

ON ENHANCING EFFICIENCY AND ACCURACY OF PARTICLE SWARM OPTIMIZATION ALGORITHMS

SILVANO CHIARADONNA, FELICITA DI GIANDOMENICO AND NADIR MURRU

ISTI – CNR

Via G. Moruzzi 1, 56124, Pisa, Italy

{silvano.chiaradonna; f.digiandomenico; nadir.murru}@isti.cnr.it

Received August 2014; revised December 2014

ABSTRACT. *The particle swarm optimization (PSO) algorithm has been recently introduced in the non-linear programming, becoming widely studied and used in a variety of applications. Starting from its original formulation, many variants for improvement and specialization of the PSO have been already proposed, but without any definitive result, thus research in this area is nowadays still rather active. This paper goes in this direction, by proposing some modifications to the basic PSO algorithm, aiming at enhancements in aspects that impact the efficiency and accuracy of the optimization algorithm. In particular, variants of PSO based on fuzzy logics and Bayesian theory have been developed, which show better, or competitive, performances compared to both the basic PSO formulation and a few other optimization algorithms taken from the literature.*

Keywords: Fuzzy logics, Kalman filter, Non-linear programming, Particle swarm optimization

1. **Introduction.** Many techniques, mainly based on metaheuristic algorithms, have been developed for solving non-linear programming. In particular, chaotic algorithms [36], evolutionary programming [42, 44], genetic algorithms [3, 46], tabu search algorithm [21, 23], have been widely used in these studies. Recently, the particle swarm optimization (PSO) algorithm [20] has been introduced in the non-linear programming becoming widely studied and used in a variety of applications [12, 19, 47].

The PSO algorithm performs a metaheuristic search based on competition and cooperation among particles (which represent the search variables) belonging to an initial swarm or population that covers the search space. After the basic technique was defined, numerous proposals for improvement and specialization of the PSO have appeared (refer [33] for a partial survey). Still, further investigations are nowadays of interest and solicited in PSO, to better respond to application needs in a growing variety of sectors, with related emphasis on different aspects, such as efficiency and accuracy.

In this paper we propose some modifications to the basic PSO algorithm, aiming at enhancements in the following three aspects that determine the efficiency and accuracy of the optimization algorithm:

- the setting of the initial population, based on the partitioning of the search space in disjoint intervals, one for each particle, such that the initial position for each different particle is selected from a different interval,
- the setting of the acceleration constants and of the inertia weight parameter, defined by using a fuzzy logics based strategy,
- the calculation of the current position of each particle, based on a Bayesian approach.

By adopting approaches very suited to support the dynamics of the PSO algorithm, such as the fuzzy logics and the Bayesian theory, this paper offers novel solutions which improve

on the basic PSO formulation, thus potentially resulting more appealing to be employed, as suggested by the evaluation studies already performed. Although the literature on this family of optimization algorithms is very abundant, as already recalled, our goal is to enrich the set of alternatives to enhance the choice of the best solution that fits specific applications needs and characteristics.

The rest of the paper is structured as follows. A brief recall of the basic PSO algorithm is presented in Section 2. In Section 3, the motivation and innovative aspects of the three proposed modifications to the basic PSO algorithm are discussed. In the next three sections, we develop our enhancements to the basic PSO algorithm. In particular, the setting proposed for the initial population is described in Section 4; the approach proposed to set the acceleration constants and the inertia weight is described in Section 5, while Section 6 contains the new method to move the control variables in the search space. By composing the proposed improvements, two variants of the PSO algorithm are assembled in Section 7, which are then evaluated and compared both with the basic PSO algorithm in terms of indicators representative of their accuracy and efficiency, and with other solutions taken from the literature when employed to support power flow optimization in a simple but realistic case study in the electrical power sector. Finally, Section 8 is devoted to the conclusions.

2. Overview of the Basic PSO Algorithm. In this section, the basic PSO algorithm is briefly described, to provide the context for the variants that will be presented in the next sections. In general terms, an optimization problem is the problem of minimizing (or maximizing) an objective function $f(x_1, \dots, x_d, y_1, \dots, y_s)$, where x_i are the search variables, representing the parameters of the problem that are to be optimized, and y_i are the state variables, representing the parameters that can be derived as function of the search variables. The variables x_i and y_i can be bounded and related together by means of constraint equations or disequations. For the sake of simplicity, in the following, we omit variables y_i from the notation of the objective function and we only consider search variables x_i .

The PSO algorithm is a metaheuristic global optimization method which tries to find those values of the search variables x_1, \dots, x_d for which the value of the cost function $f(x_1, \dots, x_d)$ is minimum (or maximum). The search is based on a discrete-time stochastic process describing the competition and cooperation between a population or swarm of particles that move in a search-variables space. The PSO algorithm considers a fixed size swarm of D particles that randomly change position with velocity $\mathbf{v}_p(k)$ at each discrete instant of time (step) k , according to their own experience and that of their neighbors. The position of each particle p at step k is a vector $\mathbf{x}_p(k) = (x_{p1}(k), \dots, x_{pd}(k))$. The solution of the optimization problem found by PSO is the position \mathbf{b} , reached at a certain step by a particle for the swarm, for which the value of the cost function is minimum, i.e., $f(\mathbf{b}) = \min\{f(\mathbf{x}_p(k)) | p = 1, \dots, D, k = 1, 2, \dots\}$. Algorithm 1 describes formally in pseudo-code the basic PSO procedure of finding the optimal vector \mathbf{b} , as extracted from [20].

The starting velocity and position of the particles at step 0 are initialized at lines 5 and 6, respectively. The personal best position \mathbf{l}_p for the particle p at step k is the best position the particle p has visited since the first step, i.e., the position that minimizes the objective function among all the positions of p for different values of k . It is updated at line 12. The global best position \mathbf{b} at each step k is the best position discovered by any of the particles so far. It is usually calculated as the personal best position and it is obtained as the personal best position at step k that minimizes the objective function, as shown at line 13, where \mathbf{b} is updated.

Algorithm 1: Basic PSO algorithm

```

1 Data:  $d, D, w, c_1, c_2, k_{\max}$  and bounds  $\mathbf{x}^{\min}, \mathbf{x}^{\max}$ 
2 Search variables  $\mathbf{x} = (x_1, \dots, x_d)$  and cost function  $f(\mathbf{x})$ 
3 Result:  $\mathbf{b}, f(\mathbf{b})$ 
4 for  $p = 1, \dots, D$  do // for each particle
5    $\mathbf{v}_p(0) \leftarrow \mathbf{0}$  // initial velocity
6    $\mathbf{x}_p(0) \leftarrow \text{Rnd}(\mathbf{x}^{\min}, \mathbf{x}^{\max})$  // initial random position
7    $\mathbf{l}_p, \mathbf{b} \leftarrow \mathbf{x}_p(0)$ 
8 begin
9   for  $k = 1, \dots, k_{\max}$  do // for each step
10    for  $p = 1, \dots, D$  do // for each particle
11      if  $f(\mathbf{x}_p(k-1)) < f(\mathbf{l}_p)$  then
12         $\mathbf{l}_p \leftarrow \mathbf{x}_p(k-1)$  // new personal best position
13      if  $f(\mathbf{l}_p) < f(\mathbf{b})$  then
14         $\mathbf{b} \leftarrow \mathbf{l}_p$  // new global best position
15      for  $p = 1, \dots, D$  do // for each particle
16         $r_1 \leftarrow \mathcal{U}(0, 1)$  // new uniform random value
17         $r_2 \leftarrow \mathcal{U}(0, 1)$  // new uniform random value
18         $\mathbf{v}_p(k) = w\mathbf{v}_p(k-1) + c_1r_1(\mathbf{l}_p - \mathbf{x}_p(k-1)) + c_2r_2(\mathbf{b} - \mathbf{x}_p(k-1))$ 
19         $\mathbf{x}_p(k) = \mathbf{x}_p(k-1) + \mathbf{v}_p(k)$ 
20        for  $i = 1, \dots, d$  do // for each search variable
21          if  $x_{pi}(k) < x_i^{\min}$  then
22             $x_{pi}(k) \leftarrow x_i^{\min}$ 
23          else if  $x_{pi}(k) > x_i^{\max}$  then
24             $x_{pi}(k) \leftarrow x_i^{\max}$ 

```

The position $\mathbf{x}_p(k)$ of the particle p at step k is updated at line 19, based on the random value assigned to the velocity $\mathbf{v}_p(k)$, as shown from line 16 to line 18. The velocity vector drives the optimization process and results from the sum of three different components, as shown at line 18: the momentum component, which is the previous velocity, the cognitive component, which is proportional to the distance of the particle from the best position it has ever visited, and, finally, the social component which is proportional to the particle's distance from the best position where any of the swarm's particles has ever been.

The PSO algorithm has some critical points that heavily influence its performances, like the initial position of the particles and the choice of some parameters, such as inertia weight w and the acceleration constants c_1 and c_2 , used to scale the contribution of the moment to the velocity, and of the cognitive and social components, respectively [11]. Indeed, if the initial position of the particles is a well-distributed cover of the search space, then there is a higher probability to avoid local best position and to approach the global optimum. Moreover, parameters w, c_1, c_2 heavily influence the velocity of the particles, i.e., their movements in the search space.

In the next sections we present and evaluate three variants of the PSO algorithm, aiming at enhancements in the aspects just recalled that determine the efficiency and accuracy of the optimization algorithm.

3. Three New Variants of the PSO Algorithm: Motivation and Innovative Aspects. In this section, we discuss the motivation and the innovative aspects of the three proposed variants for the PSO algorithm.

The efficiency of PSO is influenced by the initial position of the particles, i.e., by how well particles are distributed over the search space, since if regions of the search space are not covered by the initial positions, the PSO will have difficulty in finding the optimum if it is located within an uncovered region [11]. Usually, the initial position of each particle is randomly generated, with uniform distribution, from the set of all possible positions of the particles. In this case, it is important to use a good pseudo-random generator; otherwise the performances heavily degrade. In [27], the authors propose to reinitialize the particle that has the worst performance (i.e., corresponding to the biggest objective function value). In [32], re-initialization of a certain percentage of the total population is considered if after some steps there is no significative improvement. In [47], an orthogonal design is proposed in order to sample the initial positions. In [45], the authors determine the initial position of the particles by means of the Tent chaotic map. In Section 4, we propose an original algorithm that forces a uniform distribution for the position of the initial particles, dividing the search space into small intervals where we randomly generate the initial positions.

As discussed in the previous section, the values of parameters w , c_1 , c_2 affect the performance of the PSO algorithm. Standard values for the parameters of the PSO algorithm are $c_1 = c_2 = 1.496172$ and $w = 0.72984$ [5]. However, in many works (like [4], [17], [28] and [32]) values of c_1 , c_2 , w are not constant, but they change during the execution of the algorithm. This is mainly due to the fact that at the initial steps of the algorithm large values for w are preferable so that particles have a greater possibility of movement [31]. On the contrary, small values for w are preferable towards the final steps of the algorithm. Moreover, other different components could be chosen for modifying the values of w , c_1 , c_2 , like, e.g., the distance between the current particles and the best particles. Since, there are no defined rules to determine the relationship between parameters w , c_1 , c_2 and other components (such as the current step of the algorithm), in Section 5, we use fuzzy logic in order to manage these relationships. Indeed, fuzzy logics is congenial to capture and to code expert-based knowledge in view of performing targeted simulations. Usually the fuzzy logic-based systems are tuned using heuristic criteria (see, e.g., [7, 30]). Fuzzy logic has been already used in various ways for improving the PSO algorithm [1, 29]. In [37], a fuzzy system is defined to dynamically adapt the inertia weight of the PSO. It is based on two input variables, which measure the current best performance and the current inertia weight, and on an output variable, which is the change of the inertia weight. Our approach differs from previous usages of fuzzy logics (as in [29, 37]) in the choice of the input fuzzy variables, in the tuning of the membership functions and in the fuzzy rules that combine input/output fuzzy variables together with their membership functions.

Finally, in Section 6 we propose an innovative method to update the position of the particles at each step of the algorithm. The core idea of the basic PSO algorithm is to move the particles around the search space using information provided by personal and global best positions as shown in Algorithm 1. Here, we propose to move the particles in the search space by balancing the current position of each particle with the current personal and global best positions by means of equations based on the Kalman filter [18]. The Kalman filter uses a Bayesian approach providing a posterior estimation for a system's state from the current measurement and the prior estimation of that system's state by means of the Bayes theorem. In typical applications of the Kalman filter, it takes a sequence of measurements over time, containing noise (random variations) and other

inaccuracies, and produces statistically optimal estimates of unknown variables. In our context, such a Bayesian approach has been reproduced considering the current position of the particles and the current global best position as the current measurement and the prior estimation, respectively. The posterior estimation is the new global best position of the particles that should be an improved estimation for the position of the unknown global best.

4. Variant 1: Focus on the Initial Position of the Particles. Here, we propose an easy sampling method that guarantees a good uniform distribution of the initial position of the particles in the search space, thus improving the use of pseudo-random generators.

It is based on the partitioning of each dimension of the search space in disjoint intervals, one for each particle, such that the initial position of each different particle is randomly selected for each dimension from a different random interval.

Let x_i^{\min} and x_i^{\max} be the lower and upper bound of each search variable, i.e.,

$$x_i^{\min} \leq x_{pi}(k) \leq x_i^{\max}, \quad \forall p = 1, \dots, D, \quad \forall i = 1, \dots, d, \quad \forall k.$$

Each interval $[x_i^{\min}, x_i^{\max}]$ is partitioned into D disjoint subintervals, one for each particle p

$$\left[x_i^{\min} + (p - 1) \frac{x_i^{\max} - x_i^{\min}}{D}, x_i^{\min} + p \frac{x_i^{\max} - x_i^{\min}}{D} \right], \quad \forall p = 1, \dots, D.$$

Algorithm 2: Initial position of the particles for the PSO algorithm

```

1 Data:  $d, D, x_i^{\min}, x_i^{\max}$ , for  $i = 1, \dots, d$ 
2 Result:  $x_{pi}(0)$  for  $p = 1, \dots, D$  and  $i = 1, \dots, d$ 
3 begin
4   for  $i = 1, \dots, d$  do                                     // for each search variable
5      $\rho \leftarrow$  random permutation of  $(1, \dots, D)$ 
6     for  $p = 1, \dots, D$  do                                     // for each particle
7        $x_{pi}(0) \leftarrow \mathcal{U} \left( x_i^{\min} + (\rho(p) - 1) \frac{x_i^{\max} - x_i^{\min}}{D}, x_i^{\min} + \rho(p) \frac{x_i^{\max} - x_i^{\min}}{D} \right)$ 

```

The initial position $x_{pi}(0)$ of each different particle p for each dimension i is uniformly randomly selected from a randomly selected different subinterval, as shown in Algorithm 2. At line 7 it is shown that the position of the particle p is selected by the $\rho(p)$ -th subinterval, where $\rho()$ is the random permutation function of the indexes of the particles, as shown at line 5. Using this approach, the location of each different particle is selected for each dimension (search variable) from different partitions of the range of values for that dimension, improving the distribution of the particles in the search space. This approach also protects against weaknesses of bad pseudo-random generators, which might result in a poor cover of the search space.

5. Variant 2: Fuzzy Logic-based Inertia Weight and Acceleration Constants.

Usually, as shown in [5], the following relation among w, c_1, c_2 holds:

$$w = \frac{2}{|2 - \phi - \sqrt{\phi^2 - 4\phi}|}, \quad \phi = w(c_1 + c_2),$$

from which, using equal values for c_1 , c_2 , we obtain

$$c_1 = c_2 = \frac{(w + 1)^2}{2}. \quad (1)$$

Some works, like [28], propose a varying value for w :

$$w = \frac{k_{\max} - k}{k_{\max}}(w_{\max} - w_{\min}) + w_{\min},$$

where k is the current step of the algorithm, k_{\max} is the maximum number of steps, and usually $w_{\min} = 0.8$ and $w_{\max} = 1.2$. In [32] and [48] different formulas for w have been proposed and in [17] similar formulas have been used for the coefficients c_1 and c_2 . In [4], a good survey on the problem of determining the parameters w , c_1 , c_2 is provided.

Here, we propose a novel fuzzy strategy to determine the parameters w , c_1 , c_2 based on the steps number and the current particle position. Generally, a fuzzy strategy is composed by

- fuzzy variables (input and output), fuzzy sets and membership functions,
- fuzzy rules that relate input and output variables,
- a fuzzy inference engine that combines the fuzzy rules,
- a defuzzification method that provides an output value.

Our strategy is based on a 2-input/1-output inference scheme, with the fuzzy variables k , α and w . The output is the value of w . The first input is the current step k of the algorithm. The second input is the percentage distance

$$\alpha = \frac{(f(\mathbf{x}_p(k)) - f(\mathbf{b})) \cdot 100}{f(\mathbf{b})}$$

between the objective function $f(\mathbf{x}_p(k))$ evaluated on the current position of the particle p at step k and the objective function $f(\mathbf{b})$ evaluated on the global best position at step k . When α is small, it may be convenient to move the particle around its current position, corresponding to a low value of w . On the contrary, when α is large it may be sensible moving the particle far from its current position, corresponding to a high value of w .

During the execution of the PSO algorithm, the value of w can be evaluated at each step k in the following way.

A membership degree is assigned to the current step k , as shown in Figure 1(a) for different fuzzy sets: VeryShort, Short, Moderate, Long, VeryLong. A fuzzy set characterizes the current step k in terms of the membership degree of k , according to the following rules (the choice of values is expert based):

1. if $k \leq \frac{4k_{\max}}{20}$, then current step k is VeryShort,
2. if $\frac{2k_{\max}}{20} \leq k \leq \frac{6k_{\max}}{20}$, then current step k is Short,
3. if $\frac{5k_{\max}}{20} \leq k \leq \frac{17k_{\max}}{20}$, then current step k is Moderate,
4. if $\frac{14k_{\max}}{20} \leq k \leq \frac{18k_{\max}}{20}$, then current step k is Long,
5. if $k \geq \frac{17k_{\max}}{20}$, then current step k is VeryLong.

For example, looking at Figure 1(a) if the value of k is less than or equal to $\frac{k_{\max}}{20}$, then k belongs to the fuzzy set VeryShort with membership degree 1. If $\frac{2k_{\max}}{20} \leq k \leq \frac{4k_{\max}}{20}$, then k belongs to the fuzzy sets VeryShort and Short with membership degrees given by the ordinates of the corresponding curves.

Similarly, a membership degree is assigned to α , as shown in Figure 1(b), according to the following rules:

1. if $\alpha \leq 10\%$, then distance is Small,
2. if $5\% \leq \alpha \leq 65\%$, then distance is Medium,

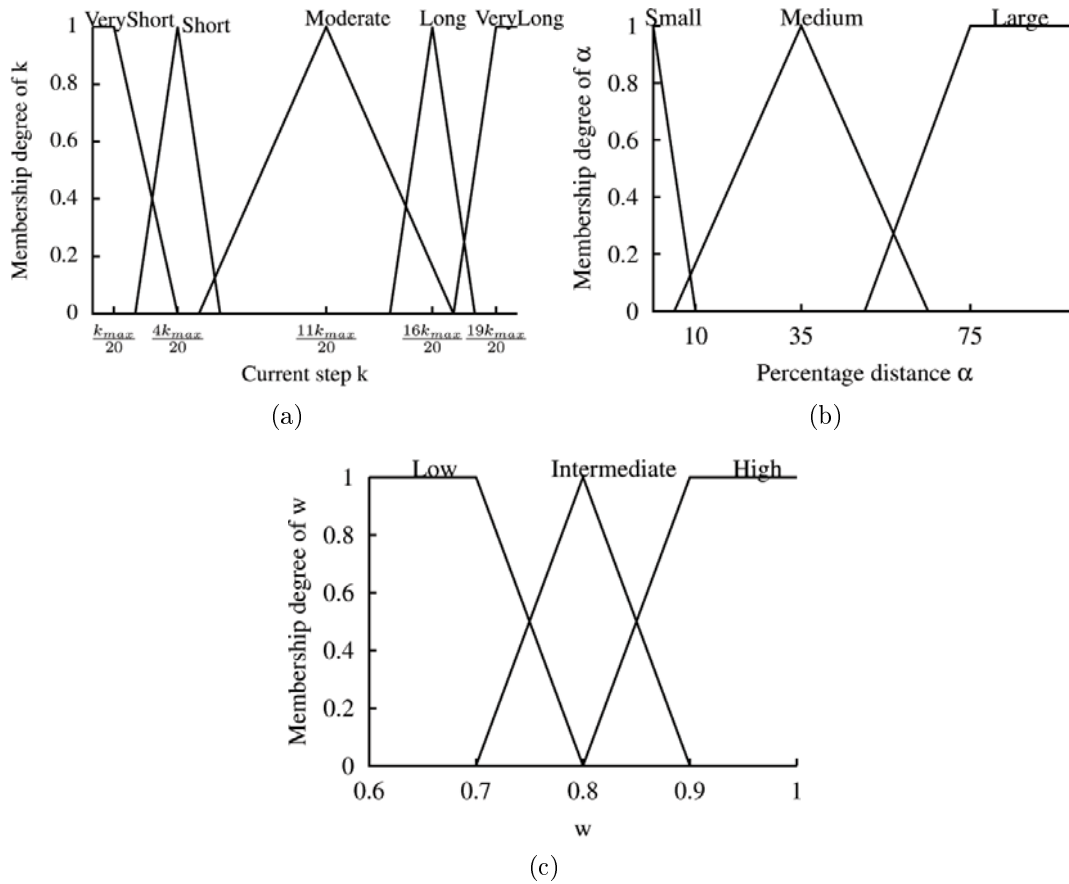


FIGURE 1. Membership degrees of the current step k , as a function of the value of k , for different fuzzy sets: VeryShort, ..., VeryLong (a), membership degrees of the percentage distance α , as a function of the value of α , for different fuzzy sets: Small, Medium, Large (b) and membership degrees of w for different fuzzy sets: Low, Intermediate, High (c).

3. if $\alpha \geq 50\%$, then current step is Large,

where Small, Medium and Large are the labels of the fuzzy sets that characterize the distance variable α . Let us observe that if $\alpha < 0$, then we set $\alpha = 0$ as input. Indeed, in this case the current particle has improved the global best and consequently we would like to perform the search around this position giving to α the maximum membership degree 1 at the fuzzy set Small.

Finally, the membership degree of w is obtained, as shown in Figure 1(c), according to the following rules (the choice of values is expert based):

1. if $0.6 \leq w \leq 0.8$, then distance is Low,
2. if $0.7 \leq w \leq 0.9$, then distance is Intermediate,
3. if $0.8 \leq w \leq 1$, then current step is High,

where Low, Intermediate and High are the labels of the fuzzy sets related to w .

The derived membership functions (depicted by triangular or trapezoidal shapes) reflect expert-based choices. The inference system, used by the inference engine to derive the value of w , is based on the eleven rules shown in Table 1.

The inference engine is the basic Mamdani model [26] with if-then rules, minimax set-operations, sum for composition of activated rules, and defuzzification based on the

TABLE 1. Inference system to derive the value of w

k	α	w
VeryShort	Small	Intermediate
VeryShort	Medium or Large	High
Short	Small	Low
Short	Medium or Large	High
Moderate	Small	Low
Moderate	Medium	Intermediate
Moderate	Large	High
Long	Small	Low
Long	Medium or Large	Intermediate
VeryLong	Small or Medium	Low
VeryLong	Large	Intermediate

centroid method that provides the abscissa of the barycentre of the fuzzy set composed according to the activated rules.

Finally, the values of c_1 and c_2 can be derived as a function of w from (1).

6. Variant 3: Bayesian Approach to Change the Position of Particles. In this section, we propose a new strategy to change the position of the particles with respect to the PSO algorithm shown in Algorithm 1. Specifically, in Subsection 6.1 we derive the new algorithm based on the Kalman filter, and in Subsection 6.2 we describe the new parameters of the algorithm and we analyze how they impact the velocity of the particles.

6.1. Bayesian PSO algorithm. In [8], the authors highlighted that the Kalman filter can be taken back to the Bayes theorem, considering random variables with normal distributions.

Let $L|X$, X and $X|L$ be random variables denoting, in terms of the Kalman filter, respectively, the current measurement (conditional on unknown parameter X), the prior estimation and the posterior estimation of the system's state. The Bayes theorem states

$$f_{X|L}(x|l) = \frac{f_{L|X}(l|x)f_X(x)}{\int_{-\infty}^{+\infty} f_{L|X}(l|u)f_X(u)du},$$

where $f_{X|L}$ is the posterior density, f_X the prior density, $f_{L|X}$ the likelihood and the denominator is a normalization factor.

Assume that $L|X$ and X have the normal distributions $\mathcal{N}(\mu_{L|X}, \sigma_{L|X}^2)$ and $\mathcal{N}(\mu_X, \sigma_X^2)$, respectively. Then, it is well-known that the posterior estimation $X|L$ has the normal distribution $\mathcal{N}(\mu_{X|L}, \sigma_{X|L}^2)$, where

$$\mu_{X|L} = \frac{\sigma_{L|X}^2 \mu_X + \sigma_X^2 \mu_{L|X}}{\sigma_{L|X}^2 + \sigma_X^2}, \quad \sigma_{X|L}^2 = \left(\frac{1}{\sigma_{L|X}^2} + \frac{1}{\sigma_X^2} \right)^{-1}.$$

Now, we define the new position of the particles at step k , for $k \geq 1$, of the PSO algorithm, in terms of the Kalman filter.

For each particle $p = 1, \dots, D$ and for each dimension $i = 1, \dots, d$, we consider that:

- the global best position $b_i(k-1)$ (for the sake of brevity denoted as b_i) is the current measurement $B|Y$,
- the position $x_{pi}(k-1)$ is the prior estimation $Y = X|L$, and
- the new position $x_{pi}(k)$ is the posterior estimation $Y|B$.

The prior estimation Y is obtained by the posterior estimation $X|L$ of the new position $x_{pi}(k)$, considering that:

- the personal best position $l_{pi}(k - 1)$ (for the sake of brevity denoted as l_{pi}) is the current measurement $L|X$,
- the position $x_{pi}(k - 1)$ is the prior estimation X , and
- the new position $x_{pi}(k)$ is the posterior estimation $X|L$.

Thus, observing that $L|X \sim \mathcal{N}(l_{pi}, \sigma_{L|X}^2)$ and $X \sim \mathcal{N}(x_{pi}(k - 1), \sigma_{pi}^2(k - 1))$, the posterior density of the random variable $Y = X|L$ is the normal distribution $\mathcal{N}(\mu_{X|L}(k), \sigma_{X|L}^2(k))$ derived by the Bayes theorem as

$$\mu_{X|L}(k) = \frac{\sigma_{L|X}^2 x_{pi}(k - 1) + \sigma_{pi}^2(k - 1) l_{pi}}{\sigma_{L|X}^2 + \sigma_{pi}^2(k - 1)},$$

$$\sigma_{X|L}^2(k) = \left(\frac{1}{\sigma_{L|X}^2} + \frac{1}{\sigma_{pi}^2(k - 1)} \right)^{-1}.$$

Finally, observing that $B|Y \sim \mathcal{N}(b_i, \sigma_{B|Y}^2)$ and $Y \sim \mathcal{N}(\mu_{X|L}(k), \sigma_{X|L}^2(k))$, the posterior density of the random variable $Y|B$ is the normal distribution $\mathcal{N}(\mu_{Y|B}(k), \sigma_{Y|B}^2(k))$ derived by the Bayes theorem as

$$\mu_{Y|B}(k) = \frac{x_{pi}(k - 1) + \delta_{L|X}(k - 1) l_{pi} + \delta_{B|Y}(k - 1) b_i}{1 + \delta_{L|X}(k - 1) + \delta_{B|Y}(k - 1)}, \tag{2}$$

$$\sigma_{Y|B}^2(k) = \frac{\sigma_{pi}^2(k - 1)}{1 + \delta_{L|X}(k - 1) + \delta_{B|Y}(k - 1)}, \tag{3}$$

where

$$\delta_{L|X}(k - 1) = \frac{\sigma_{pi}^2(k - 1)}{\sigma_{L|X}^2} \quad \text{and} \quad \delta_{B|Y}(k - 1) = \frac{\sigma_{pi}^2(k - 1)}{\sigma_{B|Y}^2}. \tag{4}$$

Using this Bayesian approach to change the position of the particles, the value for $x_{pi}(k)$, at each step k , for each particle p and each dimension i , is derived from the random variable $Y|B$ with normal distribution $\mathcal{N}(\mu_{Y|B}(k), \sigma_{Y|B}^2(k))$ as

$$x_{pi}(k) = Y|B. \tag{5}$$

The mean $\mu_{Y|B}(k)$ is derived from Equation (2) replacing recursively $x_{pi}(k - 1)$ by the value of $\mu_{Y|B}(k - 1)$ (the mean of $Y|B$ at previous step $k - 1$). The variance $\sigma_{Y|B}^2(k)$ is derived from Equation (3), replacing recursively $\sigma_{pi}^2(k - 1)$ by the value of $\sigma_{Y|B}^2(k - 1)$ (the variance of $Y|B$ at previous step $k - 1$), when the position $\mathbf{x}_p(k - 1)$ is the new global best, i.e., $f(\mathbf{x}_p(k - 1)) < f(\mathbf{b})$, otherwise, when $f(\mathbf{x}_p(k - 1)) \geq f(\mathbf{b})$, $\sigma_{Y|B}^2(k)$ is considered equal to $\sigma_{Y|B}^2(k - 1)$.

In this way, considering $h(k)$ the number of steps with a new global best until the step k , for $k \geq 1$, i.e.,

$$h(k) = \begin{cases} h(k - 1) + 1 & \text{if } f(\mathbf{x}_p(k - 1)) < f(\mathbf{b}) \\ h(k - 1), & \text{otherwise,} \end{cases} \tag{6}$$

with $h(0) = 0$, then the closed formulas of $\sigma_{Y|B}^2(k)$, $\delta_{L|X}(k)$ and $\delta_{B|Y}(k)$ for $k \geq 1$ are given by

$$\sigma_{Y|B}^2(k) = \frac{\sigma_{pi}^2(0)}{1 + h(k)(\delta_{L|X}(0) + \delta_{B|Y}(0))} = \frac{\sigma_{pi}^2(0)}{1 + h(k) \left(\frac{\sigma_{pi}^2(0)}{\sigma_{L|X}^2} + \frac{\sigma_{pi}^2(0)}{\sigma_{B|Y}^2} \right)}, \tag{7}$$

$$\delta_{L|X}(k) = \frac{\delta_{L|X}(0)}{1 + h(k)(\delta_{L|X}(0) + \delta_{B|Y}(0))} = \frac{\sigma_{pi}^2(0)}{\sigma_{L|X}^2 + h(k)\sigma_{pi}^2(0) \left(1 + \frac{\sigma_{L|X}^2}{\sigma_{B|Y}^2} \right)}, \tag{8}$$

$$\delta_{B|Y}(k) = \frac{\delta_{B|Y}(0)}{1 + h(k)(\delta_{L|X}(0) + \delta_{B|Y}(0))} = \frac{\sigma_{pi}^2(0)}{\sigma_{B|Y}^2 + h(k)\sigma_{pi}^2(0) \left(1 + \frac{\sigma_{B|Y}^2}{\sigma_{L|X}^2} \right)}. \tag{9}$$

The function $h(k)$ is monotone non-decreasing and $h(k) \leq k$.

The dimension i of the velocity of the particle p at step k , denoted by $v_{pi}(k)$, is derived as the difference of two normal random variables $x_{pi}(k) - x_{pi}(k-1)$ (as shown in Appendix A), that is a normal random variable with distribution $\mathcal{N}(\mu_V(k), \sigma_V^2(k))$ [41], where

$$\begin{aligned} \mu_V(k) &= \mu_{Y|B}(k) - \mu_{Y|B}(k-1) \\ &= \frac{\delta_{L|X}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}(l_{pi}(k-1) - x_{pi}(k-1)) \\ &\quad + \frac{\delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}(b_i(k-1) - x_{pi}(k-1)), \end{aligned} \tag{10}$$

$$\begin{aligned} \sigma_V^2(k) &= \sigma_{Y|B}^2(k) + \sigma_{Y|B}^2(k-1) \\ &= \frac{(2 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1))\sigma_{pi}^2(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}. \end{aligned} \tag{11}$$

In (10) the velocity $v_{pi}(k)$ results from the sum of the cognitive and social components of the PSO. Thus, as soon as the current position of a particle converges to the current personal and to the current global best position, the mean of the velocity $\mu_V(k)$ becomes 0 and the expected new position of the particle tends to be equal to the current one. In this case, although the mean of the new velocity is 0, a new position, that can be also significantly different from the current position, can occur with a probability that depends on the value of the variance $\sigma_{Y|B}^2(k-1)$.

A different formula for $\mu_V(k)$ can be derived from (10) (as described in Appendix A), showing that the velocity $v_{pi}(k)$ can result from the sum of three different components (the momentum, the difference between the local best positions of the last two steps and the difference between the global best positions of the last two steps), obtaining

$$\begin{aligned} \mu_V(k) &= \frac{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}\mu_V(k-1) \\ &\quad + \frac{\delta_{L|X}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}(l_{pi}(k-1) - l_{pi}(k-2)) \\ &\quad + \frac{\delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}(b_i(k-1) - b_i(k-2)). \end{aligned} \tag{12}$$

Equation (12) shows that, when the mean $\mu_V(k)$ of the velocity of a particle is greater than 0, the particle changes direction with respect to the current velocity (moment or inertia component) only when new current personal or global best positions occur. Otherwise,

when personal or global best positions do not change then the magnitude of new velocity is obtained as a fraction of the current velocity.

When the i -th dimension of the research space is bounded by x_i^{\min} and x_i^{\max} , the value $x_{pi}(k)$ is obtained from the random variable $Y|B$ with normal distribution $\mathcal{N}(\mu_{Y|B}(k), \sigma_{Y|B}^2(k))$ defined in (2) and (3), using the formula

$$x_{pi}(k) = x_i^{\min} + (Y|B - \bar{x}^{\min}) \frac{x_i^{\max} - x_i^{\min}}{\bar{x}_i^{\max} - \bar{x}_i^{\min}}, \tag{13}$$

where

$$\bar{x}_i^{\min} = \min\{x_i^{\min}, \mu_{Y|B}(k) - 3\sigma_{Y|B}(k)\}, \tag{14}$$

$$\bar{x}_i^{\max} = \max\{x_i^{\max}, \mu_{Y|B}(k) + 3\sigma_{Y|B}(k)\}. \tag{15}$$

In this way, values sampled from $Y|B$ that are in the interval $[\bar{x}_i^{\min}, \bar{x}_i^{\max}]$ are adjusted with values in the interval $[x_i^{\min}, x_i^{\max}]$ of the search variable. The interval $[\bar{x}_i^{\min}, \bar{x}_i^{\max}]$ is defined in (14) and (15) such that the probability that $Y|B$ lies out of this interval is small. This probability is less than the probability that $Y|B$ is out of the interval $[\mu_{Y|B}(k) - 3\sigma_{Y|B}(k), \mu_{Y|B}(k) + 3\sigma_{Y|B}(k)]$, that is equal to 0.003 [41]. The values sampled from $Y|B$ that are out of the interval $[\bar{x}_i^{\min}, \bar{x}_i^{\max}]$ are adjusted with the bounds x_i^{\min} or x_i^{\max} of the search variable.

Algorithm 3 describes formally in pseudocode the detailed steps to update the position of each particle at step k , with $k \geq 1$. The condition at line 7 verifies if the position of the particle p at the previous step $k - 1$ is the new global best position for the step k , i.e., if it is lower than the global best position obtained at the previous step $k - 1$. This condition, that does not depend on the value of i , is used at line 12 to decide the current value of the variance $\sigma_{Y|B}^2(k)$, for each i .

6.2. Parameters and analysis of the Bayesian PSO algorithm. As shown in (2), the new position of a particle is determined by means of a weighted mean among the previous position of the particle, the current personal best position and the current global best position. The weights of each component are determined by $\delta_{L|X}(k - 1)$ and $\delta_{B|Y}(k - 1)$, which are defined as a function of the variances $\sigma_{B|Y}^2$, $\sigma_{L|X}^2$ and $\sigma_{pi}^2(0)$. These variances represent the uncertainties associated to the current measurement at each step of b_i (conditioned to the prior estimation of l_i), the current measurement at each step of l_i (conditioned to the prior estimation of the position x_{pi}) and the posterior estimation at step 1 of the new position $x_{pi}(1)$, respectively. Parameters $\sigma_{B|Y}^2$ and $\sigma_{L|X}^2$ correspond to the acceleration coefficients of the basic PSO algorithm, since they control the impact of the cognitive and social components on the velocity of the particle. Parameter $\sigma_{pi}^2(0)$ adjusts the position of a particle with respect to the new value of the mean $\mu_{Y|B}(k)$. From (4) and (10) it follows that, the ratio of the weights of the cognitive and social components is

$$\frac{\sigma_{B|Y}^2}{\sigma_{L|X}^2}. \tag{16}$$

Thus, when $\sigma_{L|X}^2 < \sigma_{B|Y}^2$, the impact of the cognitive component on the velocity of the particle is greater than the impact of the social component, like in the basic PSO algorithm when $c_1 > c_2$. In this case, the particle is more attracted to its own personal best position, resulting in excessive wandering [11]. On the other hand, if $\sigma_{L|X}^2 > \sigma_{B|Y}^2$ then, like in the basic PSO algorithm when $c_1 < c_2$, the particle is strongly attracted to the global best position, causing particles to converge prematurely towards global best position [11]. Considering different values of $\sigma_{L|X}^2$ and $\sigma_{B|Y}^2$ while maintaining the same

Algorithm 3: Bayesian approach to update the position of each particle at step k , with $k \geq 1$

```

1 Data: Step  $k$ ,  $\sigma_{B|Y}^2$ ,  $\sigma_{L|X}^2$ ,  $\mu_{Y|B}(k-1)$  and  $\sigma_{Y|B}^2(k-1)$  for  $p$  and  $i$ 
2  $f(\mathbf{x}_p(k-1))$ ,  $f(\mathbf{b}(k-1))$ 
3 Bounds  $x_i^{\min}$  and  $x_i^{\max}$  of the  $i$ -th dimension of the particle's position
4 Result:  $x_{pi}(k)$ ,  $\mu_{Y|B}(k)$  and  $\sigma_{Y|B}^2(k)$  for  $p$  and  $i$ 
5 begin
6   for  $p = 1, \dots, D$  do // for each particle
7     new_global_best =  $f(\mathbf{x}_p(k-1)) < f(\mathbf{b}(k-1))$ 
8     for  $i = 1, \dots, d$  do // for each search variable
9        $x_{pi}(k-1) \leftarrow \mu_{Y|B}(k-1)$ 
10       $\sigma_{pi}^2(k-1) \leftarrow \sigma_{Y|B}^2(k-1)$ 
11       $\mu_{Y|B}(k) \leftarrow$  the right expression of Equation (2)
12      if new_global_best then
13         $\sigma_{Y|B}^2(k) \leftarrow$  the right expression of Equation (3)
14      else
15         $\sigma_{Y|B}^2(k) \leftarrow \sigma_{Y|B}^2(k-1)$ 
16         $x \leftarrow \mathcal{N}(\mu_{Y|B}(k), \sigma_{Y|B}^2(k))$  // new normal random value
17         $\bar{x}_i^{\min} \leftarrow \min\{x_i^{\min}, \mu_{Y|B}(k) - 3\sigma_{Y|B}(k)\}$ 
18         $\bar{x}_i^{\max} \leftarrow \max\{x_i^{\max}, \mu_{Y|B}(k) + 3\sigma_{Y|B}(k)\}$ 
19        if  $x < \bar{x}_i^{\min}$  then
20           $x_{pi}(k) \leftarrow x_i^{\min}$ 
21        else if  $x > \bar{x}_i^{\max}$  then
22           $x_{pi}(k) \leftarrow x_i^{\max}$ 
23        else
24           $x_{pi}(k) \leftarrow x_i^{\min} + (x - \bar{x}_i^{\min}) \frac{x_i^{\max} - x_i^{\min}}{\bar{x}_i^{\max} - \bar{x}_i^{\min}}$ 

```

ratio, different values for $\mu_V(k)$ can be obtained without changing the relative impacts of the cognitive and social components.

From (10) it follows that, for $k \geq 1$:

- the mean of the new position $\mu_{Y|B}(k)$ remains equal to the mean of the current position $\mu_{Y|B}(k-1)$, i.e., the mean of the velocity $\mu_V(k)$ tends to 0, as $\delta_{L|X}(k-1)$ and $\delta_{B|Y}(k-1)$ approach to 0, and
- $\mu_{Y|B}(k)$ tends to $l_{pi}(k-1)$ or $b_i(k-1)$ when, respectively, $\delta_{L|X}(k-1)$ or $\delta_{B|Y}(k-1)$ tends to infinity.

In fact, $\delta_{L|X}(k)$ or $\delta_{B|Y}(k)$ can assume very high values only when $k = 0$. Definitions (4) for $k = 0$ imply that $\delta_{L|X}(0)$ or $\delta_{B|Y}(0)$ varies from 0 to infinity as $\sigma_{L|X}^2$ or $\sigma_{B|Y}^2$, respectively, varies from infinity to 0, or as $\sigma_{Y|B}^2(0)$ varies from 0 to infinity.

From the closed Formulas (8) and (9) we derive that, depending on the initial value assigned to $\sigma_{pi}^2(0)$, for $h(k) \geq 1$

$$0 < \delta_{L|X}^2(k) < \frac{1}{h(k) \left(1 + \frac{\sigma_{L|X}^2}{\sigma_{B|Y}^2}\right)}, \quad 0 < \delta_{B|Y}^2(k) < \frac{1}{h(k) \left(1 + \frac{\sigma_{B|Y}^2}{\sigma_{L|X}^2}\right)} \quad (17)$$

where $\delta_{L|X}(k)$ and $\delta_{B|Y}(k)$ become 0 when $\sigma_{pi}^2(0)$ tends to 0, and they tend to their respective (less than 1) upper bounds when $\sigma_{pi}^2(0)$ approaches infinity. Thus, the maximum value that $\delta_{L|X}^2(k)$ and $\delta_{B|Y}^2(k)$ can reach for $h(k) \geq 1$ depends on the ratio between the values of $\sigma_{L|X}^2$ and $\sigma_{B|Y}^2$ and on the value of $h(k)$, independently from the value of $\sigma_{pi}^2(0)$.

Replacing the bounds of Equation (17) in Equation (10) we get

$$\begin{aligned} 0 \leq \mu_V(k) \leq & \frac{1}{h(k) \left(1 + \frac{\sigma_{L|X}^2}{\sigma_{B|Y}^2}\right)} (l_{pi}(k-1) - x_{pi}(k-1)) \\ & + \frac{1}{h(k) \left(1 + \frac{\sigma_{B|Y}^2}{\sigma_{L|X}^2}\right)} (b_i(k-1) - x_{pi}(k-1)), \end{aligned} \quad (18)$$

where, for $h(k) \geq 1$, the mean $\mu_V(k)$ of the velocity becomes 0, when $\sigma_{pi}^2(0)$ approaches 0, and it tends to the upper bound of (18), when $\sigma_{pi}^2(0)$ approaches infinity.

From (8) and (9) it follows also that, for $h(k) \geq 1$, $\delta_{L|X}^2(k)$ and $\delta_{B|Y}^2(k)$ have opposite trends, with $\delta_{L|X}^2(k)$ decreasing and $\delta_{B|Y}^2(k)$ increasing for increasing values of $\sigma_{L|X}^2$, and with $\delta_{L|X}^2(k)$ increasing and $\delta_{B|Y}^2(k)$ decreasing for increasing values of $\sigma_{B|Y}^2$. This implies that, when the two (cognitive and social) components of the velocity in (10) are positive and for a fixed value of $\sigma_{pi}^2(0)$ and $h(k)$, the mean $\mu_V(k)$ of the velocity is greater than a positive lower bound; it is not dependent on the values of $\sigma_{L|X}^2$ and $\sigma_{B|Y}^2$, but only on the value of $\sigma_{pi}^2(0)$ and $h(k)$.

Thus, when the cognitive and social components of the velocity in (10) are positive, $\mu_V(k)$ for $h(k) \geq 1$ is limited by a positive lower bound, that depends only on $h(k)$ and $\sigma_{pi}^2(0)$, and by the upper bound of (18), that depends only on $h(k)$ and on the ratio between $\sigma_{L|X}^2$ and $\sigma_{B|Y}^2$.

Depending on the initial value assigned to $\sigma_{pi}^2(0)$, from (7) we have for $h(k) \geq 1$

$$0 < \sigma_{Y|B}^2(k) < \frac{1}{h(k) \left(\frac{1}{\sigma_{L|X}^2} + \frac{1}{\sigma_{B|Y}^2}\right)} = \frac{\sigma_{L|X}^2}{h(k) \left(1 + \frac{\sigma_{L|X}^2}{\sigma_{B|Y}^2}\right)} \quad (19)$$

where $\sigma_{Y|B}^2(k)$ becomes 0, when $\sigma_{pi}^2(0)$ approaches 0, and it tends to the upper bound of (19), when $\sigma_{pi}^2(0)$ approaches infinity. Thus, the maximum value that the variance $\sigma_{Y|B}^2(k)$ and $\sigma_V^2(k)$ can reach is limited by the values of the parameters $\sigma_{L|X}^2$ and $\sigma_{B|Y}^2$, independently from the value of $\sigma_{pi}^2(0)$. From (7) we derive that, for values of $\sigma_{L|X}^2 \gg \sigma_{pi}^2(0)$ or $\sigma_{B|Y}^2 \gg \sigma_{pi}^2(0)$ and for a limited value of $h(k)$, the variance $\sigma_{Y|B}^2(k)$ can be approximated by the constant $\sigma_{pi}^2(0)$.

As an example of parameter setting, considering $\sigma_{L|X}^2 = \sigma_{B|Y}^2$ and $\sigma_{pi}^2(0) \gg \sigma_{L|X}^2$, we get for $h(k) \geq 1$

$$\delta_{L|X}^2(k) = \delta_{B|Y}^2(k) \simeq \frac{1}{2h(k)} \quad \text{and} \quad \sigma_{Y|B}^2(k) \simeq \frac{\sigma_{L|X}^2}{2h(k)},$$

where the right side expressions of the two equations are derived from the upper bounds of (17) and (19), respectively.

When the number of new global best positions $h(k)$ of the algorithm increases, the uncertainties relative to the current personal best position and the current global best position decrease, since improved positions for the personal and global best should be reached. From the previous discussion about the trends of $\mu_{Y|B}(k)$ and from Equations (7), (8) and (9) it follows that, for increasing values of $h(k)$, the mean $\mu_{Y|B}(k)$ tends to be constant and the variance $\sigma_{Y|B}^2(k)$ decreases to 0. In order to avoid that the mean $\mu_V(k)$ and the variance $\sigma_V^2(k)$ of the velocity decrease below a certain minimum value before the optimal position is reached, the following condition must be verified:

$$\sigma_{Y|B}^2(k_{\max}) \geq \epsilon, \quad (20)$$

where $\epsilon > 0$ and k_{\max} is the maximum number of steps for which the optimal position is reached. Given the values for $\sigma_{L|X}^2$, $\sigma_{B|Y}^2$, k_{\max} and ϵ , then the minimum value for $\sigma_{pi}^2(0)$ for which the variance $\sigma_{Y|B}^2(k_{\max}) \geq \epsilon$ is derived from (7) as

$$\frac{\epsilon}{1 - \epsilon h(k_{\max}) \left(\frac{1}{\sigma_{L|X}^2} + \frac{1}{\sigma_{B|Y}^2} \right)}, \quad (21)$$

with

$$\frac{1}{\sigma_{L|X}^2} + \frac{1}{\sigma_{B|Y}^2} < \frac{1}{\epsilon h(k_{\max})}.$$

For large values of $\sigma_{L|X}^2$ or $\sigma_{B|Y}^2$, Formula (21) is approximated by ϵ . On the contrary, low values of $\sigma_{L|X}^2$ and $\sigma_{B|Y}^2$ imply smaller values for ϵ or $h(k_{\max})$ and values for $\sigma_{pi}^2(0)$ higher than ϵ .

7. Evaluation of the Proposed Solutions. In this section, we carry on an evaluation study to assess performance and quality indicators of our solutions via simulation. To this purpose, we set up two variants of the basic PSO algorithm by assembling the previously described features. Specifically, the two PSO variants we concentrate on are generated respectively by i) employing the novel Algorithm 2 for the sample of the initial population and the fuzzy approach described in Section 5 (referred in the following as PSOF variant), and ii) the novel Algorithm 2 and the Algorithm 3 to change the position of particles (referred in the following as PSOB variant). Two kinds of evaluation are performed. First, the PSOF and PSOB solutions are compared with the basic PSO algorithm (indicated as PSOC) in terms of performance and quality indicators. Then, the new algorithms are employed in a simple but realistic use case in the electrical power sector taken from the literature and compared with a few other optimization solutions already adopted in the referred study. These analyses are detailed in the next two subsections.

7.1. Analysis and comparison with the basic PSO. In this study, we consider the Rosenbrock function [34] and the Griewank function [14], which are functions widely used to test the quality of optimization algorithms. In particular, we will use their multi-dimensional generalizations, i.e.,

$$f(x_1, \dots, x_d) = \sum_{i=1}^{d-1} ((1 - x_i)^2 + 100(x_{i+1} - x_i^2)^2)$$

and

$$g(x_1, \dots, x_d) = 1 + \frac{1}{4000} \sum_{i=1}^d x_i^2 - \prod_{i=1}^d \cos \frac{x_i}{\sqrt{i}}$$

respectively. Moreover, we consider -10 and 10 as the lower and upper bounds for each variable of the Rosenbrock function and $-20, 20$ as the lower and upper bounds for each variable of the Griewank function.

The two indicators we analyzed are *accuracy* A , which represents the ability of the algorithm to better approximate (or possibly reach) the optimal solution \mathbf{b}^* within a given number of steps, and *efficiency* K , which represents the promptness (in number of steps) of the algorithm in reaching the optimal solution \mathbf{b}^* or a certain approximated value \bar{a} of the objective function. A and K are random variables defined as

$$A = f(\mathbf{b}(k_{\max})) - f(\mathbf{b}^*),$$

$$K = \min\{k | f(\mathbf{b}(k)) \leq \bar{a} \text{ and } k \text{ is not limited}\}.$$

When the minimum of the objective function is 0, i.e., $f(\mathbf{b}^*) = 0$, then A is equal to the minimum value of the objective function obtained by the PSO algorithm, i.e., $f(\mathbf{b}(k_{\max}))$.

To enrich the analyses and the comparison between the three variants, for both indicators we evaluated the minimum and maximum values obtained out of the set of simulations performed, as well as the mean value, i.e.,

$$A_X^{\min} = \min\{A\}, A_X^{\max} = \max\{A\} \text{ and } A_X^{\text{mean}} = E[A],$$

$$K_X^{\min} = \min\{K\}, K_X^{\max} = \max\{K\} \text{ and } K_X^{\text{mean}} = E[K],$$

where X is the name of the variant for which the measure is derived: PSOC, PSOF or PSOB.

The maximum numbers of steps k_{\max} and the swarm dimension D are the two parameters that have been varied in the analyses, as specified in the tables summarizing the obtained results. Default values for the algorithms parameters, assumed in the simulations when not otherwise specified, are respectively: $k_{\max} = 150$ and $D = 35$.

Moreover, in the simulations we use the following standard values for the PSOC parameters:

$$w = 0.72984, \quad c_1 = c_2 = 1.496172.$$

For the PSOF algorithm, the coefficients w, c_1, c_2 are determined at each step k by using the fuzzy inference scheme described in Section 5.

Finally, in the PSOB algorithm, the parameters $\sigma_{p_i}^2(0), \sigma_{L|X}^2, \sigma_{B|Y}^2$ are set by expert-based choice as follows:

$$\sigma_p^2(0) = \sigma_{L|X}^2 = \frac{|x_i^{\max} - x_i^{\min}|}{2D}, \quad \sigma_{B|Y}^2 = \frac{|x_i^{\max} - x_i^{\min}|}{D},$$

for each component i of the p -th particle.

7.1.1. Accuracy evaluation results. Results of 100 simulation runs to evaluate the accuracy parameter are summarized in Table 2 for the PSOC, PSOF, PSOB algorithms used to minimize the Rosenbrock function with $d = 3$. The Rosenbrock function with $d = 3$ takes the minimum value, that is zero, at the point $(0, 0, 0)$.

We can observe that for the default values of the parameters (first row of the table), PSOF has a better accuracy than PSOB and PSOC with respect to the minimum value. Instead, PSOB shows a better accuracy than PSOF and PSOC for both A_{PSOB}^{mean} and A_{PSOB}^{\max} . Similar results are obtained for different values of the parameters D and k_{\max} . Although there is no definitive rank among the three variants of the PSO under analysis, the accuracy of the basic PSO version is never better than both the other two. Especially

TABLE 2. Minimization of the Rosenbrock function with $d = 3$ using PSOC, PSOF, PSOB algorithms, for 100 simulations

	A_{PSOC}^{\min}	A_{PSOC}^{mean}	A_{PSOC}^{\max}	A_{PSOF}^{\min}	A_{PSOF}^{mean}	A_{PSOF}^{\max}	A_{PSOB}^{\min}	A_{PSOB}^{mean}	A_{PSOB}^{\max}
Default	$4.8 \cdot 10^{-5}$	0.5730	8.7900	$4.2 \cdot 10^{-7}$	0.04701	7.3425	0.0008	0.0339	0.1523
D = 20	$8.8 \cdot 10^{-5}$	0.9532	8.0209	$8.0 \cdot 10^{-5}$	0.3920	9.2987	0.0033	0.1473	0.6662
D = 50	$1.1 \cdot 10^{-5}$	0.1951	4.9227	$2.1 \cdot 10^{-7}$	0.0134	0.0645	$2.1 \cdot 10^{-5}$	0.0127	0.1057
$k_{\max} = 100$	$6.2632 \cdot 10^{-5}$	0.6534	5.8641	0.0002	0.0727	0.2868	0.0005	0.0552	0.2147
$k_{\max} = 200$	$6.1 \cdot 10^{-6}$	0.3933	5.2793	$4.3 \cdot 10^{-9}$	0.0284	0.1110	$6.5 \cdot 10^{-5}$	0.0258	0.1179

TABLE 3. Minimization of the Griewank function with $d = 5$ using PSOC, PSOF, PSOB algorithms, for 100 simulations

	A_{PSOC}^{\min}	A_{PSOC}^{mean}	A_{PSOC}^{\max}	A_{PSOF}^{\min}	A_{PSOF}^{mean}	A_{PSOF}^{\max}	A_{PSOB}^{\min}	A_{PSOB}^{mean}	A_{PSOB}^{\max}
Default	0.0004	0.0757	0.2234	$2.2 \cdot 10^{-6}$	0.0118	0.0236	0.0038	0.0071	0.0481
D = 20	0.0072	0.0901	0.2280	0.0001	0.0839	0.2228	0.0005	0.0369	0.1073
D = 50	$9.1 \cdot 10^{-5}$	0.0626	0.2046	$6.4 \cdot 10^{-10}$	$8.2 \cdot 10^{-5}$	0.0005	$6.9 \cdot 10^{-5}$	0.0008	0.0035
$k_{\max} = 100$	0.0114	0.0941	0.2612	0.0007	0.0832	0.2490	0.0006	0.0075	0.0481
$k_{\max} = 200$	0.0003	0.0699	0.1626	$4.7 \cdot 10^{-8}$	0.0007	0.0153	$3.0 \cdot 10^{-5}$	0.0004	0.0495

when focusing on A_{PSOC}^{mean} , it is constantly (and significantly) higher than A_{PSOF}^{mean} and A_{PSOB}^{mean} , meaning lower accuracy.

Similarly, results of 100 simulation runs are summarized in Table 3 for the PSOC, PSOF, PSOB algorithms used to minimize the Griewank function with $d = 5$, again to evaluate the accuracy indicator. The Griewank function with $d = 5$ takes the minimum value, that is zero, in several different points.

Comments similar to those already made with reference to the previous table are applicable also to Table 3. PSOC accuracy is always outperformed by either PSOB or PSOF (actually, by both, unless for the minimum value in the first row when the default parameters are used, for which PSOC is better than PSOB). Therefore, also in this use case the proposed variants PSOF and PSOB should be preferred.

7.1.2. Efficiency evaluation results. To assess the efficiency indicator, we considered the Griewank function only and calculated the (minimum, mean and maximum) number of steps required by each of the three PSO variants to reach a predefined value \bar{a} for this function. Again, 100 simulations have been performed; each run is completed either when the specified value is obtained by the execution of the PSO algorithm, or a maximum of 150 steps is exceeded. Table 4 summarizes the simulation results. The values chosen for \bar{a} are the minimum, mean and maximum values obtained in Table 3 by PSOC, in correspondence of the row “Default”, that is, $A_{PSOC}^{\min} = 0.0004$, $A_{PSOC}^{\text{mean}} = 0.0757$ and $A_{PSOC}^{\max} = 0.2234$, respectively. The values 0 in the table indicate that the initial positions randomly selected by the corresponding algorithms already provide an outcome better than \bar{a} . Instead, when “not found” appears, it means that \bar{a} is not reached within 150 steps (maximum number of steps set in the simulations). Looking at the results in Table 4, we can immediately appreciate the significantly better efficiency shown in general by the variants PSOF and PSOB. In fact, for both $\bar{a} = A_{PSOC}^{\text{mean}}$ and $\bar{a} = A_{PSOC}^{\max}$, the number of steps requested by PSOF and PSOB is much lower than that required by PSOC, being indeed 0 for $\bar{a} = A_{PSOC}^{\max}$. The only exception is PSOB for $\bar{a} = A_{PSOC}^{\min}$, which is not surprising, since looking at Table 3, PSOB shows lower accuracy than the minimum value obtained by PSOC, when default parameters are used (which is 0.0004).

TABLE 4. Minimization of the Griewank function with $d = 5$ using PSOC, PSOF, PSOB algorithms, for 100 simulations

	K_{PSOC}^{\min}	$K_{\text{PSOC}}^{\text{mean}}$	K_{PSOC}^{\max}	K_{PSOF}^{\min}	$K_{\text{PSOF}}^{\text{mean}}$	K_{PSOF}^{\max}	K_{PSOB}^{\min}	$K_{\text{PSOB}}^{\text{mean}}$	K_{PSOB}^{\max}
$\bar{a} = A_{\text{PSOC}}^{\min}$	127	131	not found	16	28	50	not found	not found	not found
$\bar{a} = A_{\text{PSOC}}^{\text{mean}}$	17	66	144	4	14	21	8	12	31
$\bar{a} = A_{\text{PSOC}}^{\max}$	1	26	122	0	0	0	0	0	0

7.2. Evaluation in a practical case study: Optimization of electrical grids. In this section, we have evaluated our modified PSO algorithms employed in electrical grid optimizations, in terms of minimizing the total power loss over the lines. The evaluation is performed in terms of the accuracy indicator, as introduced in 7.1. In particular, we focused on the IEEE-6 bus system in Figure 2, typically used as a simple case study to deal with optimization in electrical grids. A survey on the general optimal power flow problem (i.e., the problem to optimize electrical grids) can be found in [13] and [39]. In [38], the authors have reported the results of the optimization of the IEEE-6 bus system by using different optimization algorithms: Genetic Algorithm (GA), Particle Swarm Optimization (PSO), Artificial Bee Colony (ABC), Differential Evolution (DE). The optimization has been implemented through minimizing the total power loss over the lines of the electrical grid, expressed as follows:

$$P_{\text{loss}} = \frac{1}{2} \sum_{i=1}^N \sum_{j=1}^N g_{i,j} (V_i^2 + V_j^2 - 2V_i V_j \cos(\delta_i - \delta_j))$$

where N is the total number of bus (nodes), V_i is the voltage magnitude at bus i , δ_i is the voltage angle at bus i , and $g_{i,j}$ is the conductance of the line connecting buses i and j . For optimization purposes, seven control variables have been used in these papers: P_6 active power of bus 6, V_1 and V_6 voltages of buses 1 and 6, Q_3 and Q_4 reactive power of buses 3 and 4, T_{24} and T_{35} tap changers of transformers over lines connecting busses 2, 4 and 3, 5. We have instantiated our PSOF and PSOB solutions to this problem, adopting the same control variables, which correspond to the search variables of our algorithms. For the sake of comparison with the previously analyzed optimization algorithms, the same ranges of values as in [38] have been assigned to the (both control and state) variables of our algorithms; for brevity, they are omitted here.

In Table 5, we report the results obtained in [38] for the GA, PSO, ABC and DE algorithms, with the addition of the results that we have obtained using our modified algorithms PSOB and PSOF. These results (expressed in terms of min, mean and maximum

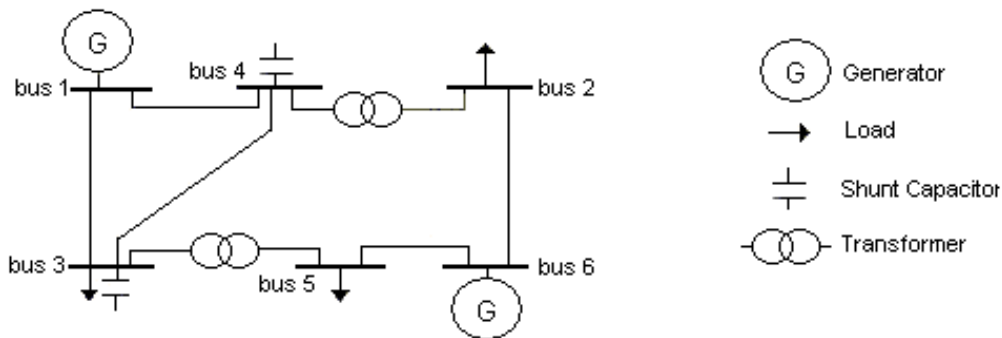


FIGURE 2. IEEE-6 bus system

TABLE 5. OPF solutions obtained by using GA, PSO ABC, DE, PSOB, and PSOF for the 6-bus test system

Method	$P_{\text{loss}}^{\text{Min}}$	$P_{\text{loss}}^{\text{Mean}}$	$P_{\text{loss}}^{\text{Max}}$
GA	6.7747	6.9705	7.5292
PSO	6.7486	6.8425	7.1517
ABC	6.7361	6.7361	6.7364
DE	6.7361	6.7361	6.7368
PSOF	6.7329	6.7557	6.9042
PSOB	6.7329	6.7331	6.7333

TABLE 6. Values of control variables corresponding to minimum solution obtained by PSOB and PSOF algorithms

Control variable	Value
V_1	1.1
V_6	1.1
P_6	27.6
Q_3	43
Q_4	27
T_{24}	1.0475
T_{35}	0.9975

total power loss) are obtained with $k_{\max} = 200$, $D = 30$, running 30 simulation experiments. From this table, it can be observed that both PSOB and PSOF show the best minimum value for the total power loss. Moreover, PSOB is the best also in terms of the average and maximum values, so ranking as the most accurate power flow optimization algorithm among the 6 considered alternatives for this scenario. Although dependent on the specific setting adopted in the setup of this case study, the observed results are encouraging to investigate more deeply the suitability of our proposed solutions to support power flow optimizations, which we plan as future work.

For completeness on the presented case study, in Table 6 we report the optimal configuration for the control variables that yields the minimum value of the total power loss obtained by the PSOB and the PSOF algorithms.

8. Conclusions. In this paper, the problem of performing non-linear optimization has been treated by means of the PSO algorithm. Moving from the original formulation of the PSO algorithm, improvements have been proposed on the sampling of the initial population, the setting of the algorithm's parameters and the method for moving the control variables in the search space. Well consolidated techniques have been adopted to cover the critical points of the algorithm we worked on, namely the fuzzy logics and the Bayesian theory. A simulation study has been carried on to show the benefits of the new proposed PSO solutions with respect to the classical PSO formulation, using well-known testing functions like the Rosenbrock function and the Griewank function. Moreover, a case study in the electrical field has been worked out, to show results on the suitability of our proposed solutions to support optimization needs in this context, in comparisons with other already adopted alternatives. The obtained results are encouraging, and we expect that the new features we introduced into the PSO algorithm are actually relevant in a variety of application contexts, especially those that are sensitive to the input sampling and the coverage of the search space.

Several research extensions are foreseen. Further simulations devoted to improve the understanding of the sensitivity of the three compared solutions to the algorithms parameters would be undoubtedly a valuable direction to explore. Other refinements would be also interesting, such as focus on the setting of variances and introducing correlation factors in the Bayesian approach, as well as setting of expert-based choices in the fuzzy logics strategy. Of course, addressing more deep investigations on the usage of the proposed solutions in specific application contexts, such as the electrical power system already tackled in this paper, is another planned research line. Finally, practical support to the the selection of the most suited algorithm to solve optimization aspects would be also very helpful by exploring the characterization of the PSO family with respect to typical needs raised in optimizations problems.

Acknowledgements. This work has been partially supported by the European Project SmartC2Net (n. ICT-318023) and the TENACE PRIN Project (n. 20103P34XC) funded by the Italian Ministry of Education, University and Research.

REFERENCES

- [1] A. M. Abdelbar, S. Abdelshahid and D. C. Wunsch, Fuzzy PSO: A generalization of particle swarm optimization, *Proc. of Int. Joint Conference on Neural Networks*, vol.2, pp.1086-1091, 2005.
- [2] G. Andersson, *Modelling and Analysis of Electric Power Systems*, Lecture 227-0526-00, ETH Zurich, 2004.
- [3] A. G. Bakirtzis, P. N. Biskas, C. E. Zoumas and V. Petridis, Optimal power flow by enhanced genetic algorithm, *IEEE Trans. Power Systems*, vol.17, pp.229-236, 2002.
- [4] J. C. Bansal, M. Singh, P. K. Saraswat, A. Verma, S. S. Jadon and A. Abraham, Inertia weight strategies in particle swarm optimization, *The 3rd World Congress on Nature and Biologically Inspired Computing*, pp.640-647, 2011.
- [5] D. Bratton and J. Kennedy, Defining a standard for particle swarm optimization, *IEEE Swarm Intelligence Symposium*, pp.120-127, 2007.
- [6] G. B. Dantzig, *Linear Programming and Extensions*, Princeton University Press, 1963.
- [7] G. E. D'Errico and N. Murru, Fuzzy treatment of candidate outliers in measurements, *Advances in Fuzzy Systems*, Article ID 783843, 2012.
- [8] G. E. D'Errico and N. Murru, An algorithm for concurrent estimation of time-varying quantities, *Meas. Sci. Technol.*, vol.3, no.4, 2012.
- [9] J. Dongarra and F. Sullivan, Guest editors' introduction: The top 10 algorithms, *Computing in Science and Engineering*, vol.2, no.1, pp.22-23, 2000.
- [10] D. Dubois and H. M. Prade, *Fuzzy Sets and Systems*, Academic Press, New York, 1980.
- [11] A. P. Engelbrecht, *Computational Intelligence: An Introduction*, 2nd Edition, John Wiley and Sons, Ltd., Chichester, England, 2007.
- [12] A. Esmín and G. Lambert-Torres, Application of particle swarm optimization to optimal power systems, *International Journal of Innovative Computing, Information and Control*, vol.8, no.3(A), pp.1705-1716, 2012.
- [13] J. J. Grainger and W. D. Stevenson, *Power System Analysis*, McGraw-Hill, 1994.
- [14] A. O. Griewank, Generalized descent for global optimization, *J. Opt. Th. Appl.*, vol.34, pp.11-39, 1931.
- [15] G. Harrison and A. R. Wallace, Optimal power flow evaluation of distribution network capacity for the connection of distributed generation, *IEEE Proc. of Gener. Trans. Distrib.*, vol.152, no.1, pp.115-122, 2005.
- [16] S. Helwig and R. Wanka, Theoretical analysis of initial particle swarm behavior, *The 10th Int. Conf. on Parallel Problem Solving from Nature*, Dortmund, Germany, pp.889-898, 2008.
- [17] X. Jiang, H. Ling, J. Yan, B. Li and Z. Li, Forecasting electrical energy consumption of equipment maintenance using neural network and particle swarm optimization, *Mathematical Problems in Engineering*, 2013.
- [18] R. E. Kalman, A new approach to linear filtering and prediction problems, *J. Basic Eng. Trans. ASME*, vol.82, no.D, pp.35-45, 1960.

- [19] Q. Kang, M. Zhou, J. An and Q. Wu, Swarm intelligence approaches to optimal power flow problem with distributed generator, *IEEE Trans. Automation Science and Engineering*, vol.10, no.2, pp.343-353, 2013.
- [20] J. Kennedy and R. Eberhart, Particle swarm optimization, *IEEE Int. Conference on Neural Networks*, Perth, Australia, vol.IV, pp.1942-1948, 2012.
- [21] T. Kulworawanichpong, Optimal power flow solution using adaptive tabu search, *Int. J. of the Physical Sciences*, vol.6, no.28, pp.6394-6409, 2011.
- [22] T. R. Kuphaldt, *Lessons in Electric Circuits, Volume II*, 6th Edition, 2007.
- [23] T. Lantharthong and N. Rugthiacharoencheep, Network configuration for load balancing in distribution system with distributed generation and capacitor placement, *Engineering and Technology*, vol.64, pp.435-440, 2012.
- [24] J. Lavaei and S. H. Low, Zero duality gap in optimal power flow problem, *IEEE Trans. Power Systems*, vol.27, no.1, pp.92-107, 2012.
- [25] U. Leeton, U. Kwannetr, D. Uthitsunthorn and T. Kulworawanichpong, Power loss minimization using optimal power flow based on particle swarm algorithm, *International Conference on Electrical Engineering/Electronics Computer Telecommunications and Information Technology 2010*, pp.440-444, 2010.
- [26] E. H. Mamdani and S. Assilian, An experiment in linguistic synthesis with a fuzzy logic controller, *International Journal of Man-Machine Studies*, vol.7, no.1, pp.1-13, 1975.
- [27] A. Mu, D. Cao and X. Wang, A modified particle swarm optimization algorithm, *Natural Science*, vol.2, no.1, pp.151-155, 2009.
- [28] A. Nickabadi, M. Ebadzadeh and R. Safabakhsh, A novel particle swarm optimisation algorithm with adaptive inertia weight, *Applied Soft Computing*, vol.11, pp.3658-3670, 2011.
- [29] T. Niknam, H. D. Mojarrad and M. Nayeripour, A new hybrid fuzzy adaptive particle swarm optimization for non-convex economic dispatch, *International Journal of Innovative Computing, Information and Control*, vol.7, no.1, pp.189-202, 2011.
- [30] T. Da Silva Noronha, T. F. De Oliveira, A. M. Da Silveira, R. R. Da Silva and A. C. F. Saraiva, Knowledge acquisition of vibrations in high power transformers using statistical analyses and fuzzy approaches: A case study, *Electric Power Systems Research*, vol.104, pp.110-115, 2013.
- [31] R. Ojha and M. Das, An adaptive approach for modifying inertia weight using particle swarm optimisation, *International Journal of Computer Science*, vol.9, no.2, pp.105-111, 2012.
- [32] B. K. Panigrahi, V. R. Pandi and S. Das, Adaptive particle swarm optimization approach for static and dynamic economic load dispatch, *Energy Conversion and Management*, vol.49, pp.1407-1415, 2008.
- [33] R. Poli, Analysis of the publications on the applications of particle swarm optimisation, *Journal of Artificial Evolution and Applications*, vol.2008, no.4, 2008.
- [34] H. H. Rosenbrock, An automatic method for finding the greatest or least value of a function, *The Computer Journal*, vol.3, pp.175-184, 1960.
- [35] S. S. Sharif, J. H. Taylor and E. F. Hill, On-line optimal power flow by energy loss minimization, *The 35th Conf. on Decision and Control*, Kobe, Japan, 1996.
- [36] S. Liu, Z. Hou and M. Wang, A hybrid algorithm for optimal power flow using the chaos optimization and the linear interior point algorithm, *International Conference on Power System Technology, Proc. of PowerCon*, vol.2, pp.793-797, 2002.
- [37] Y. Shi and R. C. Eberhart, Fuzzy adaptive particle swarm optimization, *IEEE Congress on Evolutionary Computation*, pp.101-106, 2001.
- [38] N. Sinsuphun, U. Leeton, U. Kwannetr, D. Uthitsunthorn and T. Kulworawanichpong, Loss minimization using optimal power flow based on swarm intelligences, *Trans. Electrical Eng., Electronics, and Communications*, vol.9, no.1, pp.212-222, 2011.
- [39] B. Stott, Review of load-flow calculation methods, *Proc. of the IEEE*, vol.62, no.7, pp.916-929, 1974.
- [40] P. L. Swe, W. Swe and K. M. Lin, Effects of tap changing transformer and shunt capacitor on voltage stability enhancement of transmission networks, *World Academy of Science, Engineering and Technology*, vol.5, no.3, pp.433-437, 2011.
- [41] K. S. Trivedi, *Probability and Statistics with Reliability, Queueing and Computer Science Applications*, John Wiley and Sons, New York, 2002.
- [42] J. Yuryevich and K. P. Wong, Evolutionary programming based optimal power flow algorithm, *IEEE Trans. Power Systems*, vol.14, pp.1245-1250, 1999.

- [43] J. B. Ward and H. W. Hale, Digital computer solutions of power flow problems, *Trans. of the American Institute of Electrical Engineers. Part III: Power Apparatus and Systems*, vol.75, no.3, pp.398-404, 1956.
- [44] Q. H. Wu, Power system optimal reactive power dispatch using evolutionary programming, *IEEE Trans. Power Systems*, vol.10, no.3, pp.1243-1249, 1995.
- [45] M. Wu, X. Zhang and D. Zhang, A novel chaotic PSO algorithm based on Tent map and its application to mechanical design, *Journal of Information and Computational Science*, vol.10, no.6, pp.1789-1795, 2013.
- [46] H. Zhang, L. Zhang and F. Meng, Reactive power optimization based on genetic algorithm, *International Conference on Power System Technology*, vol.2, pp.1448-1453, 1998.
- [47] B. Zhao, C. X. Guo and Y. J. Cao, An improved particle swarm optimization algorithm for optimal reactive power dispatch, *IEEE Power Engineering Society General Meeting*, vol.1, pp.272-279, 2005.
- [48] Y. Zheng, H. Ling and Q. Guan, Adaptive parameters for a modified comprehensive learning particle swarm optimiser, *Mathematical Problems in Engineering*, 2012.

Appendix A. Appendix. We use Equation (2) for deriving Formula (10) for $\mu_V(k)$ as follows

$$\begin{aligned}
 \mu_V(k) &= \mu_{Y|B}(k) - \mu_{Y|B}(k-1) \\
 &= \frac{x_{pi}(k-1) + \delta_{L|X}(k-1)l_{pi} + \delta_{B|Y}(k-1)b_i}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} - x_{pi}(k-1) \\
 &= \frac{x_{pi}(k-1) + \delta_{L|X}(k-1)l_{pi} + \delta_{B|Y}(k-1)b_i}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 &\quad - \frac{x_{pi}(k-1) + \delta_{L|X}(k-1)x_{pi}(k-1) + \delta_{B|Y}(k-1)x_{pi}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 &= \frac{\delta_{L|X}(k-1)(l_{pi}(k-1) - x_{pi}(k-1)) + \delta_{B|Y}(k-1)(b_i(k-1) - x_{pi}(k-1))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}.
 \end{aligned}$$

Formula (11) for $\sigma_V^2(k)$ is derived from (3) as follows

$$\begin{aligned}
 \sigma_V^2(k) &= \sigma_{Y|B}^2(k) + \sigma_{Y|B}^2(k-1) \\
 &= \frac{\sigma_{pi}^2(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} + \sigma_{pi}^2(k-1) \\
 &= \frac{(2 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1))\sigma_{pi}^2(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}.
 \end{aligned}$$

Formula (12) for $\mu_V(k)$ is derived from Equations (10) and (2) as follows. From (3) we get

$$\frac{1}{\sigma_{Y|B}^2(k)} = \frac{1}{\sigma_{Y|B}^2(k-1)} + \frac{1}{\sigma_{L|X}^2} + \frac{1}{\sigma_{B|Y}^2}, \tag{22}$$

$$\frac{1}{\sigma_{Y|B}^2(k-1)} = \frac{1}{\sigma_{Y|B}^2(k)} - \frac{1}{\sigma_{L|X}^2} - \frac{1}{\sigma_{B|Y}^2}, \tag{23}$$

$$\frac{\sigma_{Y|B}^2(k)}{\sigma_{Y|B}^2(k-1)} = 1 - \delta_{L|X}(k) + \delta_{B|Y}(k), \tag{24}$$

$$\frac{\sigma_{Y|B}^2(k-1)}{\sigma_{Y|B}^2(k)} = 1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1). \tag{25}$$

From (4) for each value of k and h we get

$$\sigma_{Y|B}^2(k)\delta_{L|X}(h) = \frac{\sigma_{Y|B}^2(k)\sigma_{Y|B}^2(h)}{\sigma_{L|X}^2} = \delta_{L|X}(k)\sigma_{Y|B}^2(h) \tag{26}$$

$$\sigma_{Y|B}^2(k)\delta_{B|Y}(h) = \frac{\sigma_{Y|B}^2(k)\sigma_{Y|B}^2(h)}{\sigma_{B|Y}^2} = \delta_{B|Y}(k)\sigma_{Y|B}^2(h). \tag{27}$$

From (10) we obtain a new formula for $\mu_V(k)$ as

$$\begin{aligned} \mu_V(k-1) &= \frac{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}\mu_V(k-1) \\ &= \frac{1}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}\mu_V(k-1) \\ &\quad - \frac{\delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}\mu_V(k-1) \\ &= \frac{\sigma_{Y|B}^2(k-1)\sigma_{Y|B}^2(k-2)}{\sigma_{Y|B}^2(k-2)\sigma_{Y|B}^2(k-1)} \frac{1}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}\mu_V(k-1) \\ &\quad - \frac{\delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}\mu_V(k-1) \\ &= \frac{\sigma_{Y|B}^2(k-1)}{\sigma_{Y|B}^2(k-2)} \frac{1 + \delta_{L|X}(k-2) + \delta_{B|Y}(k-2)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}\mu_V(k-1) \\ &\quad - \frac{\delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}\mu_V(k-1), \text{ from (25)} \\ &= \frac{\sigma_{Y|B}^2(k-1)}{\sigma_{Y|B}^2(k-2)} \frac{1 + \delta_{L|X}(k-2) + \delta_{B|Y}(k-2)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \\ &\quad \frac{\delta_{L|X}(k-2)(l_{pi}(k-2) - x_{pi}(k-2)) + \delta_{B|Y}(k-2)(b_i(k-2) - x_{pi}(k-2))}{1 + \delta_{L|X}(k-2) + \delta_{B|Y}(k-2)} \\ &\quad - \frac{\delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}(x_{pi}(k-1) - x_{pi}(k-2)), \text{ from (10)} \\ &= \frac{\sigma_{Y|B}^2(k-1)}{\sigma_{Y|B}^2(k-2)} \\ &\quad \frac{\delta_{L|X}(k-2)(l_{pi}(k-2) - x_{pi}(k-2)) + \delta_{B|Y}(k-2)(b_i(k-2) - x_{pi}(k-2))}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \\ &\quad + \frac{\delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}(x_{pi}(k-2) - x_{pi}(k-1)) \\ &= \frac{\sigma_{Y|B}^2(k-2)}{\sigma_{Y|B}^2(k-2)} \\ &\quad \frac{\delta_{L|X}(k-1)(l_{pi}(k-2) - x_{pi}(k-2)) + \delta_{B|Y}(k-2)(b_i(k-2) - x_{pi}(k-2))}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \\ &\quad + \frac{\delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}(x_{pi}(k-2) - x_{pi}(k-1)), \text{ from (26) and (27)} \end{aligned}$$

$$\begin{aligned}
 &= \frac{\delta_{L|X}(k-1)l_{pi}(k-2) - \delta_{L|X}(k-1)x_{pi}(k-2)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \\
 &+ \frac{\delta_{B|Y}(k-2)b_i(k-2) - \delta_{B|Y}(k-2)x_{pi}(k-2)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \\
 &+ \frac{\delta_{L|X}(k-1)x_{pi}(k-2) + \delta_{B|Y}(k-1)x_{pi}(k-2)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \\
 &+ \frac{-\delta_{L|X}(k-1)x_{pi}(k-1) - \delta_{B|Y}(k-1)x_{pi}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \\
 &= \frac{\delta_{L|X}(k-1)(l_{pi}(k-2) - x_{pi}(k-1)) + \delta_{B|Y}(k-1)(b_i(k-2) - x_{pi}(k-1))}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}
 \end{aligned} \tag{28}$$

Finally from (10) and (28) we derive Formula (12) for $\mu_V(k)$ as

$$\begin{aligned}
 \mu_V(k) &= \frac{\delta_{L|X}(k-1)(l_{pi}(k-1) - x_{pi}(k-1)) + \delta_{B|Y}(k-1)(b_i(k-1) - x_{pi}(k-1))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 &= \frac{\delta_{L|X}(k-1)l_{pi}(k-1) - \delta_{L|X}(k-1)x_{pi}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 &+ \frac{\delta_{B|Y}(k-1)b_i(k-1) - \delta_{B|Y}(k-1)x_{pi}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 &+ \frac{\delta_{L|X}(k-1)l_{pi}(k-2) - \delta_{L|X}(k-1)l_{pi}(k-2)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 &+ \frac{\delta_{B|Y}(k-1)b_i(k-2) - \delta_{B|Y}(k-1)b_i(k-2)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 &= \frac{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \\
 &\left(\frac{\delta_{L|X}(k-1)l_{pi}(k-2) - \delta_{L|X}(k-1)x_{pi}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \right. \\
 &\left. + \frac{\delta_{B|Y}(k-1)b_i(k-2) - \delta_{B|Y}(k-1)x_{pi}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \right) \\
 &+ \frac{\delta_{L|X}(k-1)l_{pi}(k-1) - \delta_{L|X}(k-1)l_{pi}(k-2)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 &+ \frac{\delta_{B|Y}(k-1)b_i(k-1) - \delta_{B|Y}(k-1)b_i(k-2)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 &= \frac{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 &\left(\frac{\delta_{L|X}(k-1)l_{pi}(k-2) - \delta_{L|X}(k-1)x_{pi}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \right. \\
 &\left. + \frac{\delta_{B|Y}(k-1)b_i(k-2) - \delta_{B|Y}(k-1)x_{pi}(k-1)}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \right) \\
 &+ \frac{\delta_{L|X}(k-1)l_{pi}(k-1) - \delta_{L|X}(k-1)l_{pi}(k-2)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}
 \end{aligned}$$

$$\begin{aligned}
& + \frac{\delta_{B|Y}(k-1)b_i(k-1) - \delta_{B|Y}(k-1)b_i(k-2)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
& = \frac{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
& \quad \frac{\delta_{L|X}(k-1)(l_{pi}(k-2) - x_{pi}(k-1)) + \delta_{B|Y}(k-1)(b_i(k-2) - x_{pi}(k-1))}{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)} \\
& \quad + \frac{\delta_{L|X}(k-1)(l_{pi}(k-1) - l_{pi}(k-2)) + \delta_{B|Y}(k-1)(b_i(k-1) - b_i(k-2))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
& = \frac{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \mu_V(k-1) \\
& \quad + \frac{\delta_{L|X}(k-1)(l_{pi}(k-1) - l_{pi}(k-2))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
& \quad + \frac{\delta_{B|Y}(k-1)(b_i(k-1) - b_i(k-2))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}, \text{ from (28)} \tag{29}
\end{aligned}$$

Another interesting formula for $\mu_V(k)$, is derived from (28) observing that

$$\begin{aligned}
l_{pi}(k-1) - l_{pi}(k-2) &= (x_{pi}(k-1) - x_{pi}(k-2)) \\
& \quad + (l_{pi}(k-1) - x_{pi}(k-1)) \\
& \quad - (l_{pi}(k-2) - x_{pi}(k-2)), \tag{30}
\end{aligned}$$

$$\begin{aligned}
b_i(k-1) - b_i(k-2) &= (x_{pi}(k-1) - x_{pi}(k-2)) \\
& \quad + (b_i(k-1) - x_{pi}(k-1)) \\
& \quad - (b_i(k-2) - x_{pi}(k-2)). \tag{31}
\end{aligned}$$

Replacing (30) and (31) in Equation (28) we get

$$\begin{aligned}
\mu_V(k-1) &= \frac{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \mu_V(k-1) \\
& \quad + \frac{\delta_{L|X}(k-1)(l_{pi}(k-1) - l_{pi}(k-2))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
& \quad + \frac{\delta_{B|Y}(k-1)(b_i(k-1) - b_i(k-2))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
& = \frac{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \mu_V(k-1) \\
& \quad + \frac{\delta_{L|X}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} (x_{pi}(k-1) - x_{pi}(k-2)) \\
& \quad + \frac{\delta_{L|X}(k-1)((l_{pi}(k-1) - x_{pi}(k-1)) - (l_{pi}(k-2) - x_{pi}(k-2)))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
& \quad + \frac{\delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} (x_{pi}(k-1) - x_{pi}(k-2)) \\
& \quad + \frac{\delta_{B|Y}(k-1)((b_i(k-1) - x_{pi}(k-1)) - (b_i(k-2) - x_{pi}(k-2)))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
& = \frac{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \mu_V(k-1)
\end{aligned}$$

$$\begin{aligned}
 & + \frac{\delta_{L|X}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \mu_V(k-1) \\
 & + \frac{\delta_{L|X}(k-1)((l_{pi}(k-1) - x_{pi}(k-1)) - (l_{pi}(k-2) - x_{pi}(k-2)))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 & + \frac{\delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \mu_V(k-1) \\
 & + \frac{\delta_{B|Y}(k-1)((b_i(k-1) - x_{pi}(k-1)) - (b_i(k-2) - x_{pi}(k-2)))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 = & \frac{1 - \delta_{L|X}(k-1) - \delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \mu_V(k-1) \\
 & + \frac{\delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \mu_V(k-1) \\
 & + \frac{\delta_{L|X}(k-1)((l_{pi}(k-1) - x_{pi}(k-1)) - (l_{pi}(k-2) - x_{pi}(k-2)))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 & + \frac{\delta_{B|Y}(k-1)((b_i(k-1) - x_{pi}(k-1)) - (b_i(k-2) - x_{pi}(k-2)))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 = & \frac{1}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \mu_V(k-1) \\
 & + \frac{\delta_{L|X}(k-1)((l_{pi}(k-1) - x_{pi}(k-1)) - (l_{pi}(k-2) - x_{pi}(k-2)))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)} \\
 & + \frac{\delta_{B|Y}(k-1)((b_i(k-1) - x_{pi}(k-1)) - (b_i(k-2) - x_{pi}(k-2)))}{1 + \delta_{L|X}(k-1) + \delta_{B|Y}(k-1)}. \tag{32}
 \end{aligned}$$

Equation (32) shows also that, when the mean $\mu_V(k)$ of the velocity of a particle is greater than 0, the particle changes direction with respect to the current velocity (moment) only when the distance between the current position and the personal or global best position is changed with respect to the previous step. Otherwise, the magnitude of new velocity is obtained as a fraction of that of the current velocity.