Particle Swarm Optimization: Stability Analysis using N-Informers under Arbitrary Coefficient Distributions

Christopher W. Cleghorn[a,*], Belinda Stapelberg[b]

[a]School of Computer Science and Applied Mathematics, University of the Witwatersrand, Johannesburg, South Africa
[b]Department of Mathematics and Applied Mathematics, University of Pretoria, Pretoria, South Africa

ARTICLE INFO

Keywords:
Particle Swarm Optimization
Stability Analysis
Stability Criteria

ABSTRACT

This paper derives, under minimal modelling assumptions, a simple to use theorem for obtaining both order-1 and order-2 stability criteria for a common class of particle swarm optimization (PSO) variants. Specifically, PSO variants that can be rewritten as a finite sum of stochastically weighted difference vectors between a particle's position and swarm informers are covered by the theorem. Additionally, the use of the derived theorem allows a PSO practitioner to obtain stability criteria that contains no artificial restriction on the relationship between control coefficients. The majority of previous stability results for PSO variants provided stability criteria under the restriction that certain control coefficients are equal; such restrictions are not present when using the derived theorem. Using the derived theorem, as demonstration of its ease of use, stability criteria are derived without the imposed restriction on the relation between the control coefficients for four popular PSO variants.

1. Introduction

The particle swarm optimization (PSO) algorithm, originally developed by Kennedy and Eberhart [1], has become a widely used optimization technique [2]. Given PSO's popularity, it has undergone a considerable amount of theoretical investigation, to list just a few, [3, 4, 5, 6, 7, 8, 9, 10, 11].

While meta-heuristics have been used to solve numerous real world problems, they typically lack the theoretical underpinning and guarantees that classical optimization techniques have. As such there is still a substantial need to bridge this gap between the theoretical work and the algorithmic design. By analyzing meta-heuristics from a theoretical perspective it allows the field to better understand how the underlying dynamics of the algorithm behavior and what the driving forces of this behavior are. This understanding can facilitate algorithmic improvements; a notable example of this is the covariance matrix adaptation evolutionary strategy (CMA-ES) [12], which has had a number of its improvements made possible directly from theoretical analysis. Theoretical understanding of the coefficient space also makes it possible to efficiently use meta-heuristics on new real world problems, as the behavioral implications of certain parameter choices can be theoretically guaranteed, which can drastically minimize the range of coefficients that need to be considered during parameter tuning.

There are a number of aspects of PSO behaviour that can be investigated from a theoretical perspective. However, the focus of this paper is on the criteria needed for order-1 and order-2 stability of PSO particles. Specifically, order-1 and order-2 stability occurs when particle positions converge to a constant in first and second order moment respectively [13][1]. The vast majority of theoretical studies have focused on reducing the modelling assumption used to obtain the stability criteria for PSO with inertia (referred to as canonical PSO (CPSO) in this paper), as proposed by Shi and Eberhart [14]. A detailed discussion of the systematic weakening of these modelling assumptions can be found in [11].

The focus of this paper is instead on providing an easy to use theorems for obtaining stability criteria for PSO variants, while using the minimal modelling assumptions[2] proposed by Cleghorn and Engelbrecht [11]. The general aim of this paper is to provide a theorem that allows a researcher to still obtain stability criteria even if they have made alterations, within reason, to the fundamental PSO algorithm. Recent empirical studies have shown that selecting PSO control coefficients that are both order-1 and order-2 stable are vital to the performance of PSO [15], and as such being able to easily obtain stability criteria for a PSO variant is an important issue for the field.

While progress in many aspects of PSO stability theory has been made there are still some common shortcomings that appear in the literature. While often not all these shortcomings are present, at least one generally is. The main limiting factors that appear in existing PSO stability research are:

- The derived stability criteria is too variant specific, making extrapolation past the variant under question challenging or impossible, often even under trivial alteration.
- There are restrictions imposed on the control coefficient in order to produce a more theoretically tractable

*Corresponding author
christopher.cleghorn@wits.ac.za (C.W. Cleghorn); belinda.stapelberg@up.ac.za (B. Stapelberg)
ORCID(s): 0000-0002-7860-0650 (C.W. Cleghorn); 0000-0002-3705-8789 (B. Stapelberg)

[1]Some authors have considered the stricter condition where the second order moment converges to zero. A detailed justification for using convergence to a constant second order moment is provided in [13].

[2]The term minimal is used, since Cleghorn and Engelbrecht [11] proved that order-1 and order-2 stability is present only if the non-stagnant distribution assumption holds, as defined in Section 4. The work of [11] is the most general in terms of modelling assumptions as well broadness of the result at present.

model. This restriction may not be fundamentally problematic if the practitioner is fully aware of what these restrictions are. However, even with proper knowledge of the imposed restrictions, using stability criteria under such coefficient relationship restrictions artificially excludes stable control coefficients, which is far from ideal.

- The derived stability criteria relies on strong modelling assumptions, such as the removal of stochasticity or assuming full informer stagnation.
- The stability result proved is sufficiently general, but application of which is far from mathematically trivial. This limiting factor increases the probability of errors in the derivation and also fundamentally dissuades PSO practitioners from deriving criteria for their proposed variant.

It should be stressed that strides have been made at reducing or removing many of these stated limitations. That said, at least one of the stated limitations is present in all existing works. At present the last stated limitation is not addressed at all in the literature, but is directly addressed within this paper without incurring the other stated limitations unnecessarily.

This paper provides a theoretical framework without these four limiting factors for a class of PSO variants. The PSO variants this paper considers are those that can be rewritten as a finite sum of stochastically weighted difference vectors between a particle's position and swarm informers. Many PSO variants can be written in this stated form. The canonical PSO is in this form naturally, with two particle informers, namely, the personal best position and the neighbourhood best position (or global best in the case of a fully connected swarm). The classic PSO variants, unified PSO (UPSO) [16] and fully informed PSO (FIPS)[17], both use multiple informers, and can be written as a finite sum of stochastically weighted difference vectors. There is also a more recent trend of adding a third informer to PSO's update equation to guide a particle's movement based on information external to the swarm itself. Specifically, in the work of Scheepers [18], a variant of PSO for multi-objective optimization utilizes a third informer from the Pareto front archive. A similar idea was also present in the work of Meier and Kramer [19], where gradient based information was used to construct a third informer to assist PSO in the training of recurrent neural networks.

The theorem presented in this paper, for obtaining stability criteria, also removes a common restriction present in existing stability work on PSO variants. Specifically, many previous order-2 stability results of PSO variants have provided stability criteria under the restriction that control coefficients are equal [20, 21, 22]; such restrictions are not present when using the provided theorem. An additional theorem is also provided for obtaining the fixed points for the expectations and variance of particle positions.

A brief description of PSO, and its general form, is given in Section 2, followed by a summary of existing relevant PSO theory in Section 3. The theoretical derivations of criteria for stability along with the limit points for particle positions are provided in Section 4. Section 5 demonstrates the use of the stability theorem by deriving the stability criteria for four PSO variants. Additionally, Section 5 provides the first order-1 and order-2 stability criteria for the comprehensive learning PSO (CLPSO)[23] as well as the first stability criteria for UPSO without restrictions on the relationship between control coefficients. An empirical demonstration of the negative impact of not meeting order-1 and order-2 stability criteria is presented in Section 6. A summary of the paper's findings is presented in Section 7.

2. Particle Swarm Optimization

Particle swarm optimization was originally inspired by the complex movement of birds in a flock. The variant of PSO this section focuses on is the CPSO algorithm [14].

The CPSO algorithm is defined as follows: Let $f : \mathbb{R}^d \to \mathbb{R}$ be the objective function that the CPSO algorithm aims to find an optimum for, where d is the dimensionality of the objective function. For the sake of simplicity, a minimization problem is assumed from this point onwards. Specifically, an optimum $\boldsymbol{o} \in \mathbb{R}^d$ is defined such that, for all $\boldsymbol{x} \in \mathbb{R}^d$, $f(\boldsymbol{o}) \leq f(\boldsymbol{x})$. In this paper the analysis focus is on objective functions where the optima exist. Let $\Omega(t)$ be a set of N particles[3] in $\mathbb{R}^d$ at a discrete time step t. Then $\Omega(t)$ is said to be the particle swarm at time t. The position $\boldsymbol{x}_i$ of particle i is updated using

$$\boldsymbol{x}_i(t+1) = \boldsymbol{x}_i(t) + \boldsymbol{v}_i(t+1), \tag{1}$$

where the velocity update, $\boldsymbol{v}_i(t+1)$, is defined as

$$\begin{aligned} \boldsymbol{v}_i(t+1) &= w\boldsymbol{v}_i(t) + c_1\boldsymbol{r}_1(t) \otimes (\boldsymbol{y}_i(t) - \boldsymbol{x}_i(t)) \\ &\quad + c_2\boldsymbol{r}_2(t) \otimes (\hat{\boldsymbol{y}}_i(t) - \boldsymbol{x}_i(t)), \end{aligned} \tag{2}$$

where $r_{1,k}(t), r_{2,k}(t) \sim U(0,1)$ for all t and $1 \leq k \leq d$. The operator $\otimes$ is used to indicate component-wise multiplication of two vectors. The position $\boldsymbol{y}_i(t)$ represents the "best" position that particle i has visited, where "best" means the location where the particle had obtained the lowest objective function evaluation. The position $\hat{\boldsymbol{y}}_i(t)$ represents the "best" position that the particles in the neighbourhood of the i-th particle have visited. The coefficients c_1, c_2, and w are the cognitive, social, and inertia weights, respectively.

There are numerous PSO variants that alter equation 2 of the CPSO algorithm. The focus of this paper is on PSO variants whose velocity update equation can be rewritten into the following form:

$$\boldsymbol{v}_i(t+1) = \boldsymbol{\theta}_0 \otimes \boldsymbol{v}_i(t) + \sum_{\iota=1}^{I} \boldsymbol{\theta}_\iota \otimes \left(\boldsymbol{\zeta}_\iota(t) - \boldsymbol{x}_i(t)\right) \tag{3}$$

$$\boldsymbol{x}_i(t+1) = \boldsymbol{x}_i(t) + \boldsymbol{v}_i(t+1) \tag{4}$$

where all $\theta_{\iota,k}$ form an arbitrary set of independent distributions with well defined mean and variance for each $0 \leq \iota \leq$

[3] The impact of the choice of N on PSO variants had recently been analyzed [24].

I, and $\boldsymbol{\zeta}_\iota$ represents each of the I particle informers. An informer is a general term for a point to which a particle is attracted to or its update is guided by. For example in the equation (2), $\boldsymbol{y}_i$ is an informer because the particle is attracted toward its personal best position due to the term $\boldsymbol{y}_i(t) - \boldsymbol{x}_i(t)$. Any positional formation that the particle is guided towards is an informer. For example, in the work done by Meier and Kramer [19], a position was used as a third informer that was constructed based on gradient information of a recurrent neural networks.

In order to make referring to this general PSO formulation easier it is referred to as N-Informer PSO (NIPSO). A full algorithm description of NIPSO variants is presented in Algorithm 1 for completeness.

Algorithm 1 PSO algorithm

Create and initialize a swarm, $\Omega(0)$, of N particles uniformly within a predefined hypercube of dimension d.
Let f be the objective function.
Let $\boldsymbol{y}_i$ represent the personal best position of particle i, initialized to $\boldsymbol{x}_i(0)$.
Let $\hat{\boldsymbol{y}}_i$ represent the neighbourhood best position of particle i, initialized to $\boldsymbol{x}_i(0)$.
Initialize $\boldsymbol{v}_i(0)$ to $\mathbf{0}$.
Let $t = 0$
repeat
  for all particles $i = 1, \cdots, N$ **do**
    if $f(\boldsymbol{x}_i) < f(\boldsymbol{y}_i)$ **then**
      $\boldsymbol{y}_i = \boldsymbol{x}_i$
    end if
    for all particles $\hat{i}$ with particle i in their neighbourhood **do**
      if $f(\boldsymbol{y}_i) < f(\hat{\boldsymbol{y}}_{\hat{i}})$ **then**
        $\hat{\boldsymbol{y}}_{\hat{i}} = \boldsymbol{y}_i$
      end if
    end for
  end for
  $t = t + 1$
  for all particles $i = 1, \cdots, N$ **do**
    update the velocity of particle i using equation (3)
    update the position of particle i using equation (4)
  end for
until stopping condition is met

3. Current PSO Stability Analysis

The derivation of stability criteria is one of the most theoretically studied aspects of PSO, with the vast majority of work being focused on CPSO. Early work included a number of simplifying assumptions. The two most notable of which are the deterministic assumption and the stagnation assumption.

The deterministic assumption is where it was assumed that $c_1\mathbf{r}_1(t)$ and $c_2\mathbf{r}_2(t)$ were held constant [4, 3, 25, 26]. The majority of recent work have dropped the necessity of this assumption and have derived stability criteria with the inclusion of stochasticity. The move to the stochastic context necessitated considering both the first and second order moments of particles, as pointed out by Poli [20]. The move to a fully stochastic context made finding both necessary and sufficient conditions for stability more challenging, with a number of early works only being able to provide conservative stability regions, such as the work of Kadirkamanathan *et al.* [27] and Gazi [6]. However, Poli [28] was able to obtain a non-conservative stability region by not relying on Lyapunov conditions [29]. However, Poli's work, like the other studies mentioned here, relied on the stagnation assumption.

At this point attention shifted to trying to weaken the stagnation assumption. The stagnation assumption assumes that the informers are unchanging (stagnant), which is a significant simplification to the PSO's dynamics. Liu [9] improved the modelling accuracy with weak stagnation, where it was assumed that $\mathbf{y}_{\hat{i}}(t) = \mathbf{y}_{\hat{i}}$, for all t sufficiently large, where $\hat{i}$ is the index of the particle that has obtained the best objective function evaluations. Bonyadi and Michalewicz [22] improved the modelling by proposing the stagnant distribution assumption, in which the informers were rather modelled as random positions sampled from fixed distributions. More recently, Cleghorn and Engelbrecht [11], improved the modelling by using the non-stagnant distribution assumption, in which the informers are modelled as random positions sampled from a time dependent distribution. What makes the work of Cleghorn and Engelbrecht a useful point from which to build a theoretical framework is that the modelling assumption itself was shown to be a necessary condition for order-1 and order-2 stability. What this implies is that even if an objective function dependent approach to stability analysis is considered, the criteria found in [11] would still necessarily hold.

Almost all existing work has derived of stability criteria directly for specific PSO variants, with most focusing on CPSO. A number of PSO variants have been directly studied [7, 19, 30, 31]. Recently, Cleghorn and Engelbrecht [11] proved Theorem 1 which allows for the derivation stability criteria for all PSO variants with the componentwise form:

$$x_k(t+1) = x_k(t)\alpha + x_k(t-1)\beta + \gamma_t \tag{5}$$

where α and β are well defined[4] random variables, and (γ_t) is a sequence of well defined random variables. The index k indicates the vector component. The full theorem is now stated to assist in the subsequent derivations in Section 4:

Theorem 1. *The following properties hold for all PSO variants of the form described in equation (5), where $E[\cdot]$ and $V[\cdot]$* [5] *are the expectation and variance operator respectively , and $\rho(\cdot)$ is the spectral radius of a matrix.*

[4] In the context of this work a well defined random variable is one that has a mean and variance.

[5] For the sake of completeness: the expected value of a function $G : \mathbb{R} \to \mathbb{R}$ of a continuous random variable $X \sim p(x)$ is given by $E_X[g(x)] = \int_{\mathcal{X}} g(x)p(x)dx$ where p is the probability density function. The variance of a function $G : \mathbb{R} \to \mathbb{R}$ of a continuous random variable $X \sim p(x)$ is given by $V_X[g(x)] = \int_{\mathcal{X}} (g(x) - E_X[g(x)])^2 p(x)dx$ where p is the probability density function.

1. *Assuming $\boldsymbol{i}_t$ converges, particle positions are order-1 stable for every initial condition if and only if $\rho(\boldsymbol{A}) < 1$, where*

$$\boldsymbol{A} = \begin{bmatrix} E[\alpha] & E[\beta] \\ 1 & 0 \end{bmatrix} \text{ and } \boldsymbol{i}_t = \begin{bmatrix} E[\gamma_t] \\ 0 \end{bmatrix}. \quad (6)$$

2. *The particle positions are order-2 stable if $\rho(\boldsymbol{B}) < 1$ and $(\boldsymbol{j}_t)$ converges, where*

$$\boldsymbol{B} = \begin{bmatrix} E[\alpha] & E[\beta] & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & E[\alpha^2] & E[\beta^2] & 2E[\alpha\beta] \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & E[\alpha] & 0 & E[\beta] \end{bmatrix}$$

and

$$\boldsymbol{j}_t = \begin{bmatrix} E[\gamma_t] \\ 0 \\ E[\gamma_t^2] \\ 0 \\ 0 \end{bmatrix} \quad (7)$$

under the assumption that the limits of $(E[\gamma_t \alpha])$ and $(E[\gamma_t \beta])$ exist.

3. *Assuming that $x(t)$ is order-1 stable, then the following is a necessary condition for order-2 stability:*

$$1 - E[\alpha] - E[\beta] \neq 0 \quad (8)$$

$$1 - E\left[\alpha^2\right] - E\left[\beta^2\right] - \left(\frac{2E[\alpha\beta]\,E[\alpha]}{1 - E[\beta]}\right) > 0 \quad (9)$$

4. *The convergence of $E[\gamma_t]$ is a necessary condition for order-1 stability, and the convergence of both $E[\gamma_t]$ and $E[\gamma_t^2]$ is a necessary condition for order-2 stability.*

While the generality of Theorem 1 is useful, it can make it potentially challenging for practitioners to quickly obtain stability criteria for their custom PSO variant from the theorem without a considerable amount of calculation. An example of the rigorous use of Theorem 1 can be found in [32].

In order to reduce the burden on practitioners to derive stability criteria, a specialization of Theorem 1 to the class of PSOs described by equations (3) and (4) is proposed in this paper. The intention of the specialization is to make obtaining stability criteria as easy as possible, while still maintaining a sufficient degree of generality to cater for a range of variations in the PSO update equation formulation. The overarching goal of the specialization is to reduce the need to perform full stability analysis for most simple variants of PSO. In particular if a practitioner wished to augment the PSO update equations, ideally they should be able to quickly determine what the stability criteria of their bespoke variant is. Knowledge of the stability criteria is of vital importance for parameter tuning as it has been demonstrated that the stability of PSO particles is highly correlated to the performance of PSO [15].

At present the majority of existing second order stability criteria published have restricted the relationship between control coefficients. Specifically, the coefficients of CPSO have often been restricted such that c_1 and c_2 were assumed either equal [28] or to have equal means and variances [22]. The theorem proved in the next section removes any such restriction, and can therefore produce stability criteria for arbitrary coefficient relationships.

Table 1
Summary of symbols needed in the use of novel theorem 2 and 3

Symbol	Meaning
$E[\cdot]$	Expectation operator
$V[\cdot]$	Variance operator
$\otimes$	Component-wise multiplication
$\rho(\cdot)$	Spectral radius operator
$\boldsymbol{x}_i$	Particle i's position
$\boldsymbol{v}_i$	Particle i's velocity
$\boldsymbol{\zeta}_\iota$	The ι-th particle informer
θ_0	The stochastic weighting of the old velocity
θ_ι $(\iota \geq 1)$	The stochastic weighting of $\boldsymbol{\zeta}_\iota(t) - \boldsymbol{x}_i(t)$
ϕ	$\sum_{\iota=1}^{I} V[\theta_\iota]$
ψ	$\sum_{\iota=1}^{I} E[\theta_\iota]$

4. Specialization to N-Informers

This section provides the derivation of order-1 and order-2 stability criteria for the class of PSO variants as defined in equations (3) and (4), which are collectively referred to as NIPSO. Furthermore, the order-1 and order-2 fixed points are derived. In order to facilitate ease of use for a practitioner the symbols required to use the two theorems are summarized in Table 1.

Theorem 2. *The following properties hold for all NIPSO combinations, under the non-stagnant distribution assumption for each informer.*[6]

1. *Particle positions are order-1 stable for every initial condition if and only if*

$$-1 < E[\theta_0] < 1 \quad (10)$$

and

$$0 < \sum_{\iota=1}^{I} E[\theta_\iota] < 2(E[\theta_0] + 1) \quad (11)$$

2. *Particle positions are order-2 stable for every initial condition only if*

$$-1 < \frac{E[\theta_0]}{\sqrt{1 - V[\theta_0]}} < 1 \quad (12)$$

and

$$0 < \psi < \frac{-2\left(E[\theta_0]^2 + V[\theta_0] - 1\right)}{1 - E[\theta_0] + \frac{\phi(1+E[\theta_0])}{\psi^2}} \quad (13)$$

[6]**Non-stagnant distribution assumption**:
Let $\boldsymbol{\xi}_i(t)$ be an informer. It is assumed that $\boldsymbol{\xi}_i(t)$ is a random variable sampled from a time dependent distribution, such that $\boldsymbol{\xi}_i(t)$ has a well defined expectation and variance for each t and that $\lim_{t\to\infty} E[\boldsymbol{\xi}_i(t)]$ and $\lim_{t\to\infty} V[\boldsymbol{\xi}_i(t)]$ exist. A detailed justification of this modelling choice is given by Cleghorn and Engelbrecht [11].

$$\text{where } \phi = \sum_{\iota=1}^{I} V[\theta_\iota] \text{ and } \psi = \sum_{\iota=1}^{I} E[\theta_\iota].$$

Proof (1): Let the non-stagnant distribution assumption hold for each of the I informers [7]. Rewriting equations (3) and (4) into the general form of equation (5) leads to:

$$\alpha = (1+\theta_0) - \sum_{\iota=1}^{I}\theta_\iota$$
$$\beta = -\theta_0$$
$$\gamma_t = \sum_{\iota=1}^{I}\theta_\iota \zeta_\iota(t)$$

In order to utilize part (1) of Theorem 1 to obtain the order-1 stability criteria, the matrix $\boldsymbol{A}$ and the vector $\boldsymbol{i}_t$ must be constructed as defined in equation (6). The required expectations are calculated as follows:

$$E[\alpha] = 1 + E[\theta_0] - \sum_{\iota=1}^{I} E[\theta_\iota] \tag{14}$$
$$E[\beta] = -E[\theta_0] \tag{15}$$
$$E[\gamma_t] = \sum_{\iota=1}^{I} E[\theta_\iota]E[\zeta_\iota(t)] \tag{16}$$

which leads to

$$A = \begin{bmatrix} 1 + E[\theta_0] - \sum_{\iota=1}^{I} E[\theta_\iota] & -E[\theta_0] \\ 1 & 0 \end{bmatrix} \tag{17}$$

and

$$\boldsymbol{i}_t = \begin{bmatrix} \sum_{\iota=1}^{I} E[\theta_\iota]E[\zeta_\iota(t)] \\ 0 \end{bmatrix}. \tag{18}$$

Since $E[\theta_\iota]$ is well defined for each ι and $E[\zeta_\iota(t)]$ is well defined and convergent for each ι, by the non-stagnant distribution assumption it follows that $i_{t,0}$ is convergent and therefore $\boldsymbol{i}_t$ is convergent. In order to find the criteria needed to satisfy the condition $\rho(\boldsymbol{A}) < 1$, the eigenvalues of $\boldsymbol{A}$ are required and are calculated to be:

$$\lambda_1, \lambda_2 = \frac{\eta \pm \sqrt{\eta^2 - 4E[\theta_0]}}{2} \tag{19}$$

where $\eta = 1+E[\theta_0]-\sum_{\iota=1}^{I} E[\theta_\iota]$. After some simplification it is found that $\rho(\boldsymbol{A}) < 1$ holds if and only if

$$-1 < E[\theta_0] < 1 \text{ and } 0 < \sum_{\iota=1}^{I} E[\theta_\iota] < 2(E[\theta_0]+1). \tag{20}$$

It follows from part (1) of Theorem 1 that NIPSO is order-1 stable if and only if the criteria of equations (20) and (11) hold. □

Proof (2): Let the non-stagnant distribution assumption hold for each of the I informers. In order to obtain the necessary conditions for order-2 stability, part 3 of Theorem 1 is utilized. A number of expectations are required to construct the matrix $\boldsymbol{B}$ and the vector $\boldsymbol{j}_t$. Specifically, $E[\alpha^2]$, $E[\beta^2]$, and $E[\alpha\beta]$ are required, and calculated as

$$\begin{aligned} E[\alpha^2] &= E\left[\left(1+\theta_0 - \sum_{\iota=1}^{I}\theta_\iota\right)^2\right] \\ &= 1 + 2E[\theta_0] - 2\sum_{\iota=1}^{I} E[\theta_\iota] \\ &\quad - 2E[\theta_0]\sum_{\iota=1}^{I} E[\theta_\iota] + E\left[\left(\sum_{\iota=1}^{I}\theta_\iota\right)^2\right], \end{aligned} \tag{21}$$

where

$$\begin{aligned} E\left[\left(\sum_{\iota=1}^{I}\theta_\iota\right)^2\right] &= V\left[\sum_{\iota=1}^{I}\theta_\iota\right] + \left(\sum_{\iota=1}^{I} E[\theta_\iota]\right)^2 \\ &= \sum_{\iota=1}^{I} V[\theta_\iota] + \sum_{i\neq j} cov\left(\theta_i, \theta_j\right) + \left(\sum_{\iota=1}^{I} E[\theta_\iota]\right)^2 \\ &= \sum_{\iota=1}^{I} V[\theta_\iota] + \left(\sum_{\iota=1}^{I} E[\theta_\iota]\right)^2, \end{aligned} \tag{22}$$

since each θ_i are independent. Substituting equation (22) back into equation (21) leads to,

$$\begin{aligned} E[\alpha^2] &= 1 + 2E[\theta_0] - 2(1+[\theta_0])\sum_{\iota=1}^{I} E[\theta_\iota] \\ &\quad + \sum_{\iota=1}^{I} V[\theta_\iota] + \left(\sum_{\iota=1}^{I} E[\theta_\iota]\right)^2. \end{aligned} \tag{23}$$

The expectation of β^2 and $\alpha\beta$ are easily calculated as:

$$E[\beta^2] = E[\theta_0^2] = V[\theta_0] + E[\theta_0]^2 \tag{24}$$

$$\begin{aligned} E[\alpha\beta] &= E\left[-\theta_0\left((1+\theta_0) - \sum_{\iota=1}^{I}\theta_\iota\right)\right] \\ &= -E[\theta_0] - V[\theta_0] - E[\theta_0]^2 - E[\theta_0]\sum_{\iota=1}^{I} E[\theta_\iota]. \end{aligned} \tag{25}$$

For equation (8) in part 3 of Theorem 1 to be satisfied the following condition must hold:

$$\psi = \sum_{\iota=1}^{I} E[\theta_\iota] \neq 0 \tag{26}$$

For equation (9) in part 3 of Theorem 1 to be satisfied the following condition must hold:

$$1 + 2E[\theta_0] + 2(1+E[\theta_0])\psi - \phi - \psi^2 - V[\theta_0] + E[\theta_0]^2 - \left(\frac{2\left(-E[\theta_0] - V[\theta_0] - E[\theta_0]^2 - E[\theta_0]\psi\right)\left(1+E[\theta_0]-\psi\right)}{1+E[\theta_0]}\right) > 0,$$

which is simplified using a method similar to that of Bonyadi and Michalewicz [22], to equal the criteria of equations (12)

[7] Strictly speaking, only a well defined expectation and limit point of the informer is needed to prove part 1.

and (13). The necessary condition of part 2 of Theorem 2 is therefore proved. □

It is worth verifying that the stability criteria of equations (12) and (13) are not only a necessary condition for order-2 stability, but in fact sufficient as well. This is achieved by verifying that if the criteria of equations (12) and (13) are satisfied then $\rho(\mathbf{B}) < 1$, from Theorem 1 part 2. All the expectations needed to construct matrix $\mathbf{B}$ have already been obtained while deriving the necessary condition. In order to verify that $\rho(\mathbf{B}) < 1$ the empirical approach of Bonyadi and Michalewicz [22] and Cleghorn and Engelbrecht [11] is used. Specifically, for $I = 1, 2, \cdots, 50$ informers[8] the experimental procedure followed is: $I \times 10^8$ random configurations representing $\{E[\theta_0], V[\theta_0], \cdots, E[\theta_I], V[\theta_I]\}$ are generated such that equations (12) and (13) are satisfied. In all of the cases it was found that if equations (12) and (13) were satisfied, then the condition $\rho(\mathbf{B}) < 1$ held. This finding is strong evidence that the criteria is sufficient for order-2 stability. □

Theorem 3. *The following properties hold for all NIPSO combinations:*

1. *Under order-1 stability the fixed points of the particle position expectations are:*

$$E_{x_{i,k}} = \frac{\sum_{\iota=1}^{I} E\left[\theta_\iota\right] E\left[\zeta_{\iota,k}\right]}{\sum_{\iota=1}^{I} E\left[\theta_\iota\right]} \tag{27}$$

where $E[\zeta_{\iota,k}]$ is the limit of $E[\zeta_{\iota,k}(t)]$.

2. *Under order-1 and order-2 stability, the fixed points of the particle position variances are:*

$$V_{x_{i,k}} = \frac{\left(1+E[\theta_0]\right)\left(\kappa_1 - 2\kappa_2 E_{x_{i,k}} + \kappa_3 E^2_{x_{i,k}}\right)}{2\psi\left(1 - E^2[\theta_0] - V[\theta_0]\right) - \phi\left(1+E[\theta_0]\right) + \psi^2\left(E[\theta_0]-1\right)} \tag{28}$$

where

$$\kappa_1 = \sum_{\iota=1}^{I}\left(E^2[\theta_\iota]V[\zeta_{\iota,k}] + E^2[\zeta_{\iota,k}]V[\theta_\iota] + V[\theta_\iota]V[\zeta_{\iota,k}]\right),$$

$$\kappa_2 = \sum_{\iota=1}^{I} V[\theta_\iota]E[\zeta_{\iota,k}], \quad \phi = \sum_{\iota=1}^{I} V[\theta_\iota], \quad \psi = \sum_{\iota=1}^{I} E[\theta_\iota],$$

with $E[\zeta_{\iota,k}]$ and $V[\zeta_{\iota,k}]$ as the the limit of $E[\zeta_{\iota,k}(t)]$ and $V[\zeta_{\iota,k}(t)]$ respectively.

Proof (1): Under the assumption of order-1 stability each particle i converges to a fixed point in expectation. Let such a fixed point be called $\boldsymbol{E}_{x_i}$. The fixed point is calculated by rewriting equations (3) and (4) into the following component-wise second order recurrence relation form:

$$x_{i,k}(t+1) = x_{i,k}(t)(1+\theta_0) - \theta_0 x_{i,k}(t-1) + \sum_{\iota=1}^{I}\theta_\iota\left(\zeta_{\iota,k}(t) - x_{i,k}(t)\right). \tag{29}$$

[8] While the experimental verification was only done up to 50 informers, there is no clear reason why it would fail to hold for higher informer counts. Practically speaking, a variant with more than 50 informers seems unlikely.

Applying the expectation operator leads to

$$E[x_{i,k}(t+1)] = E[x_{i,k}(t)](1+E[\theta_0]) - E[\theta_0]E[x_{i,k}(t-1)] + \sum_{\iota=1}^{I} E[\theta_\iota]\left(E[\zeta_{\iota,k}(t)] - E[x_{i,k}(t)]\right). \tag{30}$$

Then by setting $E[x_{i,k}(t-1)] = E[x_{i,k}(t)] = E[x_{i,k}(t+1)] = E_{x_{i,j}}$ and $E[\zeta_{\iota,k}(t)]$ to its limits $E[\zeta_{\iota,k}]$, equation (30) can be rearranged to find an explicit expression for $E_{x_{i,j}}$, thus obtaining equation (27). □

Proof (2): Under the assumption of order-1 and order-2 stability each particle i converges to a fixed point for each of the following sequences: $E[\boldsymbol{x}_i(t)]$, $E[\boldsymbol{x}_i(t)\boldsymbol{x}_i(t-1)]$, and $E[\boldsymbol{x}_i^2(t)]$. Let such fixed points be called $\boldsymbol{E}_{x_i}$, $\boldsymbol{E}_{x_i x_i}$, and $\boldsymbol{E}_{x_i^2}$ respectively as we will be working in the limit. First define

$$\partial x_{i,k}(t) = x_{i,k}(t) - E[x_{i,k}(t)] = x_{i,k}(t) - E_{x_{i,k}}. \tag{31}$$

It follows that $V[x_{i,k}(t)] = E[\partial^2 x_{i,k}(t)]$, where $\partial^2 x_{i,k}(t)$ denotes the square of equation (31). In order to obtain $V[x_{i,k}(t)]$, consider the class of recurrence relations as defined in equation (5), and that $\partial x_{i,k}(t)$ can be rewritten as

$$\partial x_{i,k}(t) = \alpha\partial x_{i,k}(t-1) + \beta\partial x_{i,k}(t-2) + d_{i,k}(t-1) \tag{32}$$

$$d_{i,k}(t-1) = \gamma_{t-1} - E_{x_{i,k}}(1-\alpha-\beta). \tag{33}$$

Squaring and applying the expectation operator to equation (32) leads to

$$\begin{aligned} &E[\partial^2 x_{i,k}(t)] \\ &= E[\alpha^2]E[\partial^2 x_{i,k}(t-1)] + 2E[\alpha\beta]E[\partial x_{i,k}(t-1)\partial x_{i,k}(t-2)] \\ &- E[d_{i,k}(t-1)]\left(2E[\alpha]E[\partial x_{i,k}(t-1)] + 2E[\beta]E[\partial x_{i,k}(t-2)]\right) \\ &+ E[\beta^2]E[\partial^2 x_{i,k}(t-2)] + E[d_{i,k}(t-1)]^2. \end{aligned} \tag{34}$$

In order to simplify equation (34) consider that

$$E[\partial x_{i,k}(t)\partial x_{i,k}(t-1)] = E[\partial x_{i,k}(t-1)\partial x_{i,k}(t-2)] \tag{35}$$

and

$$E[\partial x_{i,k}(t-2)] = E[\partial x_{i,k}(t-1)] = E[\partial E_{x_{i,k}}] = 0, \tag{36}$$

which follows from order-1 and order-2 stability [33]. Now

$$\begin{aligned} &E[\partial x_{i,k}(t)\partial x_{i,k}(t-1)] \\ &= E[\alpha]E[\partial^2 x_{i,k}(t-1)] + E[\beta]E[\partial x_{i,k}(t-2)\partial x_{i,k}(t-1)] \\ &+ E[d_{i,k}(t-1)]E[\partial x_{i,k}(t-1)]. \end{aligned} \tag{37}$$

Using equations (35) and (36), equation (37) can be rearranged to yield,

$$E[\partial x_{i,k}(t)\partial x_{i,k}(t-1)] = \frac{E[\alpha]E[\partial^2 x_{i,k}(t-1)]}{1-E[\beta]}. \tag{38}$$

Now all that remains is to substitute equation (38) into equation (34), and utilize the fact that $E[\partial^2 x_{i,k}(t-1)] = E[\partial^2 x_{i,k}(t)]$ (once again this is permissible in the limit), to obtain

$$E[\partial^2 x_{i,k}(t)] = \frac{E[d_{i,k}(t-1)^2]}{1 - E[\alpha^2] - E[b^2] - \frac{2E[\alpha\beta][\alpha]}{1-E[\beta]}}. \tag{39}$$

Equation (39) represents the variance fixed point for the large class of PSOs. However, our focus is on the case where

$$\alpha = (1+\theta_0) - \sum_{\iota=1}^{I} \theta_\iota$$
$$\beta = -\theta_0$$
$$\gamma_t = \sum_{\iota=1}^{I} \theta_\iota \zeta_{\iota,k}(t).$$

Substituting these specific α, β, and γ_t into equation (39) and performing a substantial amount of simplification (which is omitted for the sake of brevity), leads to equation (28) as was required to be proved. □

It is at this point worth reflecting on the four common limitations, where at least one of which is present, in existing PSO stability work mentioned in Section 1:

- The first limitation stated is that most existing stability is very variant specific. This limitation is mostly, but not completely, removed in this paper. The presented theory only applies to PSO variants that can be reformulated into the NIPSO form. That said, the NIPSO formulation allows for a considerable amount of variation. Specifically, the variants can possess any number of informers, as well as allowing any combination of well defined stochastic weighting coupled with any number of tunable coefficients. This degree of allowable variation is combinatorially substantial.

- The next limitation stated is that many stability results, for PSO variants, artificially impose restrictions on the relationship between control coefficient. This limitation is completely removed in new the theorems presented in this paper, with the criteria for stability as well as fixed points being directly obtainable without needing to assume any relationship between the coefficients. This fact is directly visible from equations (10-13) where each θ_i, $i = 0, \dots, I$, are handled directly.

- The third limitation stated was the reliance on overly strong modelling assumption. The results presented in this paper do not suffer from this problem as both Theorem 2 and 3 are derived using Theorem 1, which holds under provably minimal modelling assumptions. This implies that the presented stability theorems also hold under minimal modelling assumptions, as no additional assumptions are used to prove the presented theorems.

- The last limitation stated was the difficulty of using existing stability results, with simple alterations to the PSO potentially requiring pages of mathematical derivation. This is one of the main strengths of the theorems provided in this paper, with derivation of new results becoming comparatively trivial. The best way to illustrate this is via demonstration, which is done in the next section, where result that would have previously required a substantial amount of technical derivation are obtained in a couple of simple steps.

5. Application of Stability Results

In this section a number of existing stability criteria are re-derived to demonstrate how Theorem 2 can be easily applied to rapidly obtain stability criteria. Furthermore, many previous stability results on PSO variants have considerably restricted the allowable relationship between control coefficients, such a limitation is not present in this section. All derived criteria contained in this section have no restriction on the coefficient relations.

Stability criteria for CPSO, CLPSO, FIPS, and UPSO are derived in sections 5.1, 5.2, 5.3, and 5.4 respectively.

5.1. Canonical PSO

Consider the case of the CPSO algorithm, as defined by equations (1) and (2). After dropping the particle and component indices, without loss of generality [9], the stability criteria for CPSO can be obtained by using two informers with $\theta_0 = w$, $\theta_1 = c_1 r_1$, and $\theta_2 = c_2 r_2$, where $r_1, r_2 \sim U(0,1)$. In order to utilize Theorem 2, ψ and ϕ are required and calculated as:

$$\psi = \sum_{\iota=1}^{2} E[\theta_\iota] = \frac{c_1}{2} + \frac{c_2}{2}, \quad \phi = \sum_{\iota=1}^{2} V[\theta_\iota] = \frac{c_1^2}{12} + \frac{c_2^2}{12}. \quad (40)$$

Substituting ψ and ϕ into the criteria of equations (10-13) the following criteria for order-1 and order-2 stability are obtained:

$$-1 < w < 1 \quad \text{and} \quad 0 < c_1 + c_2 < \frac{4\left(1-w^2\right)}{1 - w + \frac{\left(c_1^2+c_2^2\right)(1+w)}{3(c_1+c_2)^2}}. \quad (41)$$

The criteria in equation (41) is CPSO's full order-1 and order-2 stability criteria without using the simplified case where $c_1 = c_2$. If the simplification is reimposed, the following commonly reported form reappears:

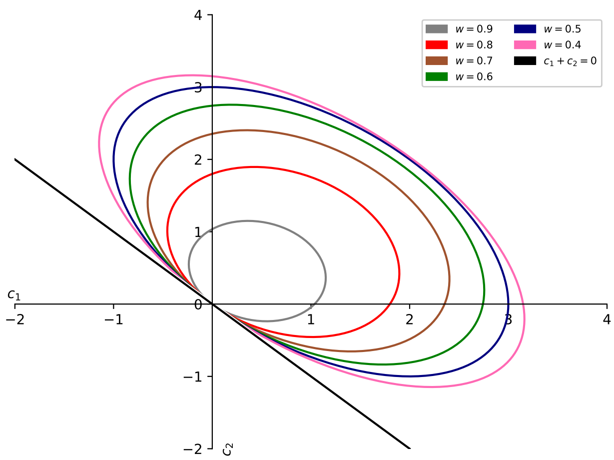
$$-1 < w < 1 \quad \text{and} \quad 0 < c_1 + c_2 < \frac{24\left(1-w^2\right)}{7-5w}. \quad (42)$$

It is interesting to observe that the weighting between c_1 and c_2 has a direct influence on the size and shape of the stability region as illustrated in Figure 1, where the cross-sections of the stability region, with fixed inertia values, are shown. Figure 2 demonstrates the complex relationship between the cross-sections of the stability region of w and c_1 pairs for differing c_2 values, with the regions of stable pairs decreasing in size and changing shape as c_2 is varied from 3 to 0.5 in figure 2a. In Figure 2b the stable region reaches its maximal volume at $c_2 = 0$ and once again begins decreasing and becomes more oval in shape as c_2 is varied from 0 to -1.5. Additionally, both figures 1 and 2 demonstrate that using equation (42) without the knowledge that it had been derived with the restriction that $c_1 = c_2$, would lead to a

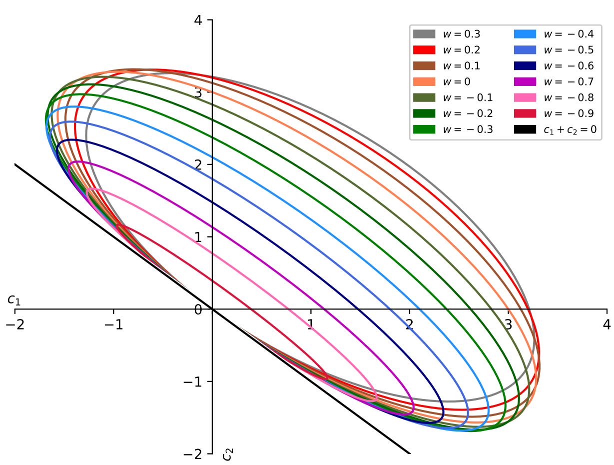
[9] Dropping the indices can be done without loss of generality, because each dimension of each particle, can be treated as its own recurrence relationship which only differs in indices. Since each of the recurrence relationships has the same set of coefficients and structure, the solution to one is the same across all of them. In terms of potential coupling within an objective function itself or among particles, the non-stagnate distribution assumption, as with the stronger stagnation assumption, induces the decoupling.

practitioner misclassifying the stability of parameter configurations, making the use of equation (41) clearly preferable.

The stable region presented in equation (41) is the largest region in the literature that guarantees both order-1 and order 2 stability for CPSO, and is believed to be maximal given the agreement found between the necessary and sufficient conditions of Theorem 1.



(a) Order-2 stable regions for $w = 0.4$ to 0.9.



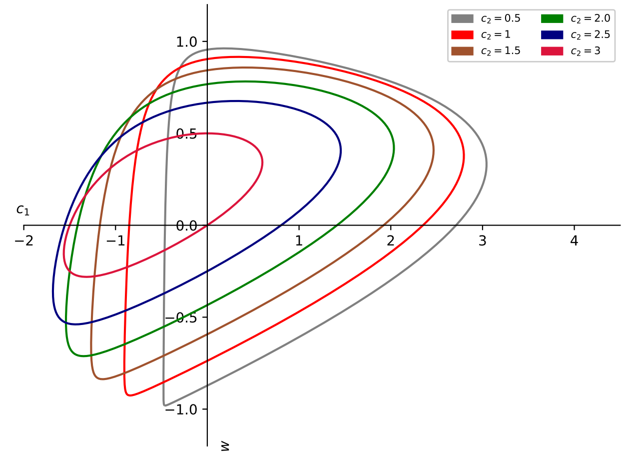
(b) Order-2 stable regions for $w = -0.9$ to 0.3.

Figure 1: Order-2 stable regions of CPSO under fixed inertia values. The interior region of the elliptic shapes correspond to where equation (41) is satisfied for a given w.

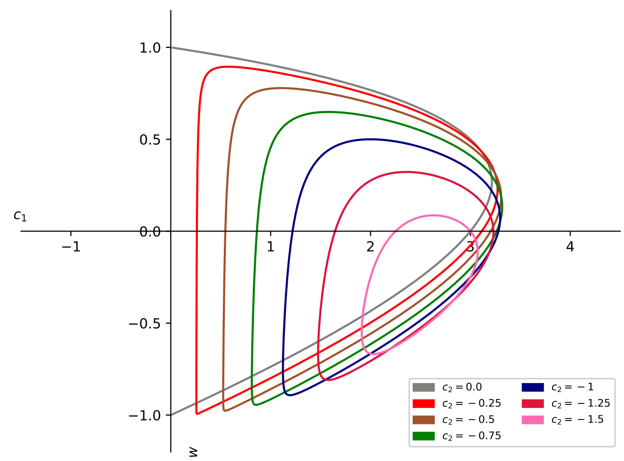
5.2. Comprehensive Learning Particle Swarm Optimizer

The CLPSO algorithm is a comparatively more recent PSO variant, proposed by Liang *et al* [23] in which the authors alter PSO's update equation to include an informer derived from all other particles' historical best information. Specifically, the velocity update equation for CLPSO is:

$$\boldsymbol{v}_i(t+1) = w\boldsymbol{v}_i(t) + c\mathbf{r} \otimes \left(\mathbf{l}_i(t) - \boldsymbol{x}_i(t)\right) \tag{43}$$



(a) Order-2 stable regions for $c_2 = 0.5$ to 3.



(b) Order-2 stable regions for $c_2 = 0$ to -1.5.

Figure 2: Order-2 stable regions of CPSO under fixed c_2 values. The interior region of the elliptic shapes correspond to where equation (41) is satisfied for a given c_2. The same visual pattern is observed if c_1 is fixed and c_2 is varied due to the symmetry present in (41).

where $\mathbf{r} \sim U(0,1)^d$ and

$$l_{i,k}(t) = \begin{cases} y_{i,k}(t) & \text{if } s > PC_i \\ p_{i,k}(t) & \text{otherwise} \end{cases} \tag{44}$$

where $s \sim U(0,1)$ and PC_i is referred to as the learning probability and $p_{i,k}(t)$ is constructed, for each dimension k, using two entity tournament selections across all personal best positions in the swarm. The selectable index is reset after a refreshing gap on non-improvement. More precise details can be found in [23].

The position $l_{i,k}(t)$, at each time step t, can be modelled as a Bernoulli distribution with support $\{y_{i,k}(t), p_{i,k}(t)\}$ that is parameterized by PC_i. In this context, under the non-stagnant distribution assumption on $l_{i,k}(t)$, stability analysis is possible. After dropping the particle and component indices, without loss of generality, the stability criteria for

CLPSO can be obtained by considering, $\theta_0 = w$, $\theta_1 = cr$, $r \sim U(0,1)$. The following calculations are required to use Theorem 2:

$$\psi = \sum_{\imath=1}^{I} E[\theta_\imath] = E[\theta_1] = \frac{c}{2} \quad (45)$$

$$\phi = \sum_{\imath=1}^{I} V[\theta_\imath] = V[\theta_1] = \frac{c}{12} \quad (46)$$

Substituting ψ and ϕ into the criteria of equations (10-13) the following criteria for order-1 and order-2 stability are obtained:

$$-1 < w < 1 \quad \text{and} \quad 0 < c < \frac{6\left(1-w^2\right)}{2-w}. \quad (47)$$

The size and the shape of the region corresponds with the line $c_2 = 0$ shown in Figure 2b, if $c_1 = c$.

5.3. Fully Informed PSO

The FIPS algorithm is an early PSO variant proposed by Kennedy and Mendes [17], based on the observation that in human society individuals are not influenced by only a single individual, but rather by a statistical summary of the state of their neighbourhood. In the FIPS algorithm, the velocity equation of CPSO is altered such that each particle is influenced by all its neighbours. Specifically, the velocity update equation for FIPS is:

$$\boldsymbol{v}_i(t+1) = w\boldsymbol{v}_i(t) + \sum_{m=1}^{|\mathcal{N}_i|} \boldsymbol{\gamma}_m \otimes \frac{(\boldsymbol{y}_m(t) - \boldsymbol{x}_i(t))}{|\mathcal{N}_i|}, \quad (48)$$

where $\mathcal{N}_i$ is the set of particles in particle i's neighbourhood, $\boldsymbol{y}_m(t) \in \mathcal{N}_i$, and $\gamma_{m,k} \sim U(0,\hat{c})$, were $c_1 + c_2 = \hat{c}$.

After dropping the particle and component indices, without loss of generality, the stability criteria for FIPS can be obtained by considering $I = |\mathcal{N}|$ informers and setting $\theta_0 = w$ and $\theta_\imath = \frac{\gamma_\imath}{|\mathcal{N}|}$ for $1 \le \imath \le |\mathcal{N}|$. The following calculations are required to use Theorem 2:

$$\psi = \sum_{\imath=1}^{I} E[\theta_\imath] = \sum_{\imath=1}^{|\mathcal{N}|} \frac{E[\gamma_\imath]}{|\mathcal{N}|} = \sum_{\imath=1}^{|\mathcal{N}|} \frac{\hat{c}}{2|\mathcal{N}|} = \frac{\hat{c}}{2} \quad (49)$$

and

$$\phi = \sum_{\imath=1}^{I} V[\gamma_\imath] = \sum_{\imath=1}^{|\mathcal{N}|} V\left[\frac{\gamma_\imath}{|\mathcal{N}|}\right] = \sum_{\imath=1}^{|\mathcal{N}|} \frac{\hat{c}^2}{12|\mathcal{N}|^2} = \frac{\hat{c}^2}{12|\mathcal{N}|}. \quad (50)$$

Substituting ψ and ϕ into the criteria of equations (10-13) the following criteria for order-1 and order-2 stability are obtained:

$$-1 < w < 1 \quad \text{and} \quad 0 < \frac{\hat{c}}{2} < \frac{12|\mathcal{N}|\left(1-w^2\right)}{3|\mathcal{N}| + 1 + w(1-3|\mathcal{N}|)}. \quad (51)$$

The derived criteria is in agreement with existing criteria of both Cleghorn and Engelbrecht [30] and García-Gonzalo and Fernández-Martinez [34], but are obtained with minimal calculations, and under a weaker modelling assumption.

5.4. Unified PSO

The UPSO algorithm was designed by Parsopoulos and Vrahatis [16] as a weighted merger between the local best PSO and the global best PSO. The PSO variant utilizes the additional control parameter, $u \in [0,1]$, called the unification factor, to control the importance placed on either the global best PSO update or the local best PSO. Specifically, the update equations for UPSO are:

$$\begin{aligned} \boldsymbol{g}_i(t+1) &= w\boldsymbol{v}_i(t) + c_1\boldsymbol{r}_1 \otimes (\boldsymbol{y}_i(t) - \boldsymbol{x}_i(t)) \\ &\quad + c_2\boldsymbol{r}_2 \otimes (\mathbf{g}(t) - \boldsymbol{x}_i(t)) \quad (52) \\ \boldsymbol{l}_i(t+1) &= w\boldsymbol{v}_i(t) + c_1\boldsymbol{r}'_1 \otimes (\boldsymbol{y}_i(t) - \boldsymbol{x}_i(t)) \\ &\quad + c_2\boldsymbol{r}'_2 \otimes (\hat{\boldsymbol{y}}_i(t) - \boldsymbol{x}_i(t)) \quad (53) \\ \boldsymbol{v}_i(t+1) &= u\boldsymbol{g}_i(t+1) + (1-u)\boldsymbol{l}_i(t+1) \quad (54) \\ \boldsymbol{x}_i(t+1) &= \boldsymbol{x}_i(t) + \boldsymbol{v}_i(t+1), \quad (55) \end{aligned}$$

where $r_{1,k}, r_{2,k}, r'_{1,k}, r'_{2,k} \sim U(0,1)$, and both $\mathbf{y}_i(t)$ and $\hat{\mathbf{y}}_i(t)$ are defined as before with the addition of $\mathbf{g}(t)$ as the global best position with the swarm at time step t.

Without loss of generality the particle and component indices are dropped again. In order to rewrite UPSO into the NIPSO form, substitute equations (52) and (53) into the velocity update equation (54) to arrive at

$$\begin{aligned} v(t+1) &= wv(t) + c_1(ur_1 + (1-u)r'_1)(y(t) - x(t)) \quad (56) \\ &\quad + c_2ur_2(g(t) - x(t)) + c_2(1-u)r'_2(\hat{y}_i(t) - x(t)). \end{aligned}$$

Now equation (56) is in the NIPSO form with $I = 3$ and $\theta_0 = w$, $\theta_1 = c_1(ur_1 + (1-u)r'_1)$, $\theta_2 = c_2ur_2$, and $\theta_3 = c_2(1-u)r'_2$. In order to calculate ψ the following additional terms are required:

$$\begin{aligned} E[\theta_1] &= c_1uE[r_1] + c_1(1-u)E[r'_1] \\ &= \frac{c_1u}{2} + \frac{c_1(1-u)}{2} = \frac{c_1}{2} \quad (57) \\ E[\theta_2] &= c_2uE[r_2] = \frac{c_2u}{2} \quad (58) \\ E[\theta_3] &= c_2(1-u)E[r'_2] = \frac{c_2(1-u)}{2}. \quad (59) \end{aligned}$$

The summation of equations (57), (58), and (59) leads to

$$\psi = \sum_{\imath=1}^{3} E[\theta_\imath] = \frac{c_1}{2} + \frac{c_2u}{2} + \frac{c_2(1-u)}{2} = \frac{c_1+c_2}{2}. \quad (60)$$

In order to calculate ψ the following additional terms are required:

$$\begin{aligned} V[\theta_1] &= c_1^2V[ur_1 + (1-u)r'_1] \\ &= c_1^2\left(u^2V[r_1] + (1-u)^2V[r_2] + 2u(1-u)COV[r_1, r'_2]\right) \\ &= c_1^2\left(\frac{u^2}{12} + \frac{(1-u)^2}{12}\right) = c_1^2\left(\frac{u^2+(1-u)^2}{12}\right) \quad (61) \end{aligned}$$

$$V[\theta_2] = V[c_2ur_2] = \frac{c_2^2u^2}{12} \quad (62)$$

$$V[\theta_2] = V[c_2(1-u)r'_2] = \frac{c_2^2(1-u)^2}{12}. \quad (63)$$

The summation of equations (61), (62), and (63) leads to

$$\phi = \sum_{\imath=1}^{3} V[\theta_\imath] = \frac{(c_1^2+c_2^2)\left(u^2+(1-u)^2\right)}{12}. \quad (64)$$

Table 2
Summary of the Order-1 and Order-2 stability criteria explicitly derived

Variant	Number of Informers	Number of Coefficients	Order-1 and Order-2 Criteria
CPSO	2	3	$\lvert w\rvert < 1$ and $0 < c_1 + c_2 < \frac{4(1-w^2)}{1-w+(c_1^2+c_2^2)(1+w)/(3(c_1+c_2)^2)}$
CLPSO	1	2	$\lvert w\rvert < 1$ and $0 < c < \frac{6(1-w^2)}{2-w}$
FIPS	$\lvert\mathcal{N}\rvert$	2	$\lvert w\rvert < 1$ and $0 < \frac{\hat{c}}{2} < \frac{12\lvert\mathcal{N}\rvert(1-w^2)}{3\lvert\mathcal{N}\rvert+1+w(1-3\lvert\mathcal{N}\rvert)}$
UPSO	3	4	$\lvert w\rvert < 1$ and $0 < c_1 + c_2 < \frac{4(1-w^2)}{1-w+(c_1^2+c_2^2)\left(u^2+(1-u)^2\right)(1+w)/(3(c_1+c_2)^2)}$

Substituting ψ and ϕ into the criteria of equations (10), (12), and (13) the following criteria for order-1 and order-2 stability are obtained:

$$-1 < w < 1 \tag{65}$$

$$0 < c_1 + c_2 < \frac{4\left(1-w^2\right)}{1 - w + \frac{(c_1^2+c_2^2)\left(u^2+(1-u)^2\right)(1+w)}{3(c_1+c_2)^2}}. \tag{66}$$

The criteria of equations (65) and (66) is the first derivation of full USPO stability criteria without artificial restrictions on the control coefficients. As with the CPSO case, in Section 5.1, the weighting between c_1 and c_2 has a clear influence on the size and shape of the stability region, as illustrated in Figure 3, where the cross-sections of the stability region, with fixed inertia values are shown.

In the restricted case where $c_1 = c_2$ is considered, the following criteria are obtained:

$$-1 < w < 1 \tag{67}$$

$$0 < c_1 + c_2 < \frac{24\left(1-w^2\right)}{7 - 5w + 2(u^2-u)(1+w)}. \tag{68}$$

which is in agreement with the derived criteria of Cleghorn and Engelbrecht [31] with minimal calculations needed, and under a weaker modelling assumption.

5.5. Derived Order-1 and Order-2 Criteria

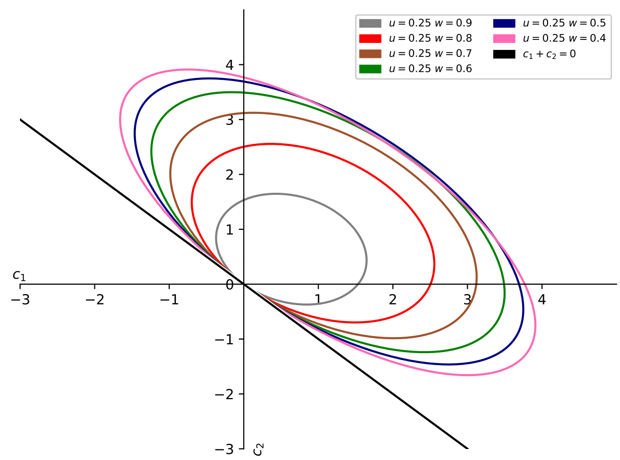
A compact summary of the order-1 and order-2 criteria explicitly derived in this section are presented in Table 2 for the reader's convenience.

6. Impact of Unstable Coefficients on performance

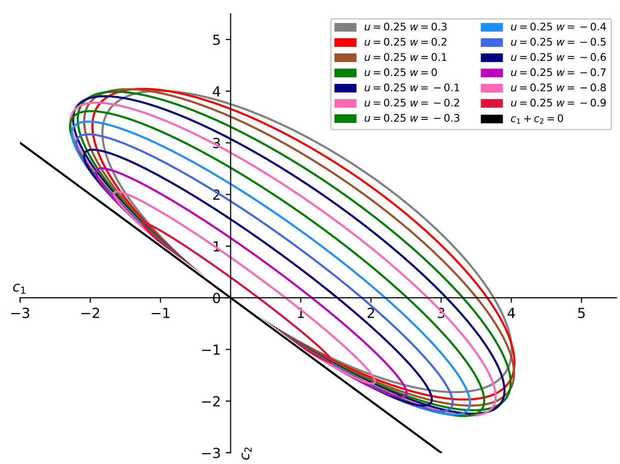
In this section a succinct empirical demonstration of how selecting coefficients that fail to meet the order-1 and order-2 stability criteria has a substantially negative effect on PSO performance is presented. One of the main findings of Cleghorn and Engelbrecht [15] was that the CPSO actually performed worse than random search for the vast majority of unstable coefficient choices [10]. In this section a similar test is done for the four PSO variants presented in Section 5.

In order to examine the effect of unstable parameter configurations on performance, a set of unstable parameter configurations is needed. The approach taken to construct this set for each PSO variant considered is as follows:

[10] The study of Cleghorn and Engelbrecht [15] only considered the $c_1 = c_2$ case



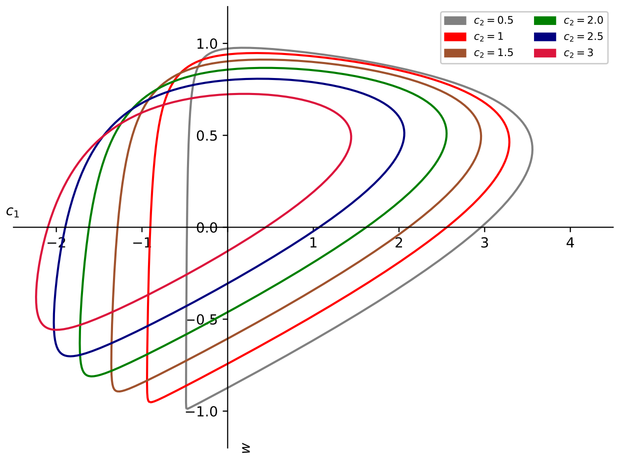
(a) Order-2 stable regions for $u = 0.25$ and $w = 0.4$ to 0.9.



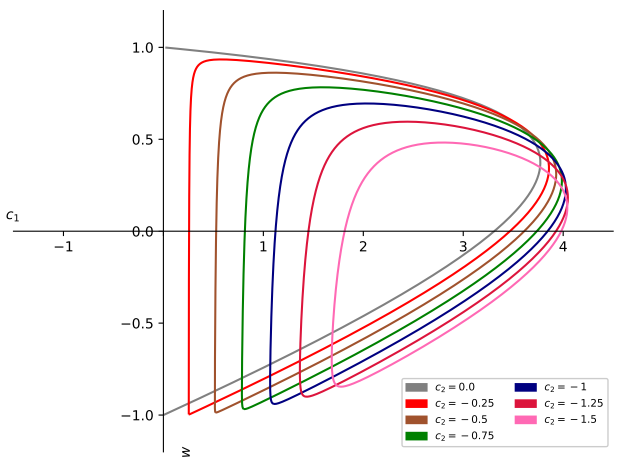
(b) Order-2 stable regions for $u = 0.25$ and $w = -0.9$ to 0.3.

Figure 3: Order-2 stable regions of UPSO under fixed inertia and unification values. The interior region of the elliptic shapes correspond to where equations (65) and (66) are satisfied for a given w and u.

- Let S be the set of coefficient choices that satisfy the order-1 and order-2 stability criteria of the variant.
- For each coefficient in the PSO variant, let l_i and u_i represent the infimum and supremum of the coefficient when satisfying the corresponding order-1 and



(a) Order-2 stable regions for $u = 0.25$ and $c_2 = 0.5$ to 3.



(b) Order-2 stable regions for $u = 0.25$ and $c_2 = 0$ to -1.5.

Figure 4: Order-2 stable regions of UPSO under fixed inertia and unification values. The interior region of the elliptic shapes correspond to where equations (65) and (66) are satisfied for a given u and c_2. The same visual pattern is observed if c_1 is fixed and c_2 is varied due to the symmetry present in (66).

order-2 stability criteria respectively[11].

- The set of unstable parameter configurations, US is defined as,

$$
\begin{aligned}
US =& C \setminus S \\
C =& [l_1 - d_1, u_1 + d_1] \times \cdots \times [l_p - d_p, u_p + d_p] \\
d_i =& (u_1 - l_1)/2
\end{aligned}
\tag{69}
$$

  where $\times$ is the Cartesian product, $\setminus$ is the standard set minus, and p is the number of coefficients in the variant.

The sizing of US is selected to provide a reasonably large coverage of possible unstable configurations without using coefficients that have unjustifiably large magnitudes, given their commonly used ranges. For all variants considered, all the originally recommended coefficient combinations are contained within $S \cup US$.

Table 3
Lower and upper bound on coefficient values for each variants that could satisfy order-1 and order-2 stability criteria . Note that these values should not be interpreted as stability criteria but merely the limits of what could satisfy them. Entries annotated with * were numerically derived.

Variant	Coefficients	Infimum	Supremum
CPSO	c_1, c_2	-1.69*	3.32*
	w	-1	1
CLPSO	c	0	$12(2-\sqrt{3})$
	w	1	-1
FIPS ($\lvert\mathcal{N}\rvert = 4$)	$\hat{c}$	0	$(6912 - 1632\sqrt{6})/605$
	w	1	-1
UPSO	c_1, c_2	-2.64*	4.44*
	w	-1	1

The CPSO, FIPS, UPSO, and CLPSO variants, for which stability criteria were derived are empirically examined. For each variant we uniformly sample 2000 configurations from the corresponding unstable set, US. The infimum and supremum of the coefficient value that could satisfy the corresponding order-1 and order-2 stability criteria is summarized in table 3. In the case of UPSO a minimum and maximum value of 0 and 1, respectively, is used for the unification factor to preserve algorithmic intent of the variant. Each sampled configuration is run 50 times on each of the 30 CEC-2014 Benchmark problems in 30-dimensions. Each variant is given a function evaluation budget of 10^5. Each variant is set to have a population size of 20 and the Von-Neumann topology is used when appropriate. A direct way to seeing failure of the PSO configuration is if the PSO performs worse than a trivial random search[12]. As such, for each CEC-2014 benchmark a random search is performed and repeated 50 times.

The intention is not to demonstrate which part of the stable region, S, that results in the best performance, but rather to illustrate the danger of utilizing unstable configurations. It should be noted that failure to perform better than random search is specifically utilized to demonstrate the severity of failure that unstable parameter configuration can induce.

The following procedure is followed for all considered variants. For each configuration, on each benchmark problem, a two-tailed Mann–Whitney U test is used to check if there is any significant difference between the PSO variant with a unstable configuration and the random search. If there is a difference, single-tailed versions are used to determine which performed statistically significantly better. A 0.95 confidence level is used. In Table 4 the obtained results are summarized, where the percentage of unstable configurations that performed better than, equal to, or worse than random search based on the aforementioned statistical test

[11] The infimum is the largest lower bound and the supremum is the smallest upper bound.

[12] In this paper a trivial random search is defined to be one where the search space is sampled uniformly at random until the function evaluation budget is meet and the best found candidate solution is used.

Table 4
Performance of PSO variant's run using unstable configurations compared to a trivial random search over the CEC-2014 benchmark suite.

Variant	Unstable Config Win	Draw	Random Win
CPSO	2.456%	5.088%	92.456%
CLPSO	1.654%	7.312%	91.035%
FIPS	1.629%	4.082%	94.290%
UPSO	2.103%	4.632%	93.265%

are presented. The reported results are averaged over the 30 benchmark problems.

As illustrated in Table 4 over 91% of unstable configurations perform statically significantly worse than a trivial random search for all variants considered. In general if a practitioner is using a meta-heuristic, like PSO, they would expect it to perform better than random search, which implies that a draw with a trivial random search is clearly a failing configuration. Over the four considered PSO variants, more than 97.5% of unstable configurations performed worse or equal to random search. The small fraction of configurations that performed better than random search were all just marginally over the relevant stability criteria boundaries.

One aspect that table 4 does not fully capture is the extremely poor objective function values that unstable configuration can actually result in. Specifically, on average, the objective function values obtained by unstable parameter configurations were found to be orders of magnitude worse than the stable configurations. As an illustrative example, consider the CEC2014's benchmark function, Rotated Discuss. Almost all stable configurations found the optimal solution (with the value of 0), whereas the vast majority of unstable configurations resulting in the the best found objective function values being over 10^6.

The findings presented in this section imply that failure to satisfy stability criteria results in a very high probability of algorithmic failure. It follows that being aware of a PSO variant's stability criteria is invaluable information when it comes to parameter tuning. Parameter tuning of a metaheuristics is often a necessary, but computationally expensive step, when trying to tackle challenging real world problems. This computational cost is driven by two primary factors; the underlying objective function's evaluation time, and the number of configurations considered. By knowing where not to look, namely at unstable configurations, significantly better parameter settings can be obtained, as resources are not wasted testing parameter configurations that will almost certainly fail to perform well.

7. Conclusion and Future Work

In this paper general theorems for rapidly obtaining order-1 and order-2 stability criteria and fixed points for a class of PSO variants are derived. Specifically, PSO variants that can be rearranged into a sum of difference vectors between informers and the current particle positions, are catered for. From this general derivation, stability criteria can be obtained for a set of custom PSO variants in a direct manner without substantial mathematical calculation. Given the direct link between PSO performance and the satisfaction of order-1 and order-2 stability criteria, the theorems provided in this paper will be directly applicable to the PSO community as a whole.

Furthermore, the proven theorems allow for stability criteria to be derived without unnecessary restrictions on the relationship between control coefficients. In this vein, stability criteria for the canonical PSO and, for the first time, the unified PSO are derived without restrictions on the relationship between control coefficients in this paper. The first stability order-1 and order-2 stability criteria for comprehensive learning PSO was also derived using the proven theorems.

One of the main lingering issues in PSO theory is the lack of a tractable mathematical approach to analyzing the few PSO variants with component-wise coupling, such as SPSO2011[35]. Which to date has required empirical approaches to analysis. It would be particularly impactful if the theorems presented in this paper could be extended to deal with this case in future work.

References

[1] J. Kennedy and R.C. Eberhart. Particle swarm optimization. In *Proceedings of the IEEE International Joint Conference on Neural Networks*, pages 1942–1948, Piscataway, NJ, 1995. IEEE Press.

[2] R. Poli. Analysis of the publications on the applications of particle swarm optimisation. *Journal of Artificial Evolution and Applications*, 2008:1–10, 2008.

[3] E. Ozcan and C.K. Mohan. Analysis of a simple particle swarm optimization system. *Intelligent Engineering Systems through Artificial Neural Networks*, volume 8:253–258, 1998.

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

[5] 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*, 102(1):8–16, 2007.

[6] V. Gazi. Stochastic stability analysis of the particle dynamics in the PSO algorithm. In *Proceedings of the IEEE International Symposium on Intelligent Control*, pages 708–713, Piscataway, 2012. IEEE Press.

[7] E. García-Gonzalo and J.L. Fernández-Martinez. Convergence and stochastic stability analysis of particle swarm optimization variants with generic parameter distributions. *Applied Mathematics and Computation*, 249:286–302, 2014.

[8] C.W. Cleghorn and A.P. Engelbrecht. A generalized theoretical deterministic particle swarm model. *Swarm Intelligence*, 8(1):35–59, 2014.

[9] Q. Liu. Order-2 stability analysis of particle swarm optimization. *Evolutionary Computation*, 23(2):187–216, 2015.

[10] J. Liu, X. Ma, T. Shi, and P. Li. Random convergence analysis of particle swarm optimization algorithm with time-varying attractor. *Swarm and Evolutionary Computation*, 61:100819, 2021.

[11] C.W. Cleghorn and A.P. Engelbrecht. Particle swarm stability: a theoretical extension using the non-stagnate distribution assumption. *Swarm Intelligence*, 12(1):1–22, 2018.

[12] N. Hansen. *The CMA evolution strategy: a comparing review. In J.A. Lozano, P. Larranaga, I. Inza, and E. Bengoetxea, editors, Towards a new evolutionary computation. Advances on estimation of distribution algorithms*. Springer, Berlin, 2006.

[13] C.W. Cleghorn. Particle swarm optimization: Understanding order-2 stability guarantees. In *Proceedings of the International Conference*

on the Applications of Evolutionary Computation, pages 535–549, Switzerland, 2019. Springer.

[14] Y. Shi and R.C. Eberhart. A modified particle swarm optimizer. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pages 69–73, Piscataway, NJ, 1998. IEEE Press.

[15] C.W. Cleghorn and A.P. Engelbrecht. Particle swarm optimizer: The impact of unstable particles on performance. In *Proceedings of the IEEE Symposium Series on Swarm Intelligence*, pages 1–7, Piscataway, NJ, 2016. IEEE Press.

[16] K.E. Parsopoulos and M.N. Vrahatis. UPSO: A unified particle swarm optimization scheme. In *Proceedings of the International Conference on Computational Methods in Sciences and Engineering*, pages 868–873, Netherlands, 2004. VSP International Science Publishers.

[17] J. Kennedy and R. Mendes. Neighborhood topologies in fully-informed and best-of-neighborhood particle swarms. In *Proceedings of the IEEE International Workshop on Soft Computing in Industrial Applications*, pages 45–50, Piscataway, NJ, 2003. IEEE Press.

[18] C Scheepers. *Multi-guided Particle Swarm Optimization: A Multi-objective Particle Swarm Optimizer*. Phd thesis, University of Pretoria, 2018.

[19] A. Meier and O. Kramer. Recurrent neural network-predictions for pso in dynamic optimization. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 29–36, New York, NY, 2018. ACM Press.

[20] R. Poli and D. Broomhead. Exact analysis of the sampling distribution for the canonical particle swarm optimiser and its convergence during stagnation. In *Proceedings of the Genetic and Evolutionary Computation Conference*, pages 134–141, New York, NY, 2007. ACM Press.

[21] T. Blackwell. A study of collapse in bare bones particle swarm optimzation. *IEEE Transactions on Evolutionary Computation*, 16(3):354–372, 2012.

[22] M.R. Bonyadi and Z. Michalewicz. Stability analysis of the particle swarm optimization without stagnation assumption. *IEEE Transactions on Evolutionary Computation*, 20(5):814–819, 2016.

[23] J.J. Liang, A.K. Qin, P.N. Suganthan, and S. Baskar. Comprehensive learning particle swarm optimizer for global optimization of multimodal functions. *IEEE Transactions on Evolutionary Computation*, 10(3):281–295, 2006.

[24] A.P. Piotrowski, J.J. Napiorkowski, and A.E. Piotrowska. Population size in particle swarm optimization. *Swarm and Evolutionary Computation*, 58:100718, 2020.

[25] F. Van den Bergh and A.P. Engelbrecht. A study of particle swarm optimization particle trajectories. *Information Sciences*, 176(8):937–971, 2006.

[26] I.C Trelea. The particle swarm optimization algorithm: Convergence analysis and parameter selection. *Information Processing Letters*, 85(6):317–325, 2003.

[27] V. Kadirkamanathan, K. Selvarajah, and P.J. Fleming. Stability analysis of the particle dynamics in particle swarm optimizer. *IEEE Transactions on Evolutionary Computation*, 10(3):245–255, 2006.

[28] R. Poli. Mean and variance of the sampling distribution of particle swarm optimizers during stagnation. *IEEE Transactions on Evolutionary Computation*, 13(4):712–721, 2009.

[29] B. Kisacanin and G.C. Agarwal. *Linear Control Systems: With Solved Problems and Matlab Examples*. Springer, New York, 2001.

[30] C.W. Cleghorn and A.P. Engelbrecht. Fully informed particle swarm optimizer: Convergence analysis. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pages 164–170, Piscataway, NJ, 2015. IEEE Press.

[31] C.W. Cleghorn and A.P. Engelbrecht. Unified particle swarm optimizer: Convergence analysis. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pages 448–454, Piscataway, NJ, 2016. IEEE Press.

[32] C.W. Cleghorn, C. Scheepers, and A.P. Engelbrecht. Stability analysis of the multi-objective multi-guided particle swarm optimizer. In *Proceedings of International Swarm Intelligence Conference (ANTS), Swarm Intelligence*, pages 201–212, Switzerland, 2018. Springer International Publishing.

[33] M. Kendall and J. Ord. *Time Series, 3rd edition*. Edward Arnold, London, UK, 1990.

[34] E. García-Gonzalo and J.L. Fernández-Martinez. Convergence and stochastic stability analysis of particle swarm optimization variants with generic parameter distributions. *Applied Mathematics and Computation*, 249:286–302, 2014.

[35] M. Zambrano-Bigiarini and M Clerc. Standard particle swarm optimisation 2011 at CEC-2013: A baseline for future PSO improvements. In *Proceedings of the IEEE Congress on Evolutionary Computation*, pages 2337–2344, Piscataway, NJ, 2013. IEEE Press.