Impacts of coefficients on movement patterns in the particle swarm optimization algorithm

Mohammad Reza Bonyadi and Zbigniew Michalewicz

***Abstract*—In this paper we investigate movement patterns of a particle in the particle swarm optimization (PSO) algorithm. We characterize movement patterns of the particle by two factors: the correlation between it's consecutive positions and it's range of movement. We introduce the base frequency of movement as a measure for the correlation between positions and the variance of movement as a measure for the range of movement. We determine the base frequency and the variance of movement theoretically and we show how they change with the values of coefficients. We extract a system of equations that enables practitioners to find coefficients' values to guarantee achieving a given base frequency and variance of movement, i.e., control the movement pattern of particles. We also show that if the base frequency of movement for a particle is small, mid range, or large then the particle's position at each iteration is positively correlated (smooth movement), uncorrelated (chaotic movement), or negatively correlated (jumping at each iteration) with its previous positions, respectively. We test the effects of the base frequency and variance of movement on the search ability of particles and we show that small base frequencies (i.e., smooth movement) are more effective when the maximum number of function evaluations is large. We found that the most frequently-used coefficient values in PSO literature impose mid-range base frequencies that correspond with a chaotic movement. We also provide new sets of coefficients that outperform existing ones on a set of benchmark functions.**

***Index Terms*—Particle swarm optimization; base frequency; correlation**

I. Introduction

PARTICLE swarm optimization (PSO) is a stochastic population-based optimization algorithm developed by Kennedy and Eberhart [1]. PSO has been applied to many optimization problems such as artificial neural network training, pattern classification, and function optimization [2], [3], to name a few. Since 1995, different aspects of the original version of PSO have been investigated and many variants of the algorithm have been proposed (see [4] for a comprehensive review on PSO). Although there have been many studies related to local convergence, invariance, and stability of particles, there are not many articles on the movement patterns of particles before convergence to their equilibrium [4].

The aim of this paper is to study two factors, i.e., the frequency and the variance of movement, to characterize movement patterns of particles. We investigate the frequency and variance for a formulation that represents a wide range of PSO position update rules and we show how these factors affect the patterns of movement when that formulation is used. We also formulate these two factors in terms of coefficients of particles for a specific type of PSO so that a given frequency and variance are achieved during the run. This formulation can be used to change the movement patterns of particles during the run so that an adequate pattern can be selected according to the current state of the search (e.g., search space characteristics, iteration number). We experiment with different frequencies and variances to understand which values for each of these factors are more effective for optimizing a benchmark of optimization problems.

Without loss of generality, this paper only considers minimization problems defined as follows:

$$\text{find } \vec{x} \in S \subseteq \mathbb{R}^d \text{ such that } \forall \vec{y} \in S, f(\vec{x}) \leq f(\vec{y}) \quad (1)$$

where S is the *search space* defined by $\{\vec{x} | l_i \leq x_i \leq u_i \text{ for all } i\}$, l_i and u_i are lower bound and upper bound of the values of the i^{th} dimension of S, d is the number of dimensions, and $f : \mathbb{R}^d \to \mathbb{R}$ is the objective function. The set of points that are generated by $f(\vec{x})$ for all $\vec{x} \in S$ is called the *landscape*.

The rest of this paper is organized as follows. In section II, we overview the formulation of the original PSO and we give some background on calculation of variance and movement patterns. In section III we present our proposed approach that includes calculation of the base frequency of oscillation, the variance of movement, and coefficients to achieve a given base frequency and variance. In section IV we test our theoretical findings and compare our coefficient settings with various settings proposed in other articles. Section V concludes the paper and provides a discussion on some directions for further research.

II. Background

In this section we provide some background on the early variants of PSO, existing studies on the variance of movement, and different patterns of movement.

M. R. Bonyadi (reza@cai.uq.edu.au, rezabny@gmail.com) is with the Centre for Advanced Imaging (CAI), the University of Queensland, QLD 4067, Australia, and the Optimisation and Logistics Group, the University of Adelaide, SA 5005, Australia.

Z. Michalewicz (zbyszek@cs.adelaide.edu.au) is the Optimisation and Logistics Group, the University of Adelaide, SA 5005, Australia, Complexica Pty Ltd, Suite 75, 155 Brebner Drive, West Lakes, SA 5021, Australia, Polish-Japanese Academy of Information Technology, ul. Koszykowa 86, 02-008 Warsaw, Poland, and the Institute of Computer Science, Polish Academy of Sciences, ul. Ordona 21, 01-237 Warsaw, Poland.

A. Particle swarm optimization

Each particle in the Original PSO (OPSO) [5], [1] consists of three vectors:

- Position ($\vec{x}_t{}^i$) — is the position of the i^{th} particle in the t^{th} iteration. This is used to evaluate the particle quality;
- Velocity ($\vec{v}_t{}^i$) — is the direction and length of movement of the i^{th} particle in the t^{th} iteration; and
- Personal best ($\vec{p}_t{}^i$) — is the best position (in terms of objective value) that the particle i has visited until iteration t. The role of this vector is to store the knowledge of best found solutions [5].

In OPSO, the velocity of each particle is updated for the next iteration $(t+1)$ by

$$\vec{v}_{t+1}^{\,i} = \vec{v}_t^{\,i} + c_1 R_{1t}\left(\vec{p}_t^{\,i} - \vec{x}_t^{\,i}\right) + c_2 R_{2t}\left(\vec{g}_t - \vec{x}_t^{\,i}\right) \quad (2)$$

where c_1 and c_2 are two real numbers known as *acceleration coefficients*, $\vec{p}_t{}^i$ is the personal best of the particle i at iteration t. We define:

$$\tau_t = argmin_i\{f(\vec{p}_t^{\,i})\}$$

as the *the global best particle* and $\vec{g}_t = \vec{p}_t^{\,\tau_t}$ as the *global best vector*. The vector $\vec{p}_t{}^i$ for each particle i is updated by Eq. 3.

$$\vec{p}_{t+1}^{\,i} = \begin{cases} \vec{x}_{t+1}^{\,i} & f\left(\vec{x}_{t+1}^{\,i}\right) < f\left(\vec{p}_t^{\,i}\right) - \epsilon_0 \text{ and } \vec{x}_{t+1}^{\,i} \in S \\ \vec{p}_t^{\,i} & \text{otherwise} \end{cases} \quad (3)$$

where $\epsilon_0 > 0$ is an arbitrarily small real value that represents the precision of the calculations. This constant can be set to the smallest possible value in the simulations.

Particles are attracted by $\vec{p}_t^{\,i} - \vec{x}_t^{\,i}$ (personal influence) and $\vec{g}_t - \vec{x}_t^{\,i}$ (social influence) to move toward known quality solutions found until iteration t, i.e., $\vec{p}_t^{\,i}$ and $\vec{g}_t$. Further, R_{1t} and R_{2t} are two randomly (a uniform distribution in the interval [0, 1]) generated $d \times d$ diagonal matrices [6], [7]. These two matrices are generated for each particle i at every iteration t separately. The position of a particle i is updated by

$$\vec{x}_{t+1}^{\,i} = \vec{x}_t^{\,i} + \vec{v}_{t+1}^{\,i} \quad (4)$$

OPSO was studied by many researchers since 1995 and many new variants were proposed. For example, it was proposed [8] to multiply the previous velocity ($\vec{v}_t{}^i$) by an inertia weight (ω) to control the impact of $\vec{v}_t{}^i$ on the movement of particles (we refer to this PSO variant as the Inertia PSO, IPSO, in this paper). The velocity update rule for IPSO was written as

$$\vec{v}_{t+1}^{\,i} = \omega \vec{v}_t^{\,i} + c_1 R_{1t}\left(\vec{p}_t^{\,i} - \vec{x}_t^{\,i}\right) + c_2 R_{2t}\left(\vec{g}_t - \vec{x}_t^{\,i}\right) \quad (5)$$

where ω is inertia weight.

B. Variance of movement

Most articles investigated the variance of movement of an arbitrary particle in OPSO or IPSO in the one-dimensional case [4]. The reason is that the position and velocity update for these methods is conducted at each dimension separately, hence, analyses in a one-dimensional space is generalizable to multidimensional space. Thus, we drop the vector notation in equations when we present ideas that are based on such update rules.

Perhaps the first study that investigated the variance of particle positions for IPSO was [9] that was extended further in [10]. These studies proved that if the variance of the sequence of positions generated by an arbitrary particle (denoted by x as it is in a one dimensional space) converges to a fixed point then $c < \frac{12(\omega^2-1)}{5\omega-7}$ where $c = c_1 = c_2$. Also, the fixed point for the variance of these positions (shown by V_x) was calculated as

$$V_x = \frac{c(\omega+1)}{4(c(5\omega-7) - 12\omega^2 + 12)}(g-p)^2 \quad (6)$$

The assumption in these studies was that p and g are not updated during the run (so called the *stagnation assumption*). The convergence of variance of the global best particle in IPSO under a more general assumption was investigated in [11] and it was proven that the convergence boundaries found by [9] as well as [10] are valid under a weak stagnation assumption.

The analysis of variance in [9], [10], and [11] was restricted to a uniform distribution for the acceleration coefficients and constant values for the inertia weight. This restriction was eliminated in [12] where it was assumed that the inertia weight is a random variable with the expected value μ_ω and the variance σ_ω^2 and acceleration coefficients are random variables with expected values μ_{ϕ_1} and μ_{ϕ_2} and variances $\sigma_{\phi_1}^2$ and $\sigma_{\phi_2}^2$. Also, it was assumed that both p and g are random variables (with some mean and variance) that are updated during the run (see Eq. 7), that weakens the stagnation assumption. The position update rule for a particle in PSO was then written as:

$$x_{t+1} = (1 + \omega - \phi_1 - \phi_2)x_t - \omega x_{t-1} + \phi_1 p + \phi_2 g \quad (7)$$

where ϕ_1, ϕ_2, p, g, and ω are random variables with given expected values (μ) and standard deviations (σ). Note that this formulation represents a wide range of PSO variants including IPSO and OPSO. It was proven [12] that the convergence of the variance of positions generated by this formulation is not dependent on the mean and variance of the random variables p and g. Convergence boundaries for this formulation were derived and it was proven that, even if p and g move during the run, those boundaries include all coefficients combinations that guarantee convergence of variance of positions (necessary condition). Experiments in [12] showed that there is no coefficients combination inside those boundaries that causes divergence of the variance of positions, i.e., those boundaries are necessary (proven) and sufficient (experimentally showed) for the convergence of the variance of positions. The fixed point of variance for Eq. 7 was calculated as:

$$V_x = -\frac{k_3 + k_4}{k_1 k_2} \quad (8)$$

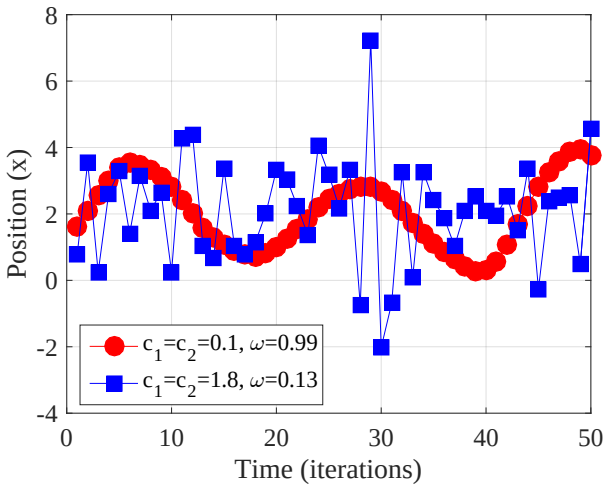

Fig. 1. Different coefficients result in different fluctuations, IPSO with $\omega = 0.99$ and $c_1 = c_2 = 0.1$ fluctuates smoothly while a particle with $\omega = 0.13$ and $c_1 = c_2 = 1.8$ oscillates more chaotically

where

- $k_1 = (\mu_{\phi_1} + \mu_{\phi_2})^2$
- $k_2 = k_1(1 - \mu_\omega) + 2(\mu_{\phi_1} + \mu_{\phi_2})(\mu_\omega^2 + \sigma_\omega^2 - 1) + (\sigma_{\phi_1}^2 + \sigma_{\phi_2}^2)(\mu_\omega + 1)$,
- $k_3 = k_1(\mu_\omega + 1)(\mu_{\phi_1}^2\sigma_p^2 + \mu_{\phi_2}^2\sigma_g^2 + \sigma_{\phi_1}^2\sigma_p^2 + \sigma_{\phi_2}^2\sigma_g^2)$,
- $k_4 = (\mu_{\phi_1}^2\sigma_{\phi_2}^2 + \mu_{\phi_2}^2\sigma_{\phi_1}^2)(\mu_\omega + 1)(\mu_g - \mu_p)^2$.

If we consider that p, g, and ω are constants and ϕ_1 and ϕ_2 follow uniform distributions with $\mu_{\phi_1} = \frac{c_1}{2}$, $\mu_{\phi_2} = \frac{c_2}{2}$, $\mu_\omega = \omega$, $\sigma_{\phi_1} = \frac{c_1}{\sqrt{12}}$, $\sigma_{\phi_2} = \frac{c_2}{\sqrt{12}}$, $\sigma_\omega = 0$ (i.e., IPSO settings), then, after simplifications, Eq. 8 will be exactly the same as what was found in [10] (Eq. 6) for IPSO.

See also [4] for a comprehensive review on theoretical investigations of PSO.

C. Patterns of movement

The trajectory of x_t (Eq. 7) exhibits different patterns that are controlled by the values of coefficients [13], [14]. For example, a particle in IPSO with $\omega = 0.99$ and $c_1 = c_2 = 0.1$ oscillates smoothly while a particle with $\omega = 0.13$ and $c_1 = c_2 = 1.8$ oscillates more chaotically (see Fig. 1). These patterns play an important role in the performance of the algorithm. As an example, a particle that moves smoothly in the search space can be potentially more effective at the latter stages of the search process than a particle that jumps all over the search space [14]. Hence, investigation of these patterns and calculation of their corresponding coefficients can lead us to further improvement of the method.

Despite the importance of these patterns, there have not been many articles to study them in detail [4]. The trajectory of x_t in OPSO was investigated in [15] where $c_1 R_{1,t}$ and $c_2 R_{2,t}$ were replaced by constants. Both the movement pattern and the magnitude of x_t were investigated theoretically for that simplified system and the effects of changing coefficients were visually illustrated.

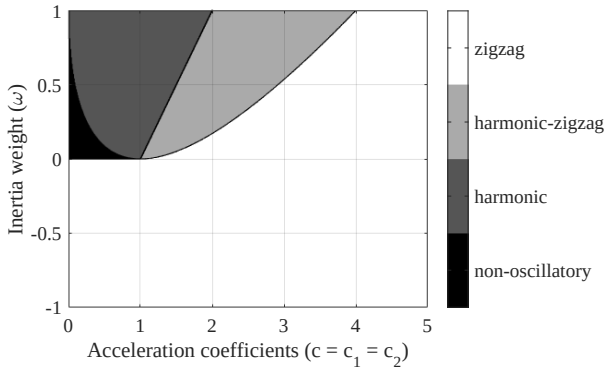

Fig. 2. Categorization of different patterns of oscillation by [13]. The gray levels indicate corresponding coefficients boundaries to non-oscillatory, harmonic, harmonic-zigzagging, and zigzagging movement patterns.

The movement patterns of the expectation of x_t in IPSO was categorized into 4 groups [13]: non-oscillatory (particle position does not oscillate during the run), harmonic (particle position oscillates smoothly similar to a wave), zigzagging (particle position oscillate significantly at each iteration), and harmonic-zigzagging (combination of significant oscillation and wave-like oscillation). It was found [13] that different patterns are observed by changing the values of coefficients. Fig. 2 illustrates the relationship between coefficients values and movement patterns of particles positions expectations found in [13].

These categories were also investigated in [14] through some experiments. It was found that the so-called *base frequency* of particle positions (the frequency of the largest amplitude among the Fourier series coefficients of the particle positions) has a direct relationship with the patterns of oscillation. This observation was used to estimate (based on some experiments) the boundaries corresponding to different oscillation patterns of IPSO and another PSO variant.

III. Patterns of Movement: Frequency and Variance

The *pattern of movement* for a particle refers to the way the position of that particle (x_t) changes during the run that is a function of the coefficients values of that particle (see Fig. 1). We study two factors related to the patterns of movement for a particle: *range of movement* and *base frequency*. The range of movement for a particle is the size of the area that bounds positions the particle experiences in its lifetime. The base frequency for a particle is the frequency of the largest amplitude among the Fourier series coefficients of the particle positions.

In this section we investigate the relationship between the factors related to movement patterns (i.e., range of movement and the base frequency) and the particle's coefficients in PSO. We propose a system of equations (section III-C) for which the solutions are coefficient

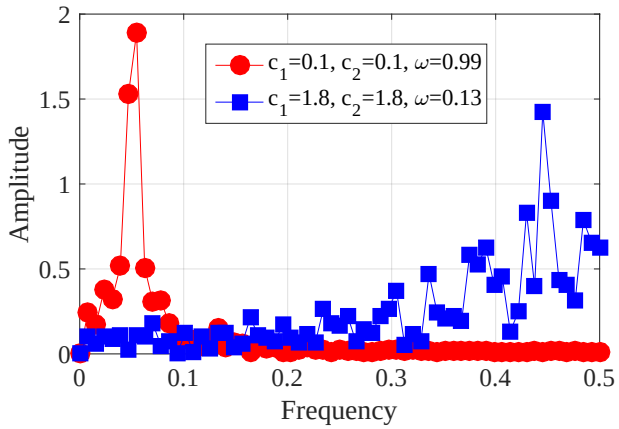

Fig. 3. The particle positions in Fig. 1 in the frequency domain.

values that guarantee to achieve a given base frequency and variance. We investigate this system of equations in more detail for a specific type of PSO, IPSO (see definition 1). We also investigate the relationship between the base frequency and the correlation between consecutive positions (section III-D) and the variance and the range of movement (section III-E) in detail.

Definition 1. *IPSO is defined by Eq. 7 where both ϕ_1 and ϕ_2 follow uniform distribution with $\mu_{\phi_1} = c_1/2$, $\mu_{\phi_2} = c_2/2$ (c_1 and c_2 are constants), $\sigma_{\phi_1} = \frac{2\mu_{\phi_1}}{\sqrt{12}}$, and $\sigma_{\phi_2} = \frac{2\mu_{\phi_2}}{\sqrt{12}}$, and ω is a constant ($\sigma_\omega = 0$, $\mu_\omega = \omega$).*

Note that Definition 1 is equivalent to the definition of IPSO presented in section II-A.

A. Base frequency of movement (F)

In PSO variants, the base frequency of positions (see section II), denoted by F throughout the paper, is directly related to the oscillation pattern of movement [14]. Fig. 3 indicates that the base frequency is small (0.05 Hz) for a particle with a smooth movement (see Fig. 1) while it is large (almost 0.45 Hz) for a particle with a more chaotic movement. In this subsection, we calculate the base frequency of the expectation of movement through theoretical analysis of Eq. 7. One should note that $F \in [0, 0.5]$ because of the discrete time in simulation, represented by the iteration number. The base frequency equal to 0 indicates that the position of the particle does not oscillate at all while the base frequency 0.5 means that the position of the particle oscillates at every iteration.

The expected position of a particle can be calculated by applying the expectation operator to Eq. 7:

$$E(x_{t+1}) = lE(x_t) - \mu_\omega E(x_{t-1}) + \mu_{\phi_1}\mu_p + \mu_{\phi_2}\mu_g \quad (9)$$

where $E(x_{t+1})$ (E_{t+1} for short) is the expectation of x_{t+1}, μ_ω, μ_{ϕ_1}, μ_{ϕ_2}, μ_p, and μ_g, are the expected values of ω, ϕ_1, ϕ_2, p, and g, respectively, and $l = 1 + \mu_\omega - \mu_{\phi_1} - \mu_{\phi_2}$.

The characteristic equation for the recursion in Eq. 9 is written as:

$$\gamma^2 - l\gamma + \mu_\omega = 0 \quad (10)$$

The roots of this equation (γ_1 and γ_2) are given by:

$$\gamma_1 = \frac{l}{2} + \frac{\sqrt{\Delta}}{2}, \gamma_2 = \frac{l}{2} - \frac{\sqrt{\Delta}}{2}$$

where $\Delta = l^2 - 4\mu_\omega$.

The expectation of position is convergent [12] if and only if the absolute value of both roots of the characteristic equation are smaller than 1 (i.e., the spectrum radius is smaller than 1). This entails:

$$-1 < \mu_\omega < 1 \text{ and } 0 < \mu_\phi < 2(\mu_\omega + 1) \quad (11)$$

Let us assume that the expected value of the position of particles is convergent and it converges to a value E_x. Thus, $E_x = lE_x + \mu_\omega E_x + \mu_{\phi_1}\mu_p + \mu_{\phi_2}\mu_g$, that results in:

$$E_x = \frac{\mu_{\phi_1}\mu_p + \mu_{\phi_2}\mu_g}{\mu_{\phi_1} + \mu_{\phi_2}} \quad (12)$$

There are three cases for the solution to the recursion in Eq. 9:

1) $E_t = r^t(a \times \cos(\theta t) + b \times \sin(\theta t)) + K$ if the characteristic equation has two distinct complex roots ($\gamma_1 = r\angle\theta$ and $\gamma_2 = r\angle -\theta$),
2) $E_t = a\gamma_1^t + b\gamma_2^t + K$ if the characteristic equation has two distinct real roots,
3) $E_t = a\gamma_1^t + bt\gamma_2^t + K$ if characteristic equation has one repeated real root,

The values of a, b, and K are constants that are calculated according to the initial conditions[1]. Clearly, if the sequence is convergent (i.e., Eq. 11 holds) then $K = E_x$.

Lemma 1. *If the roots of the characteristic equation of the recursion in Eq. 9 (formulated in Eq. 10) are complex then*

$$F = \frac{tan^{-1}(\frac{im(\sqrt{\Delta})}{l})}{2\pi} \quad (13)$$

where F is the base frequency of the expectation of positions generated by Eq. 7.

Proof. If the roots of the characteristic equation are complex then $im(\gamma_1) = im(\frac{\sqrt{\Delta}}{2})$ (where $im(\beta)$ is the imaginary component of β), $im(\gamma_2) = -im(\gamma_1)$, and $re(\gamma_1) = re(\gamma_2) = \frac{l}{2}$ (where $re(\beta)$ is the real component of β). Hence, $r = \sqrt{im(\gamma_1)^2 + re(\gamma_1)^2} = \sqrt{im(\gamma_2)^2 + re(\gamma_2)^2}$ and $\theta = tan^{-1}(im(\gamma_1)/re(\gamma_1)) = -tan^{-1}(im(\gamma_2)/re(\gamma_2))$. Because the solution of the characteristic equation is written as a summation of *sin* and *cos* functions, the

[1]As the values of x_0 and v_0 are known (initialized), the value of $E(x_0)$ and $E_(x_1)$ are also known ($E(x_0) = x_0$ and $E(x_1) = E(x_0) + E(v_0)$). Also, γ_1 and γ_2 can be found by solving Eq. 10. Hence, one can form a system of equations that involves $E(x_0)$ and $E(x_1)$ to find the values of a and b. Also, note that $K = E_t$ as we assumed that Eq. 11 holds.

sequence E_t is periodic and oscillatory. The base frequency of oscillation is related to the value of θ and it is calculated by:

$$F = \frac{\theta}{2\pi} = \frac{tan^{-1}(\frac{im(\sqrt{\Delta})}{l})}{2\pi} \tag{14}$$

Eq. 13 determines the oscillation frequency of the expectation of the particle movement when the roots are complex. □

Let us investigate the cases where the roots are real.

Lemma 2. *Assume that the roots of the characteristic equation (Eq. 10) are real. We define λ as:*

$$\lambda = \begin{cases} \gamma_1 & \text{if } |\gamma_1| > |\gamma_2| \\ \gamma_2 & \text{if } |\gamma_1| < |\gamma_2| \\ min(\gamma_1, \gamma_2) & \text{if } |\gamma_1| = |\gamma_2| \end{cases} \tag{15}$$

where γ_1 and γ_2 are the roots of the characteristic equation (Eq. 10).

- *if $\lambda < 0$ then $F = 0.5$*
- *if $\lambda \geq 0$ then $F = 0$*

Proof. There are two possibilities when the roots are real: there is one repeated roots or they are distinct.

If the roots are repeated ($\lambda = \gamma_1 = \gamma_2$) then $E_t = a\gamma^t + bt\gamma^t + K$. If $\lambda < 0$ then the sign of $a\lambda^t + bt\lambda^t$ changes at every iteration that corresponds to a zigzag pattern ($F = 0.5$). If $\lambda \geq 0$ then the sign of $a\lambda^t + bt\lambda^t$ does not change that corresponds to a no-oscillation pattern ($F = 0$).

Now we assume that the roots are distinct ($\gamma_1 \neq \gamma_2$). There are two cases: $|\gamma_1| \neq |\gamma_2|$ and $|\gamma_1| = |\gamma_2|$.

Let $|\gamma_1| \neq |\gamma_2|$. Without loss of generality, we can assume that $\lambda = \gamma_1$, i.e., $|\gamma_1| > |\gamma_2|$. In this case, for any a and b, there exists an iteration t_0 that, for any $t > t_0$ we have $|a\gamma_1^t| > |b\gamma_2^t|$, i.e., $a\gamma_1^t$ dominates $b\gamma_2^t$ from the iteration t_0. If $\gamma_1 < 0$ then, because $a\gamma_1^t$ dominates $b\gamma_2^t$ from the iteration t_0, the sign of $a\gamma_1^t + b\gamma_2^t$ changes at every iteration after t_0. Therefore, if $E_t < K$ then $E_{t+1} > K$ and if $E_t > K$ then $E_{t+1} < K$ at every iteration. This means that E_t will follow a zigzag pattern after iteration t_0, thus, $F = 0.5$. If $\gamma_1 \geq 0$ then $a\gamma_1^t + b\gamma_2^t \geq 0$ for all $t > t_0$. Hence, the value of E_t becomes closer to K from one side that refers to a non-oscillatory pattern, $F = 0$.

If $|\gamma_1| = |\gamma_2|$ then (recall that $\gamma_1 \neq \gamma_2$) $a\gamma_1^t$ and $b\gamma_2^t$ will have opposite signs at every second iteration (depending on the sign of a and b, it can be every odd or even iteration) and the same sign at the rest of iterations. Hence, as K is constant, the pattern of movement is zigzag for which $F = 0.5$. □

One should note that the case $|\gamma_1| = |\gamma_2|$ (while $\gamma_1 \neq \gamma_2$) corresponds with $\gamma_1 = -\gamma_2$ that means the characteristic equation is in the form of $\gamma^2 - \gamma_1^2 = 0$. This means that $\mu_\omega = -\gamma_1^2$ and $l = 0$.

Lemma 3. *If both roots of the characteristic equation (Eq. 10) are real then:*

- *$l \leq 0$ implies $F = 0.5$*
- *$l > 0$ implies $F = 0$*

Proof. If $l < 0$ then $\gamma_2 < 0$. Also, as $l < 0$ and $\sqrt{\Delta} \geq 0$, clearly $|l - \sqrt{\Delta}| \geq |l + \sqrt{\Delta}|$ that means $|\gamma_2| \geq |\gamma_1|$. Therefore, according to Lemma 2, $F = 0.5$ that completes the proof for this case.

If $l > 0$ then $\gamma_1 > 0$. Also, as $l > 0$ and $\sqrt{\Delta} \geq 0$, hence $|l - \sqrt{\Delta}| \leq |l + \sqrt{\Delta}|$ that means $|\gamma_1| \geq |\gamma_2|$. Therefore, according to Lemma 2, $F = 0$ that completes the proof for this case.

If $l = 0$ then $|\gamma_1| = |\gamma_2|$ that, according to the proof of Lemma 2, it imposes $F = 0.5$. □

Lemma 4. *The frequency F for the recursion in Eq. 9 can be calculated by*

$$F = \begin{cases} 0.5 & \text{if } 1 + \mu_\omega = \mu_{\phi_1} + \mu_{\phi_2},\ \mu_\omega \leq 0 \\ \frac{tan^{-1}(\frac{im(\sqrt{\Delta})}{l})}{2\pi} & \text{otherwise} \end{cases} \tag{16}$$

regardless of the type (real or complex) of the roots for its characteristic equation (Eq. 10).

Proof. The roots of the characteristic equation are either complex or real. Let us assume that the roots are complex. In this case Eq. 13 (that is equal to the second part of the Eq. 16) gives the base frequency of E_t as discussed in Lemma 1, that completes the proof for this case.

If the roots are not complex ($im(\sqrt{\Delta}) = 0$, hence, $\Delta \geq 0$), from Lemma 3, $F = 0.5$ if $l < 0$, $F = 0$ if $l > 0$, and $F = 0.5$ if $l = 0$. It is clear that Eq. 13 returns $F = 0$ if $l > 0$ and $F = 0.5$ if $l < 0$ as well, hence, Eq. 13 can calculate the value of F correctly when $l > 0$ or $l < 0$ and $\Delta \geq 0$ (two real roots). If $\Delta \geq 0$ (two real roots) and $l = 0$ (that corresponds to $1 + \mu_\omega = \mu_{\phi_1} + \mu_{\phi_2}$ and $\Delta \geq 0$), $F = 0.5$ according to Lemma 3 while Eq. 13 is undefined with this setting. Hence, Eq. 16 returns correct value in this case.

Therefore, Eq. 16 describes the base frequency of the expectation of positions in PSO (E_t) defined by Eq. 7, no matter what the roots of the characteristic equation are. □

This lemma indicates that the value of F can be calculated via Eq. 13 if there is a guarantee that the statement "$1 + \mu_\omega = \mu_{\phi_1} + \mu_{\phi_2}$ and $\mu_\omega \leq 0$" is false. In fact, the points on the line segment $1 + \mu_\omega = \mu_{\phi_1} + \mu_{\phi_2}$ when $\mu_\omega \leq 0$ cause the nominator of Eq. 13 to become $\frac{0}{0}$ which is undefined.

We use Eq. 16 to calculate frequencies for the coefficients sets in Fig. 1. We found that $F = 0.051$ when $\omega = 0.99$, $\mu_{\phi_1} = \mu_{\phi_2} = 0.1$ and $F = 0.44$ when $\omega = 0.13$, $\mu_{\phi_1} = \mu_{\phi_2} = 1.8$ which matches almost perfectly with the simulation results. Fig. 4 shows the value of F for different values of $\mu_{\phi_1} = \mu_{\phi_2} = \mu_\phi$ and μ_ω. Clearly, this figure generalizes the findings of [13] about oscillation patterns presented in Fig. 2.

Theorem 1. *For a given F and $\mu_\omega > 0$, $\mu_{\phi_1} + \mu_{\phi_2}$ can be calculated as:*

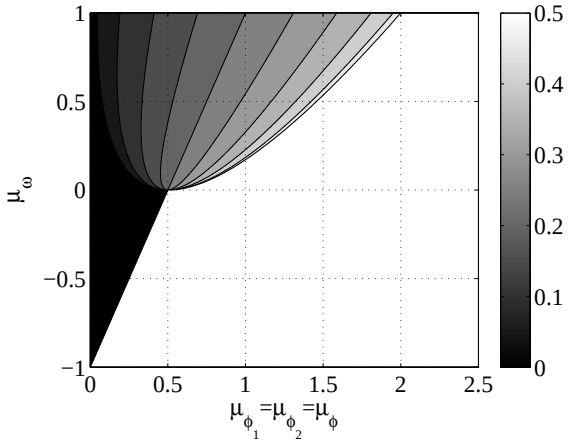

Fig. 4. The value of F for different values of μ_ω and μ_ϕ.

- *If* $F = 0$ *then* $\mu_{\phi_1} + \mu_{\phi_2} \leq 1 + \mu_\omega - 2cos(2\pi F)\sqrt{\mu_\omega}$
- *If* $F = 0.5$ *then* $\mu_{\phi_1} + \mu_{\phi_2} \geq 1 + \mu_\omega - 2cos(2\pi F)\sqrt{\mu_\omega}$
- *If* $F \in (0, 0.5)$ *then* $\mu_{\phi_1} + \mu_{\phi_2} = 1 + \mu_\omega - 2cos(2\pi F)\sqrt{\mu_\omega}$

Proof. According to Lemma 4, Eq. 13 can be used to find F if $\mu_\omega > 0$.

There are three cases if $\mu_\omega > 0$:

1) If $F = 0$ then, according to Eq. 13, $im(\sqrt{\Delta}) = 0$ (i.e., $\Delta \geq 0$) and $l > 0$. This means $l^2 \geq 4\mu_\omega$, and, as $l > 0$, hence $1 + \mu_\omega - 2\sqrt{\mu_\omega} \geq \mu_{\phi_1} + \mu_{\phi_2}$. It is then trivial to write this as: $1 + \mu_\omega - 2cos(2\pi F)\sqrt{\mu_\omega} \geq \mu_{\phi_1} + \mu_{\phi_2}$,
2) If $F = 0.5$ then, according to Eq. 13, $im(\sqrt{\Delta}) = 0$ (i.e., $\Delta \geq 0$) and $l < 0$. This means $l^2 \geq 4\mu_\omega$, and, as $l < 0$, we have $1 + \mu_\omega + 2\sqrt{\mu_\omega} \leq \mu_{\phi_1} + \mu_{\phi_2}$. It is then trivial to write this as: $1 + \mu_\omega - 2cos(2\pi F)\sqrt{\mu_\omega} \leq \mu_{\phi_1} + \mu_{\phi_2}$,
3) If $F \in (0, 0.5)$ then $im(\sqrt{\Delta}) \neq 0$ that means $\Delta < 0$ and, therefore $l^2 < 4\mu_\omega$ (both roots are complex). As $\Delta < 0$, the value of $im(\sqrt{\Delta})$ is given by $\sqrt{-\Delta}$. From Eq. 13:

$$tan(2\pi F) = \frac{\sqrt{4\mu_\omega - (1 + \mu_\omega - \mu_{\phi_1} - \mu_{\phi_2})^2}}{1 + \mu_\omega - \mu_{\phi_1} - \mu_{\phi_2}} \tag{17}$$

Obviously, the value of $tan(2\pi F)$ is positive for $F \in [0, 0.25]$ and negative for $F \in [0.25, 0.5]$. Given that the nominator is always positive, thus $1 + \mu_\omega - (\mu_{\phi_1} + \mu_{\phi_2}) < 0$ if $F \in [0, 0.25]$ and $1 + \mu_\omega - (\mu_{\phi_1} + \mu_{\phi_2}) > 0$ if $F \in [0.25, 0.5]$. We solve Eq. 17 for μ_ϕ:

$$\mu_{\phi_1} + \mu_{\phi_2} = 1 + \mu_\omega \pm 2cos(2\pi F)\sqrt{\mu_\omega}$$

In the case of $F \in [0, 0.25]$, the value of $cos(2\pi F)$ is positive. As in this case $1 + \mu_\omega - (\mu_{\phi_1} + \mu_{\phi_2}) < 0$ must be satisfied, the sign for $cos(2\pi F)\sqrt{\mu_\omega}$ should be negative. Therefore, in this case: $\mu_{\phi_1} + \mu_{\phi_2} = 1 + \mu_\omega - 2cos(2\pi F)\sqrt{\mu_\omega}$. In the case of $F \in [0.25, 0.5]$, the value of $cos(2\pi F)$ is negative. As in this case $1 + \mu_\omega - (\mu_{\phi_1} + \mu_{\phi_2}) > 0$ must be satisfied, the sign for $2cos(2\pi F)\sqrt{\mu_\omega}$ is positive. Therefore, in this case: $\mu_{\phi_1} + \mu_{\phi_2} = 1 + \mu_\omega - 2cos(2\pi F)\sqrt{\mu_\omega}$ as well. Hence, if both roots are complex then, for a given F and μ_ω:

$$\mu_{\phi_1} + \mu_{\phi_2} = 1 + \mu_\omega - 2cos(2\pi F)\sqrt{\mu_\omega} \tag{18}$$

that completes the proof. □

One should note that Eq. 18 is sufficient to find appropriate $\mu_{\phi_1} + \mu_{\phi_2}$ for any $F \in [0, 0.5]$ when $\mu_\omega > 0$, even on the boundary points ($F = 0$ and $F = 0.5$). According to this theorem, the value of $c = c_1 = c_2$ in IPSO can be calculated for a given F and $\omega > 0$ as:

$$c = 1 + \omega - 2cos(2\pi F)\sqrt{\omega} \tag{19}$$

We will use this equation later in section III-C to control the movement patterns of particles during the run.

B. Variance of movements (V_c)

The fixed point of the variance of positions for a particle with position update rule in Eq. 7 is calculated by Eq. 8. Hence, for IPSO (see section II-B for IPSO setting) and $c_1 = c_2 = c$, V_x is written as:

$$V_x = \frac{c(\omega + 1)[8(\sigma_p^2 + \sigma_g^2) + (\mu_g - \mu_p)^2]}{4(c(5\omega - 7) - 12\omega^2 + 12)} \tag{20}$$

Note that $c = c_1 = c_2$ imposes $c_1 >= 0$ and $c_2 >= 0$ as $c_1 + c_2 \geq 0$ is necessary for convergence of expectation (see Eq. 11). Equation 20 can be written as $V_x = V_c[8(\sigma_p^2 + \sigma_g^2) + (\mu_g - \mu_p)^2]$, where

$$V_c = \frac{c(\omega + 1)}{4(c(5\omega - 7) - 12\omega^2 + 12)} \tag{21}$$

V_c is called the *variance coefficient*. The main difference between V_x and V_c is that V_c is independent of the changes in p and g. Hence, one can set the range of the search through changing the value of V_c during the run. Of course the reflection of this in V_x also depends on $8(\sigma_p^2 + \sigma_g^2) + (\mu_g - \mu_p)^2$.

If $\mu_g = \mu_p$ and $\sigma_p = \sigma_g = 0$ then V_x becomes zero which means that the particle stops moving, no matter what the value for V_c is. However, as long as $\mu_g \neq \mu_p$ or $\sigma_p + \sigma_g \neq 0$, the variance of positions can be controlled by V_c. Fig. 5 shows the value of $ln(V_c)$ (natural logarithm) for different values of c and ω. Clearly, for a fixed ω, V_c is a monotonic function of c. However, this is not true for a fixed c, i.e., if c is constant then V_c is not a monotonic function of ω.

To achieve a desired V_c, one can solve Eq. 21 for c as follows:

$$c = \frac{-48V_c\omega^2 + 48V_c}{28V_c + \omega - 20V_c\omega + 1} \tag{22}$$

For a given V_c and ω, one can find c using this equation. We will use this equation later in section III-C to control the movement patterns of particles.

For parameter setting purposes, one should note that the changes of ω and c may have a large influence on

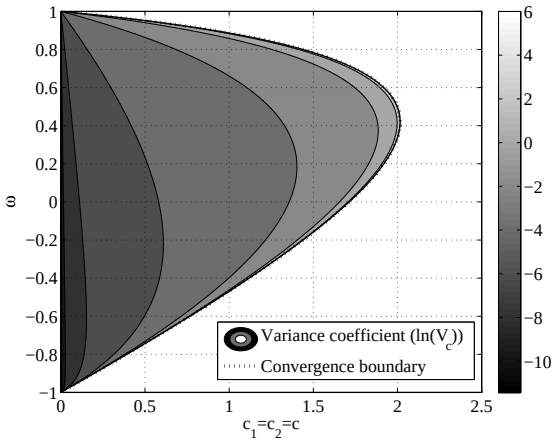

Fig. 5. The value of $ln(V_c)$ (natural logarithm). The darker the color is, the smaller the value of V_c is. The dotted line represents the convergence boundary.

the value of V_c, especially when these values are closer to the "edges" of the convergence boundary (Fig. 5). For example, for $\omega = 0.4$, the value of V_c is almost 0.77, 0.81, 3, and 23 when c is 0.043, 0.047, 1.97, and 2.01, respectively. This indicates that 2% changes in c imposes almost 90% changes on V_c when c is large, while 9% changes in c imposes almost 5% changes in V_c when c is small. This emphasizes the importance of the precision that need to be considered for parameter setting of PSO methods.

C. Calculation of coefficients for given V_c and F

For given F and V_c, one is able to calculate the values for c and $\omega > 0$ for IPSO by solving a system of equations that involves Eq. 18 and Eq. 22 (see Eq. 23).

$$\begin{cases} c &= 1 + \omega - 2cos(2\pi F)\sqrt{\omega} \\ c &= \frac{-48V_c\omega^2+48V_c}{28V_c+\omega-20V_c\omega+1} \end{cases} \tag{23}$$

The value of ω can be calculated by solving the following equation:

$$h_{F,V_c}(\omega) = \frac{-48V_c\omega^2 + 48V_c}{28V_c + \omega - 20V_c\omega + 1} - 1 - \omega + 2cos(2\pi F)\sqrt{\omega} = 0 \tag{24}$$

The value for c can be then calculated by substituting ω in one of the equations in Eq. 23.

In order to prove the existence of at least one solution (let us call it (ω_0, c_0)) for the system of equations in Eq. 23, we prove that ω_0 exists (the function $h_{F,V_c}(\omega)$ has at least one real root in $[0, 1]$ within its domain). The value of ω_0 is then substituted in one of the equations in Eq. 23 to calculate c_0. We first prove that $h_{F,V_c}(\omega)$ is continuous (Lemma 5) within the given domain ($\omega \in [0, 1]$ for any $V_c \geq 0$ and any $F \in [0, 0.5]$) and then we use the intermediate value theorem[2] to prove that this function has at least one root for any $V_c \geq 0.05$ and any $F \in [0, 0.5]$ (Theorem 2). We explain later why $V_c \in [0, 0.05)$ might result in no feasible combination of coefficients to achieve a given V_c and F.

Lemma 5. *For any $V_c \geq 0$, $F \in [0, 0.5]$, and $\omega \in [0, 1]$, $h_{F,V_c}(\omega)$ is continuous.*

Proof. We write $h_{F,V_c}(\omega) = h^1_{F,V_c}(\omega) + h^2_{F,V_c}(\omega)$ where $h^1_{F,V_c}(\omega) = \frac{-48V_c\omega^2+48V_c}{28V_c+\omega-20V_c\omega+1}$ and $h^2_{F,V_c}(\omega) = -1 - \omega + 2cos(2\pi F)\sqrt{\omega}$. We prove that both $h^1_{F,V_c}(\omega)$ and $h^2_{F,V_c}(\omega)$ are continuous in the given domain and, because adding two continuous functions results in a continuous function, it completes the proof for continuity of $h_{F,V_c}(\omega)$.

It is clear that $h^2_{F,V_c}(\omega)$ is continuous if $\omega \in [0, 1]$, hence, we focus on $h^1_{F,V_c}(\omega)$. $h^1_{F,V_c}(\omega)$ is continuous everywhere except when its denominator is zero ($28V_c + \omega - 20V_c\omega + 1 = 0$). This denominator is zero if and only if $V_c = \frac{\omega+1}{20\omega-28}$. The derivative of this formulation is $\frac{-48}{(20\omega-28)^2}$ that is always negative, meaning that it is monotonically decreasing in this domain. Thus, the maximum value of V_c for $\omega \in [0, 1]$ is found at $\omega = 0$ that results in $V_c = -1/28$ that is smaller than 0. Hence, there is no value of ω in $[0, 1]$ that causes $28V_c + \omega - 20V_c\omega + 1 = 0$ for any $V_c \geq 0$, that means $h^1_{F,V_c}(\omega)$ is continuous in this domain. Therefore, both $h^1_{F,V_c}(\omega)$ and $h^2_{F,V_c}(\omega)$ are continuous for any $V_c \geq 0$, $F \in \mathbb{R}$, and $\omega \in [0, 1]$, that means $h_{F,V_c}(\omega)$ is continuous in this domain. □

We use this lemma to prove the existence of at least one real solution for the systems of equations in Eq. 23.

Theorem 2. *For any $V_c \geq 0.05$ and $F \in [0, 0.5]$, there exists at least one solution (ω_0, c_0) for the system of equations in Eq. 23 where $\omega_0 \in [0, 1]$ and the pair (ω_0, c_0) guarantees convergence of variance.*

Proof. We prove that $h_{F,V_c}(\omega) = 0$ has at least one real solution $\omega_0 \in [0, 1]$. According to Lemma 5, $h_{F,V_c}(\omega)$ is continuous in the domain $\omega \in [0, 1]$ for any $V_c \geq 0.05$ and any F. For any V_c and $\omega = 1$, It is clear that $h_{F,V_c}(1) = 2cos(2\pi F) - 2$ which guarantees $h_{F,V_c}(1) \leq 0$. Also, we calculate $h_{F,V_c}(0)$ as follows

$$h_{F,V_c}(0) = \frac{48V_c}{28V_c + 1} - 1 \tag{25}$$

It is clear that $h_{F,V_c}(0) \geq 0$ for any F and any $V_c \geq 0.05$. This means that, for any F and $V_c \geq 0.05$, there is at least one $\omega \in [0, 1]$ that causes $h_{F,V_c}(\omega)$ to change its sign from negative to positive, hence, based on the intermediate value theorem, $h_{F,V_c}(\omega)$ has at least one real root in that domain for any F and $V_c \geq 0.05$.

As $h_{F,V_c}(\omega) = 0$ has at least one real root in $[0, 1]$, this solution can be used to calculate c in Eq. 23, hence, the

[2]If f is continuous on a closed interval $[a, b]$ and $c \in [f(a), f(b)]$ then there exists $x \in [a, b]$ such that $f(x) = c$. In a special case, if we assume that $f(a)f(b) < 0$ then there exists an x in the interval $[a, b]$ that $f(x) = c$, i.e., f has a root in the interval $[a, b]$. This special case is known as the Bolzano theorem[16].

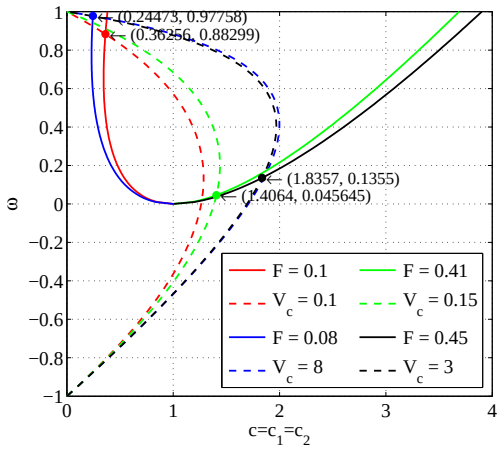

Fig. 6. The values of (c, ω) when $(F, V_c) \in \{(0.1, 0.1), (0.08, 8.0), (0.41, 0.15), (0.45, 3.0)\}$.

found value for ω and c will guarantee convergence of variance, that completes the proof. □

The reason why we only considered the cases where $V_c \geq 0.05$ is that $h_{F,V_c}(0)$ (Eq. 25) is negative for $0 \leq V_c < 0.05$, hence, with this setting, the intermediate value theorem does not provide any information about the existence of real roots for $h_{F,V_c}(\omega)$ (Eq. 24). Intuitively, small values for V_c are achievable only if c is small (see Fig. 5). However, if c is small then some large values of F are not achievable for any ω (see Fig. 4). This means that choosing small values for V_c might result in no solution for Eq. 24.

One can simplify Eq. 24 as:

$$\hat{h}_{F,V_c}(\omega) = -(28V_c + 1)b^4 + 2(1 - 20F'V_c)b^3 \quad (26)$$
$$-2(4V_c + 1)b^2 + 2F'(28V_c + 1)b + 20V_c - 1 = 0$$

where $\omega = b^2$ and $F' = cos(2\pi F)$. This equation is polynomial and can be solved by many different methods (we use the method proposed in [17] to solve this equation and find (ω_0, c_0) in this paper). Note that, according to Theorem 2, $h_{F,V_c}(\omega)$ and, consequently, $\hat{h}_{F,V_c}(\omega)$, have at least one real root in $[0, 1]$ when $V_c \geq 0.05$ and $F \in [0, 0.5]$.

Fig. 6 shows the value of ω and c for four different pairs of frequencies and variance coefficients, $(F, V_c) \in \{(0.1, 0.1), (0.08, 8.0), (0.41, 0.15), (0.45, 3.0)\}$. For each case, the curves corresponding to F and V_c have been shown (in c vs ω space). The crossing points of these curves are solutions for the system of equations given by Eq. 23.

D. F and correlation between positions

We experimentally investigate the relationship between F and the correlation between positions (the Pearson correlation between x_t and x_{t+1} for any t). We set $V_c \in \{0.5, 1.5, 2, 2.5, ..., 25\}$ and, for each V_c and $F \in \{0.01, 0.02, ..., 0.49\}$, we calculate c and ω by solving

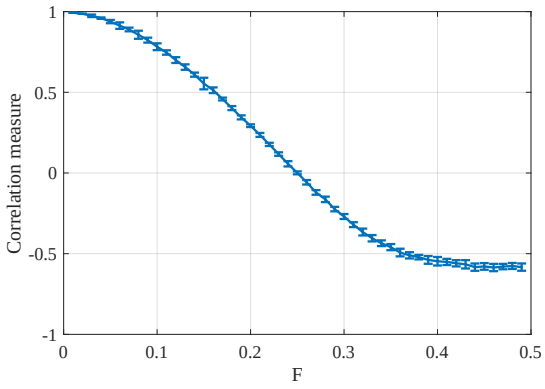

Fig. 7. The vertical axis shows the correlation measure and the horizontal axis shows F. The graph shows the average and standard deviation of the correlation measure for different values of V_c. The figure indicates that V_c does not influence the relationship between F and the correlation measure. The value of p and g were set to 1 and 5 respectively. However, our experiments showed that any other values for p and g result in the same relationship between F and the correlation coefficient.

the system of equations in Eq. 23. For each pair (ω, c), we simulate one particle of IPSO (p and g constants), 50 times, for 1,000 iterations and calculate the Pearson correlation coefficient between positions in these runs (i.e., the correlation matrix $Corr$ that is 1000×1000 where $Corr_{i,j}$ shows the correlation between positions at iteration i and iteration j estimated through 50 runs). We calculate the *correlation measure* as the average of the correlation coefficient between each position and its previous one in those runs (i.e., correlation measure $= \frac{\sum_{i=2}^{1000} Corr_{i,i-1}}{1000}$). If the positions are independent from one another (see Fig. 7) then the correlation measure is close to zero.

Fig. 7 indicates that the correlation measure is equal to zero (positions are mostly independent) when $F = 0.25$. However, the positions are positively correlated when $F < 0.25$ and negatively correlated when $F > 0.25$. Hence, for smaller F (positive correlation), the particle moves smoothly and for larger F (negative correlation) particle jumps from one place to another. Also, the most random-like movement takes place at $F = 0.25$, i.e., particle position at each iteration is independent of the position at the previous iteration. All these conclusions are in fact independent of the value of V_c, as Fig. 7 indicates (V_c applies small variance to the correlation measure indicating that the relationship between F and the correlation measure is independent of V_c).

E. V_c and the range of search

The variance of movement is closely related to the range that a particle covers during the search, i.e., for a fixed distribution, a larger range is covered when the variance is larger. Hence, one can use the variance as a measure for the range of the search by a particle.

In the case of IPSO with $c_1 = c_2$, because the expectation of x_t is between p and g (recall that $c_1 \geq 0$ and $c_2 \geq 0$), the particle's position (x_t) oscillates between p and g with the variance $V_c(g-p)^2$. If V_c is too small then, depending on the distribution of the positions[3], x_t, the particle may not sample any point outside the boundaries of $[min(p,g), max(p,g)]$. However, from the search perspective, there is no reason to expect any high quality solution between p and g. Hence, a small V_c may result in an inefficient search. It is therefore desirable to guarantee that the positions of particle have a chance to go beyond the interval $[E_x - P_1, E_x + P_2]$, where $P_1 = min\{p,g\}$ and $P_2 = max\{p,g\}$.

Theorem 3. *For any distribution of x_t generated by the recursion in Eq. 7, given that $g \neq p$, if*

$$V_c > \left(\frac{max(\mu_{\phi_1}, \mu_{\phi_2})}{\mu_{\phi_1} + \mu_{\phi_2}} \right)^2 \tag{27}$$

then the number of points generated by x_t outside the interval $(E_x - P_1, E_x + P_2)$ is non-zero where $P_1 = min\{p,g\}$ and $P_2 = max\{p,g\}$.

Proof. It is trivial to show that, for any distribution with the expectation μ and the standard deviation σ, the number of points randomly generated by that distribution outside of the interval $(\mu - \sigma, \mu + \sigma)$ is larger than zero. Hence, if we guarantee that $E_x + \sqrt{V_x} > max\{p,g\}$ and $E_x - \sqrt{V_x} < min\{p,g\}$ then there is a guarantee that the number of points outside $(E_x - P_1, E_x + P_2)$ is non-zero.

If $\mu_{\phi_1} > 0$ and $\mu_{\phi_2} > 0$ then $min\{p,g\} < E_x < max\{p,g\}$. Let us assume that $g > p$. In this case, if $E_x + \sqrt{V_x} > g$ then $\frac{\mu_{\phi_1} p + \mu_{\phi_2} g}{\mu_{\phi_1} + \mu_{\phi_2}} + \sqrt{V_c}(g-p) > g$ that leads to $\sqrt{V_c} > \frac{\mu_{\phi_1}}{\mu_{\phi_1} + \mu_{\phi_2}}$. Also, if $E_x - \sqrt{V_x} < p$ then $\frac{\mu_{\phi_1} p + \mu_{\phi_2} g}{\mu_{\phi_1} + \mu_{\phi_2}} - \sqrt{V_c}(g-p) < p$ that leads to $\sqrt{V_c} > \frac{\mu_{\phi_2}}{\mu_{\phi_1} + \mu_{\phi_2}}$. The same analogy can be followed when $p > g$ that results in the same outcome.

Hence, if $\sqrt{V_c} > \frac{max(\mu_{\phi_1}, \mu_{\phi_2})}{\mu_{\phi_1} + \mu_{\phi_2}}$ then $E_x + \sqrt{V_x} > max\{p,g\}$ and $E_x - \sqrt{V_x} < min\{p,g\}$ are guaranteed that, in fact, guarantees that the number of points generated by x_t outside the interval $(E_x - P_1, E_x + P_2)$ is non-zero. □

This theorem enables practitioners to prevent situations where the particles only search a small area between p and g. The criteria found in Theorem 3 is independent of the distribution of x_t and other parameters in Eq. 7. Specific distribution of x_t is needed to calculate a more accurate measure of the number of points generated outside of the interval $(E_x - P_1, E_x + P_2)$. Also, for specific distributions for x_t (e.g., normal), if known, there might be better lower bounds to ensure the percentage of points that are outside $(E_x - P_1, E_x + P_2)$ is larger than zero. It is easy to see that if more points are sampled inside $(E_x - P_1, E_x + P_2)$ then more points need to be sampled outside of this interval to ensure that the variance remains constant. For $\mu_{\phi_1} = \mu_{\phi_2}$, ensuring $V_c > 0.25$ guarantees that the particle positions move beyond $(E_x - P_1, E_x + P_2)$, no matter the distribution of the x_t. Hence, for IPSO, $V_c > 0.25$ guarantees searching outside of $(E_x - P_1, E_x + P_2)$. We remind that the necessary condition for this theorem is that $p \neq g$. To guarantee that $p \neq g$ one can add some perturbation to p and g, that in turn may make the algorithm locally convergent, see [19], [14], [4] for details.

[3]The exact distribution of x_t depends on the distribution of p, g, c_1, c_2, and ω and, to the best of our knowledge, it is unknown. There are experimental evidence, however, that show this distribution is bell-shaped (claimed to be normal in [18]) for IPSO settings. Although the experiments show that this distribution is bell-shaped, no proper proof has been provided to show whether it is normal.

IV. Experiments

The aim of this section is to understand which values for frequency and variance have better performance on a set of benchmark test functions. We introduce the benchmark we use for the tests and discuss how we conduct the comparisons. Then, we apply IPSO with different coefficients to these benchmarks and investigate the effects of F and V_c on the performance of the algorithm.

A. Benchmarks and comparison strategy

We use the CEC2014 benchmark set [20] that contains 30 optimization functions (for all functions, each dimension is bounded in $[-100, 100]$) to test the performance of different methods (the term "methods" refers to IPSO algorithms with different coefficient values in this section). In analyses we focus on 10 dimensional problems and we set the swarm size to 20. We use both 10 and 30 dimensional problems for comparison between our proposed coefficients combinations and existing ones. Also, we run each method 96 times in experiments.

We use the Wilcoxon test to find if the performance difference between two method on a particular function k is significant (the confidence bound $p \leq 0.05$). We define $s_{i,j,k}$ by

$$s_{i,j,k} = \begin{cases} 1 & \text{if } R_{i,k} < R_{j,k} \text{ and } p \leq 0.05 \\ 0 & p > 0.05 \\ -1 & \text{if } R_{j,k} < R_{i,k} \text{ and } p \leq 0.05 \end{cases}$$

where p is the confidence bound of the Wilcoxon test that compares methods i and j, and $R_{i,k}$ is the median of the results of the method i when it is applied to the function k. We assign a "point" M_i to each method i as follows:

$$M_i = \sum_{j=1}^{m} z_{i,j} \tag{28}$$

where $z_{i,j}$ is defined by

$$z_{i,j} = \begin{cases} a_1 & \text{if } \sum_{k=1}^{n} s_{i,j,k} > b \\ a_2 & \text{if } |\sum_{k=1}^{n} s_{i,j,k}| \leq b \\ 0 & \text{if } \sum_{k=1}^{n} s_{i,j,k} < -b \end{cases}$$

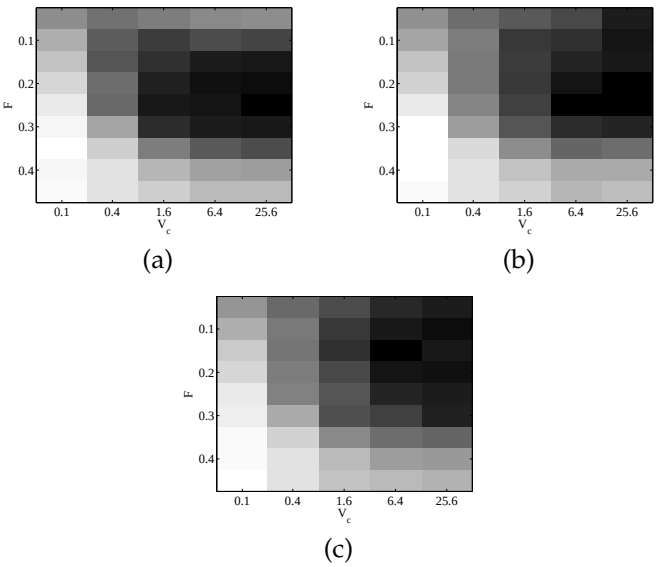

Fig. 8. The darker the square is, the better the performance of that combination of (V_c, F) (a) short run, (b) long run, (c) very long run.

In this paper, we set $a_1 = 3$, $a_2 = 1$, and $b = 0$, unless stated. Note that $\sum_{k=1}^{n} s_{i,j,k} > b$ indicates that the number of functions for which method i performs significantly better, according to the Wilcoxon test, than method j is larger than the number of test functions for which method j performs significantly better than method i by at least $b + 1$.

The best method is determined by sorting the vector M in descending order. In the case of a tie between two methods ($M_i = M_j$), the value of $S_i = \sum_{j=1}^{m} \sum_{k=1}^{n} s_{i,j,k}$ is used to determine which method is better. The value of S_i is the total *score* the method i has received against all other tested methods.

B. Effects of F and V_c on the search performance

In order to test the effect of V_c and F on the performance of IPSO, we set $V_c \in \{0.1, 0.4, 1.6, 6.4, 25.6\}$ and $F \in \{0.05, 0.1, 0.15, 0.2, 0.25, 0.3, 0.35, 0.4, 0.45\}$ and calculate ω and c for each pair of (V_c, F) by solving Eq. 23. We compare IPSO with these different coefficients (5×9 coefficient pairs) using the comparison strategy introduced in section IV-A. We conduct three sets of tests, one with a small number of function evaluations ($2500 \times d$ function evaluations, we call it short run), one with a larger number of function evaluations ($25000 \times d$ function evaluations, we call it long run), and another with even larger number of function evaluations ($50000 \times d$ function evaluations, we call it very long run). The value of M for each combination of (V_c, F) has been shown in Fig. 8.

In a short run, it is clear that the coefficients corresponding with the medium to large V_c and medium F have better performances. This indicates that, in a short run, a chaotic search (medium F) that covers a large area (large variance) has a better performance. When the number of iterations increases, the performance of the methods with coefficients corresponding with smaller F are improved (no significant changes of the best choices for V_c comparing to the short runs is observed). The reason is that, in a longer run, particles need to conduct an "accurate" search around existing high quality solutions at the latter stages of the search. When F is small, x_t is close to x_{t+1} at each iteration t that results in a more "accurate" search. Hence, we expect that, if such accurate search is needed then a coefficient combination that imposes small F shows a good performance. Note that, in all cases (short, long, and very long), the coefficients that impose small variance have a very poor performance as they cannot cover sufficiently large areas, specially at the early stages of the search that such behavior is needed.

TABLE I
TOP THREE COEFFICIENTS COMBINATIONS FOR SHORT, LONG, AND VERY LONG RUNS. RANK SHOWS THE RANKING OF IPSO THAT USES THOSE COMBINATIONS FOUND BY THE STRATEGGY EXPLAINED IN SECTION IV-A.

Run	Rank	ω	c	V_c	F
Short	1	0.711897	1.711897	25.6	0.25
	2	0.836416	1.271188	25.6	0.2
	3	0.832043	1.268295	6.4	0.2
Long	1	0.836416	1.271188	25.6	0.2
	2	0.704772	1.704772	6.4	0.25
	3	0.711897	1.711897	25.6	0.25
Very long	1	0.913799	0.790038	6.4	0.15
	2	0.96506	0.375544	25.6	0.1
	3	0.836416	1.271188	25.6	0.2

Table I shows the top three combination of coefficients in short, long, and very long runs. The results of IPSO when it uses these coefficients is compared with the results of the algorithm when it uses well-known existing coefficients in the next section.

One should note that in all experiments in this paper we use the CEC2014 benchmark problems for which each dimension of the search space is bounded in $[-100, 100]$. Therefore, to solve a new problem with different boundaries, a standardization procedure might be needed to map the boundaries of that problem to $[-100, 100]$. On potential mapping is given by $y_i = \frac{u_i - l_i}{200}(x_i + 100) + l_i$ where x_i is the i^{th} dimension of $\vec{x}$ (the solution candidate in $[-100, 100]^d$), y_i is the i^{th} dimension of the mapped vector, and l_i and u_i are the search space boundaries of the i^{th} dimension of the problem.

C. Comparison with well-known coefficients

We compare the results of the top three combinations of coefficients in short run, long run, and very long run (see Table I) with those of other coefficients for IPSO in literature. The value of F and V_c for each combination (proposed in this paper and taken from earlier research) have been shown in Table II. The value of F for most existing coefficients (A_1 to A_5) sets is around 0.2 and 0.25. This means that, most experimental approaches have picked coefficients combinations in a way that the

TABLE II
FREQUENTLY USED COEFFICIENTS FOR IPSO TOGETHER WITH PROPOSED COEFFICIENTS. FOR EACH COEFFICIENT PAIRS THE VALUES OF F AND V_c HAVE BEEN ALSO REPORTED.

Proposed in	Set	F	V_c	c	ω
[21]	A_1	0.22803	1.0581	1.494	0.729
[13]	A_2	0.26028	0.77273	1.7	0.6
[22]	A_3	0.24859	17.0298	1.7	0.715
[23]	A_4	0.19968	0.29988	1.193	0.721
[11]	A_5	0.26599	0.24047	1.55	0.42
This paper	A_6	0.25	25.6	1.711897	0.711897
This paper	A_7	0.2	25.6	1.271188	0.836416
This paper	A_8	0.2	6.4	1.268295	0.832043
This paper	A_9	0.25	6.4	1.704772	0.704772
This paper	A_{10}	0.1	25.6	0.375544	0.965060
This paper	A_{11}	0.15	6.4	0.913799	0.790038

positions generated by each particle is almost random. Also, the value for V_c in these settings is not that large except for the set A_3 that was intentionally set to impose a large variance. One potential reason behind the lack of coefficients to impose large V_c is that most coefficients have been obtained by some experiments in which the coefficient values are changed by a step size. However, as V_c is very sensitive to the changes of the values of coefficients (see section III-B), setting this step sizes to a relatively large value could skip coefficient sets that impose large variances. In addition, the parameter settings for IPSO are usually conducted for specific benchmark functions and, because the variance of movement also depends on g and p, the picked functions can play a very important role on the best found combination. As an example, if the function is smooth then p and g become very close to each other at the early stages of the search that leads to vanishing the effect of V_c at the early stages of the search. Hence, the impact of the coefficients on the variance of movement is not observed.

We test IPSO with the coefficients in Table II using the benchmark functions and comparison strategy introduced in section IV-A. We run the methods for $2500 \times d$ function evaluations (short run), $25000 \times d$ function evaluations (long run), and , $50000 \times d$ function evaluations (very long run). Table III shows that, for 10 dimensional test cases, the combination A_7 has the best performance for short runs, A_8 for long runs, and A_{10} for very long runs. Clearly, the performance of low frequency setting (A_{10}) has become better when the number of function evaluations increases. Also, the setting A_6 results in the best performance in general (third place in all tests). The coefficients combination A_5 shows very poor results comparing to the others. One explanation for this poor performance is that the value of V_c that this combination imposes is small that may result in inefficient search in the search space (see theorem 3). See section VI, Appendix, for detailed results related to these comparisons.

We also test the performance of IPSO with the coefficients in Table II on 30 dimensional problems. Results indicate almost the same conclusions as what was drawn in 10 dimensional experiments.

TABLE III
EXPERIMENTAL RESULTS FOR 10 DIMENSIONAL PROBLEMS. THE RANK OF EACH PARAMETER SET FOR DIFFERENT TESTS (SHORT, LONG, AND VERY LONG RUNS) HAS BEEN INDICATED IN PARENTHESES IN THE "POINTS" COLUMN. THE "SCORE" COLUMN INDICATES THE SCORE VALUE CALCULATED THROUGH THE PROCEDURE DESCRIBED IN SECTION IV-A.

Set	Short run		Long run		Very long run	
	Points	Scores	Points	Scores	Points	Scores
A_1	10 (8)	-22	10 (8)	-54	10 (8)	-63
A_2	7 (9)	-47	7 (9)	-67	7 (9)	-61
A_3	29 (2)	96	26 (2)	94	16 (7)	63
A_4	4 (10)	-168	4 (10)	-212	4 (10)	-205
A_5	1 (11)	-263	1 (11)	-262	1 (11)	-283
A_6	25 (3)	91	24 (3)	95	22 (3)	91
A_7	**29 (1)**	**101**	19 (5)	80	24 (2)	103
A_8	20 (4)	83	**26 (1)**	**95**	20 (5)	84
A_9	20 (5)	76	20 (4)	87	21 (4)	84
A_{10}	13 (7)	-12	17 (7)	70	**29 (1)**	**103**
A_{11}	16 (6)	65	19 (6)	74	16 (6)	84

TABLE IV
EXPERIMENTAL RESULTS FOR 30 DIMENSIONAL PROBLEMS. SEE TABLE III FOR EXPLANATION OF "POINTS" AND "SCORE" COLUMNS.

Set	Short run		Long run		Very long run	
	Points	Scores	Points	Scores	Points	Scores
A_1	16 (6)	48	13 (7)	29	13 (7)	-15
A_2	10 (8)	-29	7 (9)	-95	7 (9)	-102
A_3	**31 (1)**	**133**	26 (3)	136	10 (8)	-76
A_4	4 (10)	-204	4 (10)	-216	4 (10)	-215
A_5	1 (11)	-278	1 (11)	-288	1 (11)	-287
A_6	28 (2)	126	**29 (1)**	**140**	26 (2)	143
A_7	22 (4)	99	22 (4)	119	20 (4)	117
A_8	19 (5)	75	19 (5)	98	20 (5)	110
A_9	25 (3)	122	27 (2)	140	26 (3)	134
A_{10}	7 (9)	-87	10 (8)	-86	**31 (1)**	**164**
A_{11}	13 (7)	-5	16 (6)	23	16 (6)	27

V. CONCLUSION AND FUTURE WORKS

In this paper we studied two factors to characterize movement patterns for a particle in PSO, namely the range of movement and the correlation between positions generated by the particle. We focused our investigations on these two factors to find how they impact the performance of the method. We then showed that base frequency, F, can formulate the correlation between positions generated by particles ($F = 0.25$ causes zero correlation between positions, $F < 0.25$ causes positive and $F > 0.25$ causes negative correlation between positions) and the variance of movement, V_c, can formulate the range of movement. We found that these factors, especially V_c, are very sensitive to the changes of the coefficients values. We provided formulas to find appropriate acceleration coefficients to achieve a given F (Eq. 18) or V_c (Eq. 21) for a given ω. We also provided a system of equations (Eq. 23) to find the values of c and ω for a given V_c and F and we proved that this system of equations has at least one solution for any $V_c \geq 0.05$, F, with $\omega \in [0, 1]$ that guarantee convergence of variance. Hence, by setting the coefficients to guarantee a given V_c and F, one can

control the movement pattern of particles during the run. We experimentally tested the impact of V_c and F on the search and we found that smaller values for F are more effective when the number of iterations is large. These experiments also showed that a large V_c is usually beneficial for any number of iterations. We compared the values of V_c and F for several frequently used coefficient values in a particular type of PSO and we found that all of these coefficient values encourage chaotic movement (no correlation between consecutive positions). Also, we found new sets of coefficients for IPSO that result in better performance on the CEC2014 benchmark (Table III) for different number of iterations (short run with $2500 \times d$ function evaluations, long run with $25000 \times d$ function evaluations, and very long run with $50000 \times d$ function evaluations). We found that $c = 1.271188$ and $\omega = 0.836416$ ($F = 0.2$ and $V_c = 25.6$), $c = 1.268295$ and $\omega = 0.832043$ ($F = 0.2$ and $V_c = 6.4$), and $c = 0.375544$ and $\omega = 0.965060$ ($F = 0.1$ and $V_c = 25.6$), are the best coefficient sets for a short run, long run, and a very long run, respectively when $d = 10$. For d=30, the best found coefficients were $c = 1.7$ and $\omega = 0.715$ ($F = 0.2486$ and $V_c = 17.0298$), $c = 1.711897$ and $\omega = 0.711897$ ($F = 0.25$ and $V_c = 25.6$), and $c = 0.375544$ and $\omega = 0.965060$ ($F = 0.1$ and $V_c = 25.6$) for short, long, and very long runs. Overall, we found that the combination $c = 1.711897$ and $\omega = 0.711897$ ($F = 0.25$ and $V_c = 25.6$) performs better than other coefficients in all cases (10 and 30 dimensional for short, long, and very long runs).

Finding the relationship between F and V_c with the characteristics of the landscape is one potential direction for further research. In this paper we only focused on the effects of these factors on the performance of the method with regards to the length of the run on a set of functions. However, different values for these factors may change the performance of the method on landscapes with different characteristics (see [24] and [25]). Designing a method to find the best F and V_c at each iteration and use Eq. 23 to find the coefficient values and change these coefficients to maximize the search performance of the algorithm can be also another promising research direction. In addition, this paper investigated the relationship between frequency of movement and the correlation between consecutive generated positions. However, it might be worthwhile to study the correlation between each generated position and previous $j > 1$ positions (autocorrelation) to understand whether the patterns of movement can be characterized by those information.

VI. Appendix

This section reports detailed comparison results between methods (recall that by methods we ref to different parameter values for IPSO) when they are applied to 10 dimensional CEC2014 benchmark functions (30 functions). Three tables are listed in this section, each table shows the results for experiments that were different based on the maximum allowed number of iterations (short run: $2500 \times d$ function evaluations, long run: $25000 \times d$ function evaluations, and very long run: $50000 \times d$ function evaluations). For each table, a value x in the row "w" for a method A_i in a specific column indicates that A_i performs significantly (based on the Wilcoxon test) better than the other method (specified in the column) in x number of functions (over 30 functions). A value x' in the row "l" for a method A_i in a specific column indicates that A_i performs significantly worse than the other method (specified in the column) in x' number of functions. The total number of functions was 30, hence, the number of functions that the methods were statistically similar can be calculated by $30-w-l$. In fact, for a method A_i in comparison to another method A_j, $\sum_{k=1}^{30} s_{i,j,k} = w_{A_i,A_j} - l_{A_i,A_j}$.

TABLE V
This table lists the results for the "short" run.

		A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}
A_1	l	0	1	12	1	1	15	13	13	11	10	12
	w	0	5	1	22	26	1	1	1	1	7	2
A_2	l	5	0	17	0	1	17	17	15	15	9	16
	w	1	0	1	20	27	1	2	1	1	7	4
A_3	l	1	1	0	0	0	0	1	0	0	3	2
	w	12	17	0	24	29	1	1	2	2	11	5
A_4	l	22	20	24	0	0	23	23	23	22	19	20
	w	1	0	0	0	23	0	0	0	0	3	1
A_5	l	26	27	29	23	0	28	28	28	28	26	25
	w	1	1	0	0	0	0	0	0	0	2	1
A_6	l	1	1	1	0	0	0	3	0	0	1	1
	w	15	17	0	23	28	0	0	1	1	10	4
A_7	l	1	2	1	0	0	0	0	0	2	1	1
	w	13	17	1	23	28	3	0	1	5	15	3
A_8	l	1	1	2	0	0	1	1	0	2	2	2
	w	13	15	0	23	28	0	0	0	2	11	3
A_9	l	1	1	2	0	0	1	5	2	0	3	2
	w	11	15	0	22	28	0	2	2	0	9	4
A_{10}	l	7	7	11	3	2	10	15	11	9	0	12
	w	10	9	3	19	26	1	1	2	3	0	1
A_{11}	l	2	4	5	1	1	4	3	3	4	1	0
	w	12	16	2	20	25	1	1	2	2	12	0

VII. Acknowledgment

This work was partially funded by the ARC Discovery Grant DP130104395.

References

[1] J. Kennedy and R. Eberhart, "Particle swarm optimization," in *International Conference on Neural Networks*, vol. 4. IEEE, 1995, pp. 1942–1948.
[2] A. P. Engelbrecht, *Fundamentals of computational swarm intelligence*. John Wiley & Sons, 2006.
[3] R. Poli, "Analysis of the publications on the applications of particle swarm optimisation," *Journal of Artificial Evolution and Application*, vol. 2008, no. 3, pp. 1–10, 2008. [Online]. Available: http://www.hindawi.com/archive/2008/685175/
[4] M. R. Bonyadi and Z. Michalewicz, "Particle swarm optimization for single objective continuous space problems: a review," *Evolutionary computation*, 2016.
[5] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *International Symposium on Micro Machine and Human Science*. IEEE, 1995, pp. 39–43.
[6] M. Clerc, *Particle swarm optimization*. Wiley-ISTE, 2006.

TABLE VI
THIS TABLE LISTS THE RESULTS FOR THE "LONG" RUN.

		A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}
A_1	l	0	2	19	0	1	18	19	20	15	17	17
	w	0	6	2	26	26	2	3	1	1	5	2
A_2	l	6	0	20	0	1	20	21	21	20	17	19
	w	2	0	3	26	27	2	3	2	3	5	5
A_3	l	2	3	0	0	1	1	2	4	1	7	2
	w	19	20	0	28	29	1	3	2	2	8	5
A_4	l	26	26	28	0	2	27	26	27	29	25	26
	w	0	0	0	0	25	0	1	1	0	2	1
A_5	l	26	27	29	25	0	28	28	28	28	27	28
	w	1	1	1	2	0	1	1	1	1	2	1
A_6	l	2	2	1	0	1	0	1	0	2	6	1
	w	18	20	1	27	28	0	4	3	1	6	3
A_7	l	3	3	3	1	1	4	0	3	3	3	1
	w	19	21	2	26	28	1	0	0	4	4	0
A_8	l	1	2	2	1	1	3	0	0	3	3	1
	w	20	21	4	27	28	0	3	0	3	4	2
A_9	l	1	3	2	0	1	1	4	3	0	10	4
	w	15	20	1	29	28	2	3	3	0	7	8
A_{10}	l	5	5	8	2	2	6	4	4	7	0	4
	w	17	17	7	25	27	6	3	3	10	0	2
A_{11}	l	2	5	5	1	1	3	0	2	8	2	0
	w	17	19	2	26	28	1	1	1	4	4	0

TABLE VII
THIS TABLE LISTS THE RESULTS FOR THE "VERY LONG" RUN.

		A_1	A_2	A_3	A_4	A_5	A_6	A_7	A_8	A_9	A_{10}	A_{11}
A_1	l	0	2	18	1	1	20	21	18	19	18	19
	w	0	5	3	26	27	1	2	3	1	4	2
A_2	l	5	0	18	1	0	19	21	21	20	21	21
	w	2	0	6	25	29	4	5	3	4	5	3
A_3	l	3	6	0	2	1	1	7	2	3	9	2
	w	18	18	0	27	28	0	1	1	0	3	3
A_4	l	26	25	27	0	1	28	27	27	28	27	27
	w	1	1	2	0	27	1	1	1	1	1	2
A_5	l	27	29	28	27	0	30	29	30	29	29	29
	w	1	0	1	1	0	0	0	0	1	0	0
A_6	l	1	4	0	1	0	0	4	1	1	7	4
	w	20	19	1	28	30	0	4	2	1	5	4
A_7	l	2	5	1	1	0	4	0	2	2	5	2
	w	21	21	7	27	29	4	0	1	8	5	4
A_8	l	3	3	1	1	0	2	1	0	4	7	2
	w	18	21	2	27	30	1	2	0	0	5	2
A_9	l	1	4	0	1	1	1	8	0	0	10	4
	w	19	20	3	28	29	1	2	4	0	4	4
A_{10}	l	4	5	3	1	0	5	5	5	4	0	3
	w	18	21	9	27	29	7	5	7	10	0	5
A_{11}	l	2	3	3	2	0	4	4	2	4	5	0
	w	19	21	2	27	29	4	2	2	4	3	0

[7] M. A. Montes de Oca, T. Stützle, M. Birattari, and M. Dorigo, "Frankenstein's pso: A composite particle swarm optimization algorithm," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 5, pp. 1120–1132, 2009.

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

[9] M. Jiang, Y. P. Luo, and S. Y. Yang, "Stochastic convergence analysis and parameter selection of the standard particle swarm optimization algorithm," *Information Processing Letters*, vol. 102, no. 1, pp. 8–16, 2007.

[10] R. Poli, "Mean and variance of the sampling distribution of particle swarm optimizers during stagnation," *IEEE Transactions on Evolutionary Computation*, vol. 13, no. 4, pp. 712–721, 2009.

[11] Q. Liu, "Order-2 stability analysis of particle swarm optimization," *Evolutionary computation*, 2014.

[12] M. Bonyadi and Z. Michalewicz, "Stability analysis of the particle swarm optimization without stagnation assumption," *IEEE Transactions on Evolutionary Computation*, vol. to appear, 2015.

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

[14] M. Bonyadi and Z. Michalewicz, "Analysis of stability, local convergence, and transformation sensitivity of a variant of particle swarm optimization algorithm," *IEEE Transactions on Evolutionary Computation*, vol. to appear, 2015.

[15] E. Ozcan and C. K. Mohan, "Particle swarm optimization: surfing the waves," in *Congress on Evolutionary Computation*, vol. 3. IEEE, 1999.

[16] B. Bolzano, "Purely analytic proof of the theorem that between any two values which give results of opposite sign there lies at least one real root of the equation. translated by s. russ," *Ewald (1996)*, vol. 1, pp. 225–48, 1817.

[17] A. Edelman and H. Murakami, "Polynomial roots from companion matrix eigenvalues," *Mathematics of Computation*, vol. 64, no. 210, pp. 763–776, 1995.

[18] J. Kennedy, "Bare bones particle swarms," in *Swarm Intelligence Symposium*, 2003, pp. 80–87.

[19] M. R. Bonyadi and Z. Michalewicz, "A locally convergent rotationally invariant particle swarm optimization algorithm," *Swarm Intelligence*, vol. 8, no. 3, pp. 159–198, 2014.

[20] J. Liang, B. Qu, and P. Suganthan, "Problem definitions and evaluation criteria for the cec 2014 special session and competition on single objective real-parameter numerical optimization," *Computational Intelligence Laboratory, Zhengzhou University, Zhengzhou China and Technical Report, Nanyang Technological University, Singapore*, 2013.

[21] 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.

[22] M. Jiang, Y. Luo, and S. Yang, "Particle swarm optimization-stochastic trajectory analysis and parameter selection," in *Swarm Intelligence, Focus on Ant and Particle Swarm Optimization*, F. Chan and M. KumarTiwari, Eds. I-TECH Education and Publishing, Wien, 2007, book section 11, pp. 179–198.

[23] D. Bratton and J. Kennedy, "Defining a standard for particle swarm optimization," in *Swarm Intelligence Symposium*. IEEE, 2007, pp. 120–127.

[24] K. M. Malan and A. P. Engelbrecht, "A survey of techniques for characterising fitness landscapes and some possible ways forward," *Information Sciences*, vol. 241, pp. 148–163, 2013.

[25] K. M. Malan, "Characterising continuous optimisation problems for particle swarm optimisation performance prediction," Ph.D. dissertation, University of Pretoria, 2014.