# Explanation of Stagnation at Points that are not Local Optima in Particle Swarm Optimization by Potential Analysis

Alexander Raß  Manuel Schmitt  Rolf Wanka
Department of Computer Science, University of Erlangen-Nuremberg, Germany
{alexander.rass, manuel.schmitt, rolf.wanka}@fau.de

## ABSTRACT

This paper investigates the frequently observed phenomenon of *stagnation* which appears on particle swarm optimization (PSO). We introduce a measure of significance of single dimensions and provide experimental and theoretical evidence that the classical PSO, even with swarm parameters known (from the literature) to be *good*, almost surely does *not* converge to a local optimum (*stagnation*) if too few particles are used. Stagnation is an undesirable property of PSO.

## Categories and Subject Descriptors

I.2.8 [**Artificial Intelligence**]: Problem Solving, Control Methods, and Search—*heuristic methods*

## Keywords

Particle swarm optimization; stagnation; potential

## 1. INTRODUCTION

Particle swarm optimization (PSO), is a popular nature-inspired meta-heuristic for solving continuous optimization problems. The popularity of the PSO framework is due to the fact that it shows a good tradeoff between simplicity and quality of the results.

However, during the execution of the PSO one encounters stagnation, i. e., the particles converge to a solution which is no local optimum. To track such stagnation, we introduce a stagnation measure, which is a multidimensional extension to the potential introduced in [4, 5]. A sketch of the developed model is presented, which provides sufficient insights to prove that PSO stagnates almost surely. [3] supplies the full version of this paper with more detailed definitions, assumptions, theorems, proofs and experiments.

## 2. STAGNATION IN PSO

**Definition 1 (Classical PSO process)** *A* swarm *of $N$ particles moves through the $D$-dimensional search space $\mathbb{R}^D$. Let $f : \mathbb{R}^D \to \mathbb{R}$ be the objective function. At each time $t \in \mathbb{N}$, each particle $n$ has a position $X_t^n \in \mathbb{R}^D$, a velocity $V_t^n \in \mathbb{R}^D$ and a local attractor $L_t^n \in \mathbb{R}^D$, the best position particle $n$ has visited until time $t$. Additionally, the swarm shares its global attractor $G_t^n \in \mathbb{R}^D$, storing the best position of* any *particle until time $t$ and (as attractors are updated iteratively) the first $n-1$ particles at time $t+1$. For $t > 0$, $1 \le n \le N$ the positions and velocities are determined by the following* movement equations*:*

$$V_{t+1}^n := \chi \cdot V_t^n + c_1 \cdot r_t^n \odot (L_t^n - X_t^n) + c_2 \cdot s_t^n \odot (G_t^n - X_t^n)$$
$$X_{t+1}^n := X_t^n + V_{t+1}^n.$$

*$\chi$, $c_1$ and $c_2$ are positive constants called the* fixed parameters *of the swarm and $r_t^n$, $s_t^n$ are uniformly distributed over $[0,1]^D$ and independent. $\odot$ denotes the item-wise product of vectors. We write $\mathcal{A}_t$ for the natural filtration of the process.*

For every step the stagnation measure is a $D$-dimensional vector of potentials. It is intended that the greater the value of such a potential for a single dimension is, the greater is the portion of the change in the function value which is due to the movement in that dimension. Furthermore, the logarithmic potential is defined, which compares the impact of a specific dimension with the maximal impact along all dimensions. Since the convergence analysis in [2] implies that the general movement of a converging particle swarm drops exponentially, a logarithmic scale is used and linear decrease is expected.

**Definition 2 (Potential)** *Let $f$ be the objective function. $\Phi(t,d) := \max_{1\le n\le N} |f(X_t^n) - f(\tilde{X}_t^{n,d})|$ is called the* potential *where $\left(\tilde{X}_t^{n,d}\right)_{\tilde{d}} := X_t^{n,\tilde{d}} + V_t^{n,\tilde{d}}$, if $\tilde{d} = d$, $X_t^{n,\tilde{d}}$, otherwise. $\Psi(t,d) := \log_2\left(\Phi(t,d)/\max_{1\le\tilde{d}\le D}\Phi(t,\tilde{d})\right)$ is called the* logarithmic potential *and $I_{t,d} := \Psi(t+1,d) - \Psi(t,d)$ is called* increment of dimension *$d$.*

Here only sphere function $f_{sph}(x) := \sum_{i=1}^{D} x_i^2$ is presented (in [3], we study more functions). We use *good* swarm parameters $\chi = 0.72984$ and $c_1 = c_2 = 1.496172$, as proposed in [1]. To receive reliable results, the PSO is simulated with at least 2 000 bits precision for the mantissa, because with double precision soon all positions in single dimensions evaluate to the same value and the swarm stops moving in that dimensions. In Fig. 1 the logarithmic potential of some dimensions decreases (approximately) linearly, while the global attractor in these dimensions hardly improves (stagnation). Such periods of time, when stagnation occurs, are captured by the following definition.

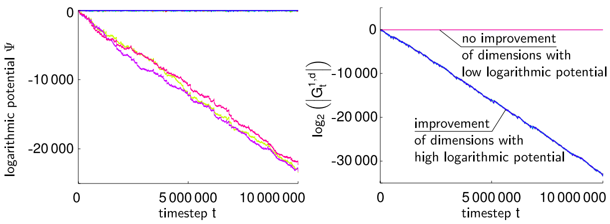


*GECCO '15 Companion, July 11–15, 2015, Madrid, Spain*



ACM ISBN 978-1-4503-3488-4/15/07.

DOI: http://dx.doi.org/10.1145/2739482.2764654




(a) Logarithmic potential $\Psi$ (b) $\log_2 |G_t^{1,d}|$

Figure 1: (a) $\Psi(t,d)$, (b) $\log_2 |G_t^{1,d}|$ on $f_{sph}$, $D = 10$, $N = 3$; each line represents a single dimension

**Definition 3** **($(N_0, c_0, c_s)$-Stagnation phase)** *Let $N_0 < D$ be a positive integer constant and $c_0 \leq c_s < 0$ negative constants. We define the random set $S(t_1, t_2) := \{d : \Psi(t_1, d) \leq c_0 \wedge \max_{t_1 \leq t \leq t_2} \Psi(t, d) \leq c_s\}$ and stopping times $\beta_{-1} := 0$ and inductively for all $i \geq 0$ $\alpha_i := \inf\{t \geq \beta_{i-1} : |S(t,t)| \geq N_0\}$, $\beta_i := \inf\{t \geq \alpha_i : |S(\alpha_i, t)| < N_0\}$. $\alpha_i$ defines the start and $\beta_i$ defines the end of the i'th ($N_0$, $c_0$, $c_s$)-stagnation phase, if the values are finite. A dimension d is called* stagnating at time $t$ *if there is an i such that $\alpha_i \leq t < \beta_i$ and $d \in S(\alpha_i, t)$.*

Informally: During a stagnation phase there are at least $N_0$ dimensions which stay insignificant. As the logarithmic potential decreases linearly (Figure 1) a model with independent increments is reasonable. The behavior of logarithmic potential in stagnating dimensions is very similar. Therefore in our model the increments are composed by independent base increments, which are equal for all dimensions, and independent dimension dependent increments. The following assumption describes a simplified version of the model. For a more refined model using weaker assumptions see [3].

**Assumption 1 (Separation of logarithmic potential)** *It is assumed that for objective functions $f$ there exist $N_0$, $c_0$, $c_s$ as in Definition 3, $(\Gamma, \Theta)$ two random probability distributions on $\mathbb{R}$, random variables $(B_t)_{t \in \mathbb{N}}$, which will be called* base increments*, and $(J_{t,d})_{t \in \mathbb{N}, d \leq D}$, which will be called* dimension dependent increments*, such that for all $t \in \mathbb{N}$ and $d \leq D$ $B_t$, $J_{t,d}$ are $\mathcal{A}_{t+1}$-measurable, $B_t$, $J_{t,1}$, $\ldots$, $J_{t,D}$ are independent of $\mathcal{A}_t$ and of each other, $B_t$ has distribution $\Gamma$ and $J_{t,d}$ has distribution $\Theta$, $\mathrm{E}[J_{0,0}] = 0$, $\mathrm{E}[B_0] < 0$, $\mathrm{E}[(J_{0,0})^k]$ and $\mathrm{E}[(B_0)^k]$ exist and are finite for all $1 \leq k \leq 6$ and if $d$ is stagnating at time $t$ then the increments of dimension $d$ fulfill the equation $I_{t,d} = B_t + J_{t,d}$*

In this model we can prove the following theorem:

**Theorem 1** *A $(N_0, c_0, c_s)$-stagnation phase does, with positive probability, never end.*

As a consequence there is a time $T$, from which on the swarm infinitely stagnates, if the expected waiting time for a stagnation phase is finite. For the proof of this theorem we mainly need that the sixth central moments of sums of $t$ consecutive increments increase only as $O(t^3)$, which is supplied by the assumed independence, but also experiments confirm that property. As one can see in Fig. 1, there is no improvement in stagnating dimensions, because they have insignificant influence on the swarm. If dimensions stagnate during long periods of time, there is no reason for any specific position in that dimension to have positive probability (measure theoretical sense), which result in a further assumption on the model.

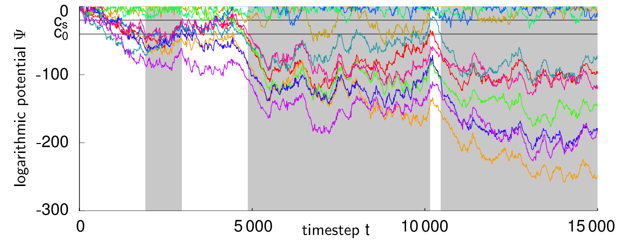



Figure 2: $\Psi(t,d)$ on $f_{sph}$, $D = 10$, $N = 2$; $c_0 = -40$, $c_s = -20$, $N_0 = 7$; each line represents a single dimension

**Assumption 2 (Insignificance of stagn. dim.)** *If a stagnation phase does not end, then dimensions which stagnate through the whole phase do not converge to any specific point with positive probability.*

With this assumption we can prove the final theorem, which proves that PSO stagnates:

**Theorem 2** *If the swarm converges almost surely, the objective function $f$ has only a countable number of local optima and for all $i$ with $\mathrm{P}(\beta_{i-1} < \infty) > 0$, the expectation $\mathrm{E}[\alpha_i - \beta_{i-1} \mid \beta_{i-1} < \infty] \leq T$ for some constant $T$, then the swarm converges not to a local optimum, almost surely.*

The bounded expectation in this theorem only specifies that the expected time from the end of a stagnation phase to the start of the next stagnation phase is bounded. Fig. 2 is an illustration how logarithmic potentials develope in the beginning, where stagnation phases are marked with a gray background.

## 3. REFERENCES


[1] M. Clerc and J. Kennedy. The particle swarm – explosion, stability, and convergence in a multidimensional complex space. *IEEE Transactions on Evolutionary Computation*, 6:58–73, 2002.

[2] M. Jiang, Y. P. Luo, and S. Y. Yang. Particle swarm optimization – stochastic trajectory analysis and parameter selection. In F. T. S. Chan and M. K. Tiwari, editors, *Swarm Intelligence – Focus on Ant and Particle Swarm Optimization*, chapter 17, pages 179–198. I-TECH Education and Publishing, Vienna, 2007.

[3] A. Raß, M. Schmitt, and R. Wanka. Explanation of stagnation at points that are not local optima in particle swarm optimization by potential analysis. *arXiv*, abs/1504.08241, 2015.

[4] M. Schmitt and R. Wanka. Particle swarm optimization almost surely finds local optima. In *Proc. 15th Genetic and Evolutionary Computation Conference (GECCO)*, pages 1629–1636, 2013.

[5] M. Schmitt and R. Wanka. Particle swarm optimization almost surely finds local optima. *Theoretical Computer Science*, 561:57–72, 2015.