\mathbf{0}_h}^2\right) = t\log\left((\frac{1}{3}c^2 - c + 1)(\frac{m-1}{m})\right) + \log\sigma_{\mathbf{0}_h}^2 \\
&< \log(\epsilon/\sigma_{0_h}^2) + \log\sigma_{\mathbf{0}_h}^2 = \log\epsilon .
\end{aligned}
$$

$\square$

We have two corollaries immediately from Theorem 4:

Corollary 5. *Given swarm size m, acceleration coefficient c, and level of convergence ϵ such that $[\frac{1}{3}c^2 - c + 1][(m-1)/m] < 1$ and $\epsilon < 1$, we have $E[\sigma_{\mathbf{t}i}^2] < \epsilon$ for $i = 1, 2, \ldots, n$ for $t = O(-\log \epsilon)$.*

Corollary 6. *Given swarm size m, c, and ϵ such that $[\frac{1}{3}c^2 - c + 1][(m-1)/m] < 1$ and $\epsilon < 1$, there exists a constant $c' < 1$ such that for $t = O(-1/\log c'(1 - 1/m))$, we have $E[\sigma_{\mathbf{t}i}^2] < \epsilon$ for $i = 1, 2, \ldots, n$.*

Corollary 5 reveals the linear relationship between the level of convergence and the convergence time, and the interpretation of Corollary 6 is that when the swarm size is sufficiently large, the effect of enlarging swarm size on the convergence time is not important. In the next section, we will empirically examine the two corollaries with a common, practical PSO configuration.

4 Experiments

In this section, we verify the validity of Corollaries 5 and 6 by running standard PSO. We use two objective functions in our experiments:

- Sphere function [22]: $f_1(\mathbf{x}) = \sum_{i=1}^{D} x_i^2$, where $\mathbf{x} \in [-100, 100]^D$.

- Schwefel's problem 1.2 [22]: $f_2(\mathbf{x}) = \sum_{i=1}^{D} \left(\sum_{j=1}^{i} x_j^2 \right)$, where $\mathbf{x} \in [-100, 100]^D$.

We have $D = 10$ for both $f_1(\mathbf{x})$ and $f_2(\mathbf{x})$ in the following experiments.

We firstly examine Corollary 5. The PSO algorithmic parameters are given as $c_p = 1$, $c_n = 1$, $w = 1/(2 \ln 2)$, and swarm size = 50. The value of ϵ is varied from 10^{-1} to 10^{-10}. For each value of ϵ, we perform 100 independent runs. For each run, we count the number of generations from initialization to the state in which variances for all dimensions are smaller than ϵ, and we calculate the mean number of generations for the 100 runs.

The comparison of these experimental results and our theoretical results is shown in Figures 3 and 4. From Figure 3, we can see that the experimental results of $f_1(\mathbf{x})$ are very close to $-4.6 \log \epsilon + 43 = O(-\log \epsilon)$, and from Figure 4, the experimental results of $f_2(\mathbf{x})$ are very close to $-4.7 \log \epsilon + 43.5 = O(-\log \epsilon)$. The experimental results agree with our estimation in Corollary 5, in which the value of $-\ln \epsilon$ and the PSO convergence time are linearly related.

After Corollary 5 is empirically verified with standard PSO, we now examine Corollary 6. The parameters we used in PSO are given as $c_p = 1$, $c_n = 1$, $w = 1/(2 \ln 2)$, and $\epsilon = 10^{-6}$. The swarm size ranges from 50 to 1000 with step 5. For each swarm size, we perform 100 independent runs and record the mean as we did in last experiment. The comparison of experimental and theoretical results is shown in Figures 5, 6,7, and 8. From Figures 5 and 6, we can see that the convergence time is close to $-64.9/\log 0.555(1 - 1/m) = \mathcal{O}(-1/\log c'(1 - 1/m))$, where $c' = 0.555$, and in Figures 7 and 8, the convergence time is close to $-99.75/\log 0.405(1 - 1/m) = \mathcal{O}(-1/\log c'(1 - 1/m))$, where $c' = 0.405$. As we can observe from these figures, when the swarm size becomes large, the increase of convergence time is insignificant, confirming our estimation in Corollary 6.

5 Conclusions

In this paper, a statistical interpretation of a simplified model of PSO was adopted to analyze the PSO convergence time. In order to capture the essence of particle interaction, the statistical model adopted in this paper assumed no fixed attractors. The effect of particle interaction were included in our analysis. Our theoretical results revealed the relationship between the convergence time and the level of convergence as well as the relationship between the convergence time and the swarm size. Numerical results, in the standard settings of PSO, were obtained
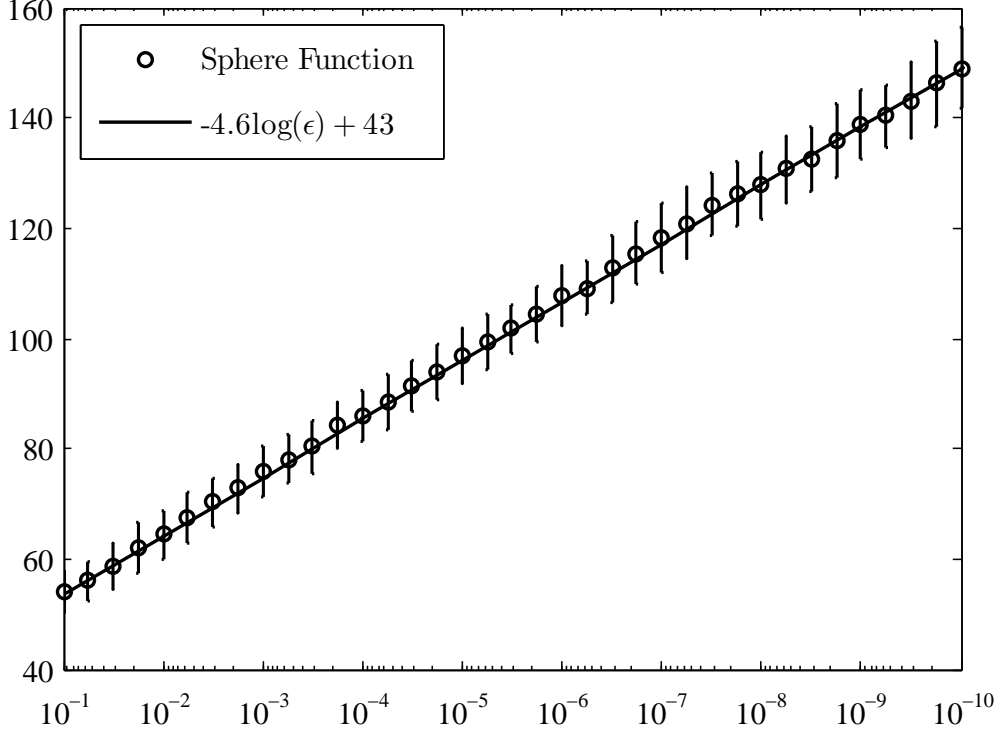
Figure 3: Comparison of experimental results and theoretical results from Corollary 5 of $f_1(\mathbf{x})$. The x-axis represents the value of ϵ, and y-axis represents the mean number of generation. The experimental results are very close to $\mathcal{O}(-\log \epsilon)$.
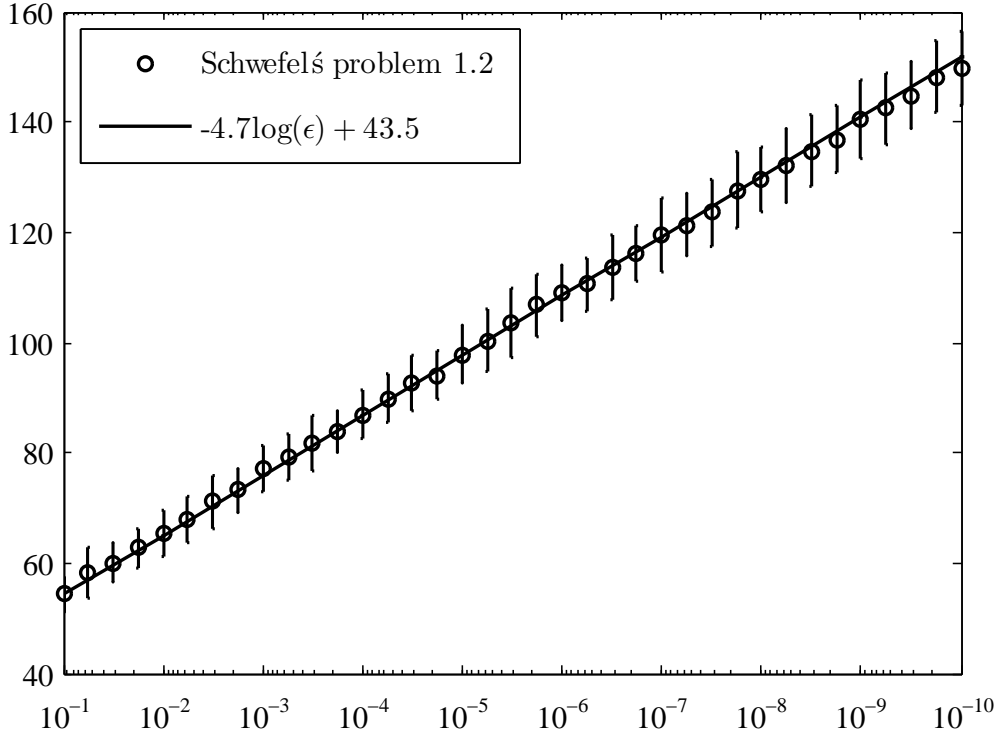

Figure 4: Comparison of experimental results and theoretical results from Corollary 5 of $f_2(\mathbf{x})$. The x-axis represents the value of ϵ, and y-axis represents the mean number of generation. The experimental results are very close to $\mathcal{O}(-\log \epsilon)$.
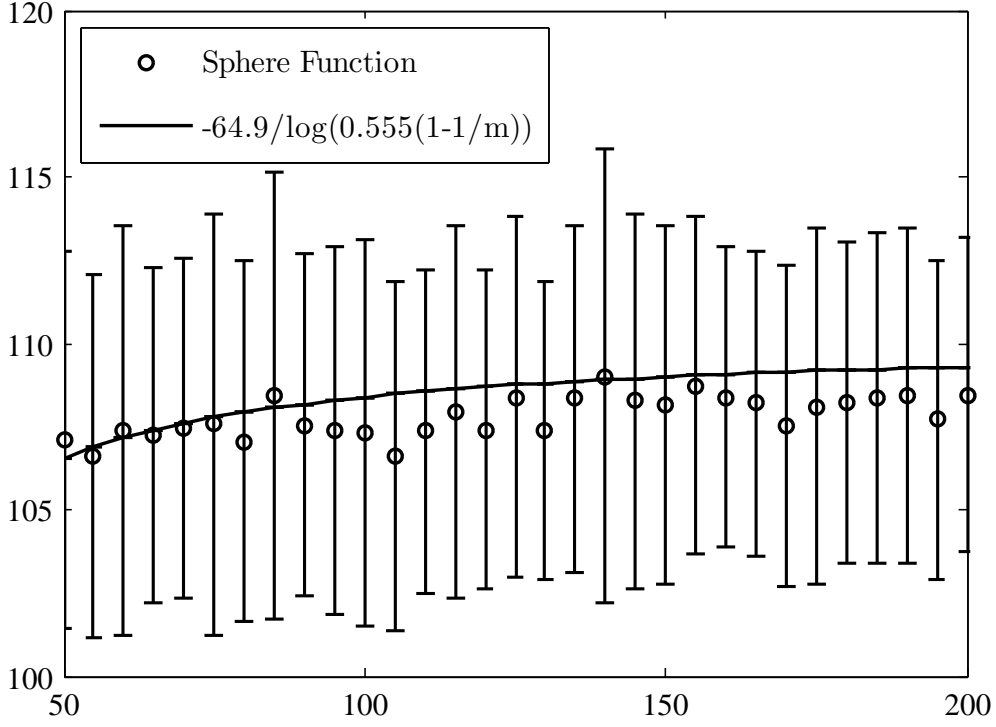
Figure 5: Comparison of experimental results and theoretical results from Corollary 6 of $f_1(\mathbf{x})$. x-axis represents the swarm size ranging from 50 to 200, and y-axis represents the mean number of generation. The experimental results are very close to $\mathcal{O}(-1/\log c'(1 - 1/m))$ with $c' < 1$.
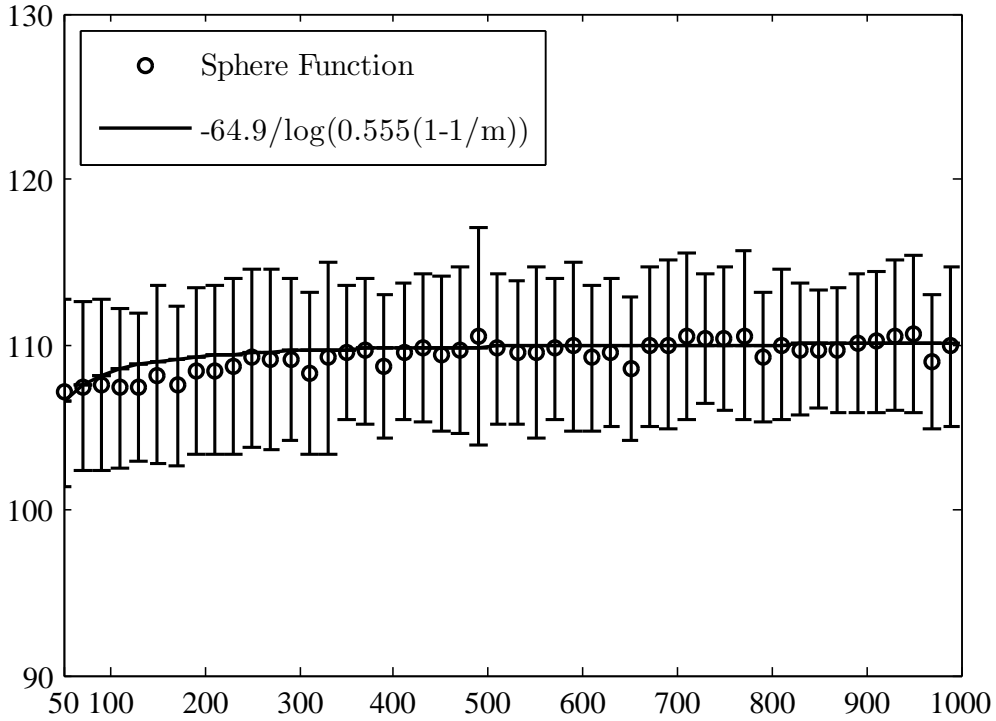

Figure 6: Comparison of experimental results and theoretical results from Corollary 6 of $f_1(\mathbf{x})$. x-axis represents the swarm size ranging from 50 to 1000, and y-axis represents the mean number of generation. The experimental results are very close to $\mathcal{O}(-1/\log c'(1 - 1/m))$ with $c' < 1$.
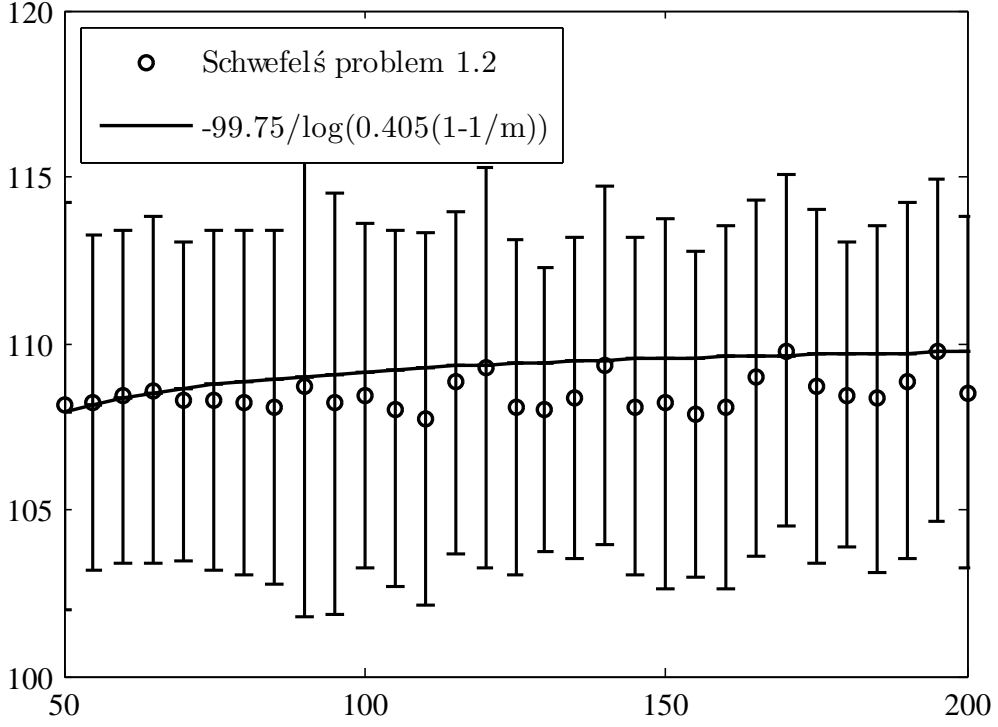
8

Figure 7: Comparison of experimental results and theoretical results from Corollary 6 of $f_2(\mathbf{x})$. x-axis represents the swarm size ranging from 50 to 200, and y-axis represents the mean number of generation. The experimental results are very close to $\mathcal{O}(-1/\log c'(1-1/m))$ with $c' < 1$.
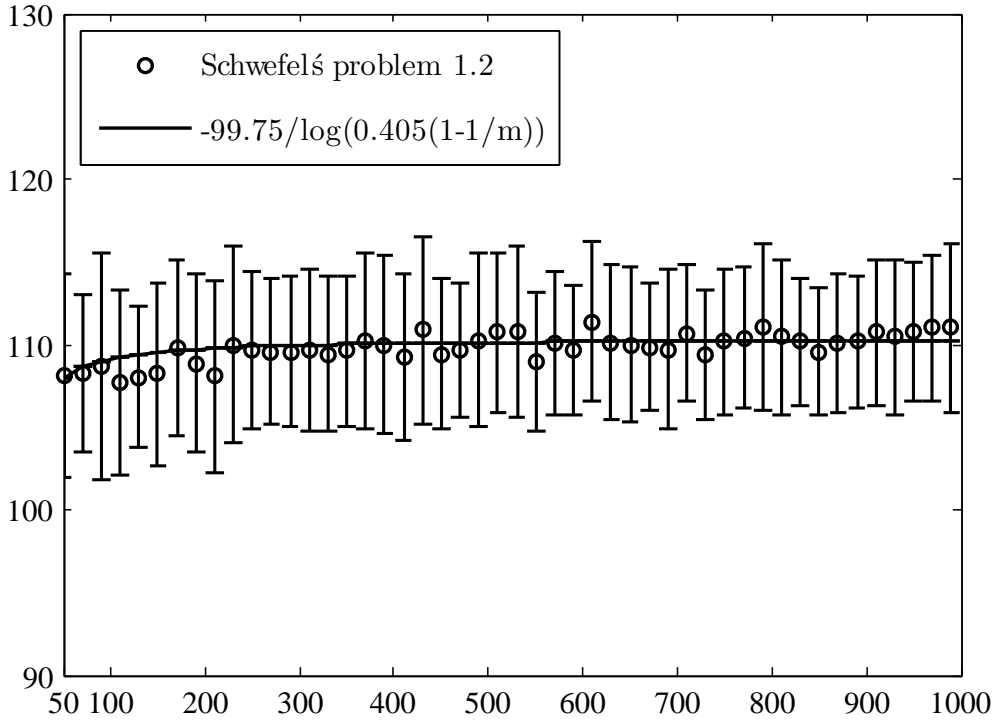

Figure 8: Comparison of experimental results and theoretical results from Corollary 6 of $f_2(\mathbf{x})$. x-axis represents the swarm size ranging from 50 to 1000, and y-axis represents the mean number of generation. The experimental results are very close to $\mathcal{O}(-1/\log c'(1-1/m))$ with $c' < 1$.

9

to empirically verify our theoretical results derived with a simplified PSO configuration. The agreement between the experimental and theoretical results indicated the importance of particle interaction in PSO. Consequently, more research effort should be invested into analyzing the working of particle interaction in order to better understand particle swarm optimization.

Acknowledgments

References

[1] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *Proceedings of IEEE International Conference on Neural Networks*, vol. 1942-1948, 1995.

[2] J. Kennedy and R. C. Eberhart, "A new optimizer using paritcle swarm theory," in *Proceeding of the Sixth Int. Symposium on Micromachine and Human Science*, Nagoya, Japan, 1995, pp. 39–43.

[3] Y. Shi and R. C. Eberhart, "Empirical study of particle swarm optimization," in *Evolutionary Computation, 1999. CEC 99. Proceedings of the 1999 Congress on*, vol. 3, 1999, pp. 1945–1950.

[4] F. Pan, G. Wang, and Y. Liu, "A multi-objective-based non-stationary UAV assignment model for constraints handling using PSO," in *GEC '09: Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation*, 2009, pp. 459–466.

[5] P. Yan and L. Tang, "PSO algorithm for a scheduling parallel unit batch process with batching," in *GEC '09: Proceedings of the first ACM/SIGEVO Summit on Genetic and Evolutionary Computation*, 2009, pp. 703–708.

[6] T. M. Alkhamis and M. A. Ahmed, "Simulation-based optimization for repairable systems using particle swarm algorithm," in *WSC '05: Proceedings of the 37th conference on Winter simulation*, 2005, pp. 857–861.

[7] R. Wrobel and P. Mellor, "Particle swarm optimisation for the design of brushless permanent magnet machines," in *Industry Applications Conference, 2006. 41st IAS Annual Meeting. Conference Record of the 2006 IEEE*, vol. 4, Oct. 2006, pp. 1891–1897.

[8] J. Kennedy, "The behavior of particles," in *Proceedings of the 7th International Conference on Evolutionary Programming*, 1998, pp. 581–589.

[9] E. Ozcan and C. K. Mohan, "Analysis of a simple particle swarm optimization system," *Intelligent Engineering Systems Through Artificial Neural Networks*, vol. 8, pp. 253–258, 1998.

[10] ——, "Particle swarm optimization: surfing the waves," in *Proceedings of 1999 IEEE Congress on Evolutionary Computation (CEC 99)*, 1999, pp. 1939–1944.

[11] 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.

[12] F. van den Bergh, "An analysis of particle swarm optimizers," Ph.D. dissertation, University of Pretoria, 2002.

[13] K. Yasuda, A. Ide, and N. Iwasaki, "Adaptive particle swarm optimization," in *Proceedings of 1999 IEEE International Conference on Systems, Man and Cybernetics*, 2003, pp. 1554–1559.

[14] Y.-L. Zheng, L.-H. Ma, L.-Y. Zhang, and J.-X. Qian, "On the convergence analysis and parameter selection in particle swarm optimization," in *Proceedings of the Second International Conference on Machine Learning and Cybernetics*, 2003, pp. 1802–1807.

[15] I. C. Trelea, "The particle swarm optimization algorithm: convergence analysis and parameter selection," *Information Processing Letters*, vol. 85, pp. 317–325, 2003.

[16] F. van den Bergh and A. P. Engelbrecht, "A study of particle swarm optimization particle trajectories," *Information Sciences*, vol. 176, no. 8, pp. 937–971, April 2006.

[17] V. Kadirkamanathan, K. Selvarajah, and P. J. Fleming, "Stability analysis of the particle dynamics in particle swarm optimizer," *IEEE Transactions on Evolutionary Computation*, vol. 10, no. 3, pp. 245–255, 2006.

[18] M. Jiang, Y. Luo, and S. Yang, "Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm," *Information Processing Letters*, vol. 102, pp. 8–16, 2007.

[19] Y.-p. Chen and P. Jiang, "Analysis on the facet of particle interaction in particle swarm optimization," *Theoretical Computer Science*, 2010, doi: 10.1016/j.tcs.2010.03.003.

[20] J. Kennedy, "The particle swarm: social adaptation of knowledge," in *Proceedings of the 1997 IEEE International Conference on Evolutionary Computation*, 1997, pp. 303–308.

[21] M. R. Spiegel, *Schaum's Outline of Theory and Problems of Probability and Statistics*. Mcgraw-Hill, 1975, ISBN: 0-07060-220-4.

[22] X. Yao, Y. Liu, and G. Lin, "Evolutionary programming made faster," *Evolutionary Computation, IEEE Transactions on*, vol. 3, no. 2, pp. 82–102, Jul 1999.