

# Multi-dimensional Particle Swarm Optimization in Dynamic Environments

Serkan Kiranyaz, Jenni Pulkkinen and Moncef Gabbouj

**Abstract**— Particle swarm optimization (PSO) was proposed as an optimization technique for static environments; however, many real problems are dynamic, meaning that the environment and the characteristics of the global optimum can change in time. In this paper, we adapt recent techniques, which successfully address several major problems of PSO and exhibit a significant performance over multi-modal and non-stationary environments. In order to address the premature convergence problem and improve the rate of PSO’s convergence to global optimum, Fractional Global Best Formation (FGBF) technique is used. FGBF basically collects all the best dimensional components and fractionally creates an artificial Global Best particle (*aGB*) that has the potential to be a better “guide” than the PSO’s native *gbest* particle. To establish follow-up of local optima, we then introduce a novel multi-swarm algorithm, which enables each swarm to converge to a different optimum and use FGBF technique distinctively. Finally for the multi-dimensional dynamic environments where the optimum dimension too changes in time, we utilize a recent PSO technique, the Multi-Dimensional (MD) PSO, which re-forms the native structure of swarm particles in such a way that they can make inter-dimensional passes with a dedicated dimensional PSO process. Therefore, in a multi-dimensional search space where the optimum dimension is unknown, swarm particles can seek for both positional and dimensional optima. This eventually pushes the frontier of the optimization problems in dynamic environments towards a global search in a multi-dimensional space, where there exists a multi-modal problem possibly in each dimension. We investigated both standalone and mutual applications of the proposed methods over the moving peaks benchmark (MPB), which originally simulates a dynamic environment in a unique (fixed) dimension. MPB is appropriately extended to accomplish the simulation of a multi-dimensional dynamic system, which contains dynamic environments active in several dimensions. An extensive set of experiments show that in traditional MPB application domain, FGBF technique applied with multi-swarms exhibits an impressive speed gain and tracks the global peak with the minimum error so far achieved with respect to the other competitive PSO-based methods. When applied over the extended MPB, MD PSO with FGBF can find optimum dimension and provide the (near-) optimal solution in this dimension.

**Index Terms**—Particle Swarm Optimization, multi-dimensional search, Fractional Global Best Formation

## I. INTRODUCTION

**M**ANY real-world problems are dynamic and thus require systematic re-optimizations due to system and/or environmental changes. Even though it is possible to handle such dynamic problems as a series of individual processes via restarting the optimization algorithm after each change, this may lead to a significant loss of useful information, especially when the change is not too drastic. Since most of such problems have multi-modal nature, which further complicates the dynamic optimization problems, the need for powerful and efficient optimization techniques is imminent. In the last decade the efforts have been focused on evolutionary algorithms (EAs) [4] such as Genetic Algorithms (GA) [19], Genetic Programming (GP) [26], Evolution Strategies (ES), [5] and Evolutionary

Programming (EP) [18]. The common point of all EAs, which have population based nature may also avoid being trapped in a local optimum. Thus they can find the optimum solutions; however, this is never guaranteed.

Conceptually speaking, Particle Swarm Optimization (PSO) [16], [24], [35], which has obvious ties with the EA family, lies somewhere in between GA and EP. Yet unlike GA, PSO has no complicated evolutionary operators such as *crossover*, *selection* and *mutation* and it is highly dependent on stochastic processes. PSO is originated from the computer simulation of individuals (particles or living organisms) in a bird flock or fish school [48], which basically show a natural behavior when they search for some target (e.g. food). Their goal is, therefore, to converge to the global optimum of a possibly nonlinear function or system. Similarly, in a PSO process, a swarm of particles (or agents), each of which represents a potential solution to an optimization problem, navigate through the search space. The particles are initially distributed randomly over the search space with a random velocity and the goal is to converge to the global optimum of a function or a system. Each particle keeps track of its position in the search space and its best solution so far achieved. This is the personal best value (the so-called *pbest* in [24]) and the PSO process also keeps track of the global best solution so far achieved by the swarm by remembering the index of the best particle (the so called *gbest* in [24]). During their journey with discrete time iterations, the velocity of each agent in the next iteration is affected by the best position of the swarm (the best position of the particle *gbest* as the *social* component), the best personal position of the particle (*pbest* as the *cognitive* component), and its current velocity (the *memory* term). Both *social* and *cognitive* components contribute randomly to the velocity of the agent in the next iteration.

There are some efforts for simulating dynamic environments in a standard and configurable way. Some early works such as [2], [3] and [15] use experimental setup introduced by Angeline in [2]. In this setup the minimum of the three-dimensional parabolic function,  $f(x, y, z) = x^2 + y^2 + z^2$ , is moved along a linear or circular trajectory or randomly. Three different update frequencies (200, 1000 and 2000 evaluations) and change severities (0.01, 0.1, 0.5) are used. However, this setup enables testing only in a uni-modal environment. Branke in [10] has provided a publicly available Moving Peaks Benchmark (MPB) to enable different dynamic optimization algorithms to be tested in a standard way in a multi-modal environment. MPB allows creation of different dynamic fitness functions consisting of a number of peaks with varying location, height and width. The primary measure for performance evaluation is offline error, which is the average difference between the optimum and the best evaluation since the last environment change. Obviously, this value is always a positive number and it is zero only for perfect tracking. Several PSO methods are developed and tested using MPB such as [8], [9], [29], and [33]. Particularly Blackwell and Branke in [8] proposed a successful multi-swarm approach. The idea behind this is that different swarms can converge to different peaks and track them when the environment changes. The swarms interact only by mutual repulsion that keeps two swarms from converging to the same peak.

Similar to the aforementioned EAs, PSO might exhibit some major problems and severe drawbacks such as parameter dependency [31] and loss of diversity [40]. Particularly the latter phenomenon increases the probability of being trapped in local optima and it is the main source of premature convergence problem especially when the dimensionality of the search space is large [47] and the problem to be optimized is multi-modal [17], [40]. Another reason for the premature convergence is that particles are flown through a single point, which is (randomly) determined by *gbest* and *pbest* positions and this point is not even guaranteed to be a local optimum [45]. Since PSO was proposed for static problems in general, effects of such drawbacks eventually become much more severe for dynamic environments. Various modifications and PSO variants have been proposed in order to address these

problems such as [1], [6], [11]-[14], [20]-[23], [27]-[32], [36]-[43], [45], [46] and [49]-[53]. Such methods usually try to improve the diversity among the particles and the search mechanism either by changing the update equations towards a more diversified versions or adding more randomization to the system (to particle velocities, positions, etc.) or simply resetting some or all of them randomly when some conditions are met. However, their performance improvement might be quite limited even in static environments and most of them use *more* parameters and/or thresholds to accomplish this whilst making the PSO variant even more parameter dependent. Therefore, they do not set a reliable solution for dynamic environments, which usually have multi-modal nature and high dimensionality.

Another major drawback of the basic PSO and the aforementioned variants is that they can only be applied to a search (solution) space with a fixed dimensionality. However, in many optimization problems, the optimum dimension is also unknown (e.g. data clustering, object extraction, optimization of the dynamic functions, etc.) and should thus be determined within the PSO process. Take for example the color-based *image* segmentation as a data clustering application, where the optimum dimension of the solution space corresponds to the true number of clusters (segments) in the data (color) space, which cannot be known beforehand. In such case the PSO process should perform a multi-dimensional search in order to determine both true (optimum) number of clusters and the optimum centroid location for each cluster. The problem becomes even more challenging when it is applied over a dynamic environment such as a *video* where both number of clusters (segments) and their centroids (dominant colors) are in due change over time. Yet since the change between consecutive frames is not drastic rather minor, instead of performing a new clustering in the color domain via multi-dimensional search for each frame in the video, for a new (next) frame, the PSO process can establish a follow-up mechanism to track the optimum (number of) clusters (segments) from the previous frame. Therefore, using the past history about global and local optima becomes a crucial information to search for the current optima (both dimension and location).

In this paper, we shall first introduce a recent technique that significantly improves the global convergence performance of PSO by forming an artificial Global Best particle (*aGB*) fractionally. This algorithm, the so-called Fractional Global Best Formation (FGBF), collects the best dimensional components from each swarm particle and fractionally creates the *aGB*, which will replace *gbest* as a guide for the swarm, if it turns out to be better than the swarm's native *gbest* particle. We then propose a novel multi-swarm algorithm, which combines multi-swarms with the FGBF technique so that each swarm can apply FGBF distinctively. Via applying the proposed techniques on conventional MPB we shall show that they can find and track the global peak significantly in an efficient way and usually in earlier stages. For the multi-dimensional dynamic environments where the optimum dimension also changes over time, we shall then introduce a Multi-Dimensional (MD) PSO technique, which re-forms the native structure of swarm particles in such a way that they can make inter-dimensional passes with a dedicated dimensional PSO process. Therefore, in a multi-dimensional search space where the optimum dimension is unknown, swarm particles can seek for both positional and dimensional optima. This eventually pushes the frontier of the optimization problems in dynamic environments, towards a *global* search in a multi-dimensional space, where the problems in each dimension are possibly multi-modal and dependent on each other in a certain manner. Since the conventional MPB is created for a unique (fixed) dimensionality, we shall propose an extension of the benchmark within a dimension range in which there exists a unique optimum dimension where the global (highest) peak is located. In order to provide a certain degree of dependency among individual dimensions, peaks in different dimensions share the common coordinates of the peak locations. This is basically accomplished by subtracting a penalty term, whose magnitude depends on a dimensional error, from the landscape height in non-

optimal dimensions. As a result, over the extended MPB, MD PSO can seek for both optimum dimension and the global peak on it simultaneously. FGBF, multi-swarms and MD PSO are generic and independent, i.e. one can be performed without others but we shall show when the multi-dimensional search is needed, the best performance is achieved by their mutual operation. In recent works, both MD PSO and FGBF have been successfully applied over *static* problems such as general data clustering [25], and evolutionary artificial neural networks [21], respectively.

The rest of the paper is organized as follows. Section II surveys related work on PSO and MPB. The proposed techniques, MD PSO, multi-swarms with FGBF and their applications over the (extended) MPB are presented in detail in Section III. Section IV provides the experiments conducted and discusses the results. Finally, Section V concludes the paper.

## II. RELATED WORK

### A. The Basic PSO algorithm

In the basic PSO method, (*bPSO*), a swarm of particles flies through an  $N$ -dimensional search space where each particle represents a potential solution to the optimization problem. Each particle with index  $a$  in the swarm,  $\xi = \{x_1, \dots, x_a, \dots, x_S\}$ , is represented by the following characteristics:

$x_{a,j}(t)$  :  $j^{\text{th}}$  dimensional component of the position of particle  $a$ , at time  $t$

$v_{a,j}(t)$  :  $j^{\text{th}}$  dimensional component of the velocity of particle  $a$ , at time  $t$

$y_{a,j}(t)$  :  $j^{\text{th}}$  dimensional component of the personal best (*pbest*) position of particle  $a$ , at time  $t$

$\hat{y}_j(t)$  :  $j^{\text{th}}$  dimensional component of the global best position of the swarm, at time  $t$

Let  $f$  denote the fitness function to be optimized. Without loss of generality assume that the objective is to find the maximum of  $f$  in an  $N$ -dimensional space. Then the personal best of particle  $a$  can be updated at iteration  $t$  as,

$$y_{a,j}(t) = \begin{cases} y_{a,j}(t-1) & \text{if } f(x_a(t)) < f(y_a(t-1)) \\ x_{a,j}(t) & \text{else} \end{cases} \quad j = 1, 2, \dots, N \quad (1)$$

Since  $gbest$  is the index of the GB particle,  $\hat{y}(t) = y_{gbest}(t)$ . Then at each iteration in a PSO process, positional updates are performed for each dimensional component,  $j \in \{1, N\}$  and for each particle,  $a \in \{1, S\}$ , as follows:

$$\begin{aligned} v_{a,j}(t+1) &= w(t)v_{a,j}(t) + c_1 r_{1,j}(t)(y_{a,j}(t) - x_{a,j}(t)) + c_2 r_{2,j}(t)(\hat{y}_j(t) - x_{a,j}(t)) \\ x_{a,j}(t+1) &= x_{a,j}(t) + v_{a,j}(t+1) \end{aligned} \quad (2)$$

where  $w$  is the inertia weight, [42] and  $c_1, c_2$  are the acceleration constants which are usually set to 1.49 or 2.

$r_{1,j} \sim U(0,1)$  and  $r_{2,j} \sim U(0,1)$  are random variables with uniform distribution. Recall from the earlier discussion that the first term in the summation is the *memory* term, which represents the role of previous velocity over the current velocity, the second term is the *cognitive* component, which represents the particle's own experience and the third term is the *social* component through which the particle is "guided" by the *gbest* particle towards the GB solution so far obtained. Although the use of inertia weight,  $w$ , was later added by Shi and Eberhart [42], into the velocity update equation, it is widely accepted as the basic form of PSO algorithm. A larger value of  $w$  favors exploration while a small inertia weight favors exploitation. As originally introduced,  $w$  is often linearly decreased from a high value (e.g. 0.9) to a low value (e.g. 0.4) during iterations of a PSO run. Depending on the problem to

be optimized, PSO iterations can be repeated until a specified number of iterations, say  $IterNo$ , is exceeded, velocity updates become zero, or the desired fitness score is achieved (i.e.  $f > \varepsilon_C$ ). Accordingly the general pseudo-code of the  $bPSO$  can be given as follows:

**$bPSO$**  ( termination criteria:  $\{IterNo, \varepsilon_C, \dots\}, V_{\max}$  )

1. For  $\forall a \in \{1, S\}$  do:
  - 1.1. Randomize  $x_a(1), v_a(1)$
  - 1.2. Let  $y_a(0) = x_a(1)$
  - 1.3. Let  $\hat{y}(0) = x_a(1)$
2. End For.
3. For  $\forall t \in \{1, IterNo\}$  do:
  - 3.1. For  $\forall a \in \{1, S\}$  do:
    - 3.1.1. Compute  $y_a(t)$  using (1)
    - 3.1.2. If  $(f(y_a(t)) > \max_{1 \leq i < a} (f(\hat{y}(t-1)), f(y_i(t))))$  then  $gbest = a$  and  $\hat{y}(t) = y_a(t)$
  - 3.2. End For.
  - 3.3. If any termination criterion is met, then **Return**.
  - 3.4. For  $\forall a \in \{1, S\}$  do:
    - 3.4.1. For  $\forall j \in \{1, N\}$  do:
      - 3.4.1.1. Compute  $v_{a,j}(t+1)$  using (2)
      - 3.4.1.2. If  $(|v_{a,j}(t+1)| > V_{\max})$  then clamp it to  $|v_{a,j}(t+1)| = V_{\max}$
      - 3.4.1.3. Compute  $x_{a,j}(t+1)$  using (2)
    - 3.4.2. End For.
  - 3.5. End For.
4. End For.
5. **Return**.

Velocity clamping also called as “dampening” with the user-defined maximum range  $V_{\max}$  (and  $-V_{\max}$  for the minimum) as in step 3.4.1.2 is one of the earliest attempts to control or prevent oscillations. Such oscillations are indeed crucial since they broaden the search capability of the swarm; however, they have a potential drawback of oscillating continuously around the optimum point. Therefore, such oscillations should be dampened and the convergence is achieved with the proper use of the velocity clamping and the inertia weight. Furthermore, this is the  $bPSO$  algorithm where the particle  $gbest$  is determined within the entire swarm. Another major topological approach, the so-called  $lbest$ , also exists where the swarm is divided into overlapping neighborhoods of particles and instead of defining  $gbest$  and  $\hat{y}(t) = y_{gbest}(t)$  over the entire swarm, for a particular neighborhood  $N_i$ , the (local) best particle is referred as  $lbest$  with the position  $\hat{y}_i(t) = y_{lbest}(t)$ . Neighbors can be defined with respect to particle indices (i.e.  $i \in \{j-l, j+l\}$ ) or by using some other topological forms [44]. It is obvious that  $gbest$  is a special case of  $lbest$  scheme where the neighborhood is defined as the entire swarm. The  $lbest$  approach is one of the earlier attempts, which usually improves the diversity; however, it is slower than the  $gbest$  approach. PSO variants for dynamic environments will be covered in the next sections.

### B. Moving Peaks Benchmark

Conceptually speaking, MPB developed by Branke in [10], is a simulation of a configurable dynamic environment changing over time. The environment consists of a certain number peaks with varying location, height and width. The dimensionality of the fitness function is fixed in advance and thus is an input parameter of the benchmark. An  $N$ -dimensional fitness function with  $m$  peaks is expressed as,

$$F(\vec{x}, t) = \max \left( B(\vec{x}), \max_{p=1 \dots m} P(\vec{x}, h_p(t), w_p(t), \vec{c}_p(t)) \right) \quad (3)$$

where  $B(\vec{x})$  is time invariant basis landscape, whose utilization is optional, and  $P$  is the function defining the height of the  $p^{\text{th}}$  peak at location  $\vec{x}$ , where each of the  $m$  peaks can have its own dynamic parameters such as height,  $h_p(t)$ , width,  $w_p(t)$  and location vector of the peak center,  $\vec{c}_p(t)$ . Each peak parameter can be initialized randomly or set to a certain value and after a time period (number of evaluations),  $T_e$ , at a time (evaluation)  $t$ , a change over a single peak,  $p$ , can be defined as follows:

$$\begin{aligned} h_p(t) &= h_p(t - T_e) + r\Delta h \\ w_p(t) &= w_p(t - T_e) + r\Delta w \\ \vec{c}_p(t) &= \vec{c}_p(t - T_e) + \vec{v}_p(t) \end{aligned} \quad (4)$$

where  $r \in N(0,1)$ ,  $\Delta h$  and  $\Delta w$  are the height and width severities and  $\vec{v}_p(t)$  is the normalized shift vector, which is a linear combination of a random vector and the previous shift vector,  $\vec{v}_p(t - T_e)$ . Type and number of peaks along with their initial heights and widths, environment (search space) dimension and size, change severity, level of change randomness and change frequency can be defined. To facilitate standard comparative evaluations among different algorithms, three standard settings of such MPB parameters, the so called *Scenarios*, have been defined. *Scenario 2* is the most widely used. Each scenario allows a range of values, among them the following are commonly used: number of peaks = 10, change severity  $vlength = 1.0$ , correlation  $lambda = 0.0$  and peak change frequency = 5000. In *Scenario 2* no basis landscape is used and peak type is a simple *cone* with the following expression,

$$\begin{aligned} P(\vec{x}, h_p(t), w_p(t), \vec{c}_p(t)) &= h_p(t) - s_p(t) \|\vec{x} - \vec{c}_p(t)\| \quad \text{where} \\ s_p(t) &= \frac{h_p(t)}{w_p(t)} \quad \text{and} \quad \|\vec{x} - \vec{c}_p(t)\| = \sqrt{\sum_{i=1}^N (x_i - c_{pi})^2} \quad \forall x_i \in \vec{x}, \forall c_{pi} \in \vec{c}_p(t) \end{aligned} \quad (5)$$

where  $s_p(t)$  is the slope and  $\|\cdot\|$  is the *Euclidean* distance between two  $N$ -dimensional vectors,  $\vec{x}$  and  $\vec{c}_p(t)$ . More detailed information on MPB and the rest of the parameters used in this benchmark can be obtained from [10].

### C. Multi-swarm PSO

The main problem of using the basic PSO algorithm in a dynamic environment is that eventually the swarm will converge to a single peak – whether global or local. When another peak becomes the global maximum as a result of an environmental change, it is likely that the particles keep circulating close to the peak to which the swarm has converged and thus they cannot find the new global maximum. Blackwell and Branke have addressed this problem in [8] and [9] by introducing multi-swarms that are actually separate PSO processes. Each particle is now a member of one of the swarms only and it is unaware of other swarms. The main idea is that each swarm can converge to a separate peak. Swarms interact only by mutual repulsion that keeps them from converging to the same peak. For a single swarm it is essential to maintain enough diversity so that the swarm can track small

location changes of the peak to which it is converging. For this purpose Blackwell and Branke introduced charged and quantum swarms, which are analogues to an atom having a nucleus and charged particles randomly orbiting it. The particles in the nucleus take care of the fine tuning of the result while the charged particles are responsible of detecting the position changes. However, it is clear that, instead of charged or quantum swarms, some other method can also be used to ensure sufficient diversity among particles of a single swarm so that the peak can be tracked despite of small location changes.

As one might expect, the best results are achieved when the number of swarms is set equal to the number of peaks. However, it is then required that the number of peaks is known beforehand. In [7] Blackwell presents self-adapting multi-swarms, which can be created or removed during the PSO process and, therefore, it is not necessary to fix the number swarms beforehand.

The repulsion between swarms is realized by simply re-initializing the worse of two swarms if they move within a certain range from each other. Using physical repulsion could lead to equilibrium, where swarm repulsion prevents both swarms from getting close to a peak. A proper limit closer to which the swarms are not allowed to move,  $r_{rep}$  is attained by using the average radius of the peak basin,  $r_{bas}$ . If  $p$  peaks are evenly distributed in  $X^N$ ,  $r_{rep} = r_{bas} = X / p^{1/N}$ . Detailed information about multi-swarms can be obtained in [8] and [9].

### III. THE PROPOSED OPTIMIZATION TECHNIQUE FOR DYNAMIC ENVIRONMENTS

In this section, we first introduce the multi-swarms with FGBF technique and its application to MPB. The multi-dimensional extension of PSO, the so-called MD PSO, will be detailed next. Finally we show their mutual application over the (extended) MPB.

#### A. FGBF Technique

For those problems where dimensional fitness evaluation is possible, Fractional Global Best Formation (FGBF) can be used to avoid the premature convergence by providing a significant diversity obtained from a proper *fusion* of the swarm's best components (the individual dimension(s) of the current position of each particle in the swarm). Some problems such as data clustering, on the other hand, can exhibit some kind of non-trivial interdependency among the dimensional components of the search space. For them, a proper *approximation* for the fractional fitness evaluation, if possible, should be designed to perform FGBF. In either case, FGBF fractionally creates an artificial GB at each iteration in a PSO process by selecting best particle (dimensional) components from the entire swarm. Therefore, especially during the initial steps, the FGBF can be and, most of the time, is a better alternative than the native *gbest* particle since it has the advantage of assessing each dimension of every particle in the swarm individually, and forming the *aGB* particle fractionally by using the best components among them. This process naturally uses the available diversity among individual dimensional components and thus it can prevent swarm from being trapped in local optima due to its ongoing and ever-varying particle creations. At each iteration FGBF is performed after the assignment of the swarm's *gbest* particle (i.e. performed between steps 3.2 and 3.3 in the pseudo-code of *bPSO*) and, if *aGB* turns out to be better than *gbest*, the personal best location of the *gbest* particle is replaced by the location of the *aGB* particle and, since  $\hat{y}(t) = y_{gbest}(t)$ , the artificially created particle is thus used to guide the swarm through the *social* component in (2). In other words, the swarm will be guided only by the best (winner) between native *gbest* and the *aGB* particle at any time. In the next iteration, a new *aGB* particle is

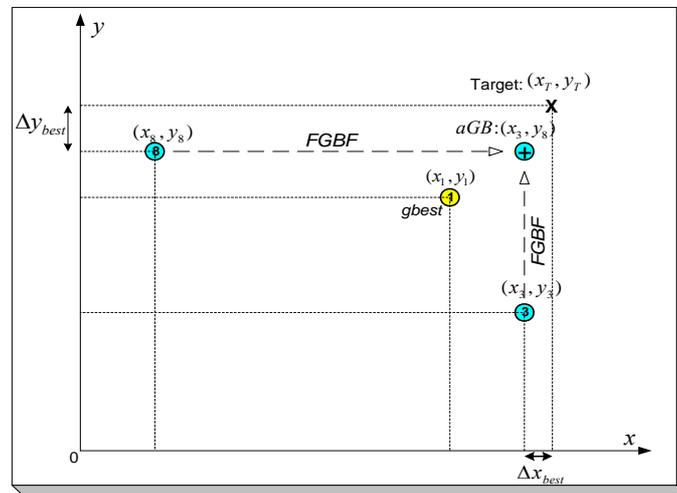
created and it will again compete against the personal best of  $gbest$  (which can be also a former  $aGB$  now).

Suppose that for a swarm  $\xi$ , FGBF is performed in a PSO process in dimension  $N$ . Recall from the earlier discussion that in a particular iteration,  $t$ , each PSO particle,  $a$ , has the following components: position  $(x_{a,j}(t))$ , velocity  $(v_{a,j}(t))$  and the personal best position  $(y_{a,j}(t))$ ,  $j \in \{1, N\}$ . As the  $aGB$  particle is fractionally (re-)created from the individual dimensions of some swarm particles at each iteration, it does not need the velocity term and, therefore, it does not have to remember its personal best location.

Let  $f(a, j)$  be the dimensional fitness score of the  $j^{\text{th}}$  component of the position of particle  $a$  and  $f(gbest, j)$  be the dimensional fitness score of the  $j^{\text{th}}$  component of the personal best position of the  $gbest$  particle. Suppose that all dimensional fitness scores  $(f(a, j), \forall a \in \{1, S\} \text{ and } f(gbest, j))$  can be computed in step 3.1 and FGBF can then be plugged in between steps 3.2 and 3.3 of  $bPSO$ 's pseudo-code. Accordingly, the pseudo-code for FGBF can be expressed as follows:

#### FGBF in $bPSO$ ( $\xi, f(a, j)$ )

1. Let  $a[j] = \arg \max_{a \in \xi, j \in \{1, N\}} (f(a, j))$  be the index of particle yielding the maximum  $f(a, j)$  for the  $j^{\text{th}}$  dimensional component.
2.  $x_{aGB,j}(t) = x_{a[j],j}(t)$  for  $\forall j \in \{1, N\}$
3. If  $f(gbest, j) > f(a[j], j)$  then  $x_{aGB,j}(t) = y_{gbest,j}(t)$
4. If  $(f(x_{aGB}(t)) > f(y_{gbest}(t)))$  then  $y_{gbest}(t) = x_{aGB}(t)$  and  $\hat{y}(t) = x_{aGB}(t)$
5. **Return.**



**Figure 1: A sample FGBF operation in 2D space.**

Note that Step 1 along with the computation of  $f(a, j)$  depends entirely on the optimization problem. It keeps track of partial fitness contributions from each individual dimensional component of each particle's position (the potential solution). Take for instance the function minimization problem as illustrated in Figure 1 where 2D space is used for illustration purposes. In the figure, three particles in a swarm are ranked as the 1<sup>st</sup> (or the  $gbest$ ), the 3<sup>rd</sup> and the 8<sup>th</sup> with respect to their proximity to the target position (or the global solution) of some function. Although  $gbest$  particle (i.e. 1<sup>st</sup> rank particle) is the closest in the overall sense, the particles ranked 3<sup>rd</sup> and 8<sup>th</sup> provide the best  $x$  and  $y$  dimensions (closest to the target's respective dimensions) in the entire swarm and hence the  $aGB$  particle via FGBF yields a better (closer) particle than the swarm's native  $gbest$ . Particularly in [25], the usage and

merits of FGBF for the optimization of several multi-modal, non-linear (benchmark) functions in high dimensions have been shown where all MD PSO runs with FGBF found the global minimum at the target dimension for all runs, over all functions, regardless of the dimension, swarm size and modality, and without any exception. This is the case where the function to be optimized is known a priori and FGBF is easier to use. For those problems where the knowledge of how each fractional dimension contributes to the overall fitness is no longer available (i.e. black-box problems), FGBF can still be adapted with respect to the problem, e.g. in [25], see the usage of FGBF over general data clustering where the optimum dimension of the search space (number of clusters) is also unknown and hence MD PSO with the proper application of FGBF is used to find out the optimum (number of) clusters in a data space.

### B. FGBF Application for MPB

The previous section introduced the principles of FGBF theory within a *bPSO* process in a single dimension and referred some of its applications in other domains, each of which is in a static environment. However, in dynamic environments this approach eventually leads the swarm to converge to a single peak (whether global or local) and therefore, it may lose its ability to track other peaks. As any of the peaks can become the optimum peak as a result of environmental changes, it is likely to lead to a suboptimal convergence. This is the basic reason of utilizing the multi-swarms along with the FGBF operation. As described in Section II.C, the mutual repulsion between swarms is performed and for computing the distance between two swarms, we use distance between the swarms' global best locations. Instead of charged or quantum swarms, FGBF is the entire mechanism to provide necessary diversity and thus to enable peak tracking if peaks' location are changed. We also re-initialize the particle velocities after each environment change to further contribute to the diversity.

Each particle with index  $a$  in a swarm  $\xi$ , represents a potential solution and therefore, the  $j^{\text{th}}$  component of an  $N$ -dimensional point  $(x_j, j \in \{1, N\})$  is stored in its positional component,  $x_{a,j}(t)$ , at a time  $t$ . The aim of the PSO process is to search for the global optimum point, which maximizes  $P(\vec{x}, h_p(t), w_p(t), \vec{c}_p(t))$ , in other words, finding the global (highest) peak in MPB environment. Recall that in *Scenario 2* of MPB the peaks used are all in *cone* shape, as given in (5). Since in (5),  $h_p(t)$  and  $s_p(t)$  are both set by MPB, finding the highest peak is equivalent to minimizing the  $\|\vec{x} - \vec{c}_p(t)\|$  term, yielding  $f(a, j) = -(x_j - c_{pj})^2$ . Step 3.1 in *bPSO*'s pseudo-code computes the (dimensional) fitness scores  $(f(a, j), f(gbest, j))$  of the  $j^{\text{th}}$  components  $(x_{a,j}, y_{gbest,j})$  and in Step 1 of the FGBF process, the dimensional component yielding maximum  $f(a, j)$  is then placed in  $aGB$ . In Step 3 these dimensional components are replaced by dimensional components of the personal best position of the *gbest* particle, if they yield higher dimensional fitness scores. We do not expect that dimensional fitness scores can be evaluated with respect to the optimum peak since this requires the *a priori* knowledge of the global optimum, instead we use either the *current* peak where the particle resides on or the peak to which the swarm is converging (swarm peak). We shall thus consider and evaluate both *modes* separately.

### C. MD PSO Algorithm

Instead of operating in a fixed dimension  $N$ , the MD PSO algorithm is designed to seek both positional and dimensional optima within a dimension range,  $(D_{\min} \leq d \leq D_{\max})$ . In MD PSO each particle has two sets of components, each of which is subjected to one of two independent and consecutive processes. The first one is the

regular positional PSO, i.e. the traditional velocity updates and due positional shifts in  $d$ -dimensional search space. The second one is a dimensional PSO, which allows the particle to navigate through dimensions. Accordingly, each particle keeps track of its last position, velocity and personal best position ( $pbest$ ) in a particular dimension so that when it re-visits the same dimension at a later time, it can perform its regular “positional” fly using this information. The dimensional PSO process of each particle may then move the particle to another dimension where it will remember its earlier positional status and keep “flying” within the positional PSO process in this dimension, and so on. The swarm, on the other hand, keeps track of the  $gbest$  particles in all dimensions, each of which respectively indicates the best (global) position so far achieved and can thus be used in the regular velocity update equation for that dimension. Similarly the dimensional PSO process of each particle uses its personal best dimension in which the personal best fitness score has so far been achieved. Finally, the swarm keeps track of the global best dimension,  $dbest$ , among all the personal best dimensions. The  $gbest$  particle in  $dbest$  dimension represents the optimum solution and dimension, respectively.

In a MD PSO process at time (iteration)  $t$ , a particle with index  $a$  in the swarm,  $\xi = \{x_1, \dots, x_a, \dots, x_S\}$ , is represented by the following characteristics:

$xx_{a,j}^{xd_a(t)}(t)$  :  $j^{\text{th}}$  component (dimension) of the position of particle  $a$ , in dimension  $xd_a(t)$

$vx_{a,j}^{xd_a(t)}(t)$  :  $j^{\text{th}}$  component (dimension) of the velocity of particle  $a$ , in dimension  $xd_a(t)$

$xy_{a,j}^{xd_a(t)}(t)$  :  $j^{\text{th}}$  component (dimension) of the personal best ( $pbest$ ) position of particle  $a$ , in dimension  $xd_a(t)$

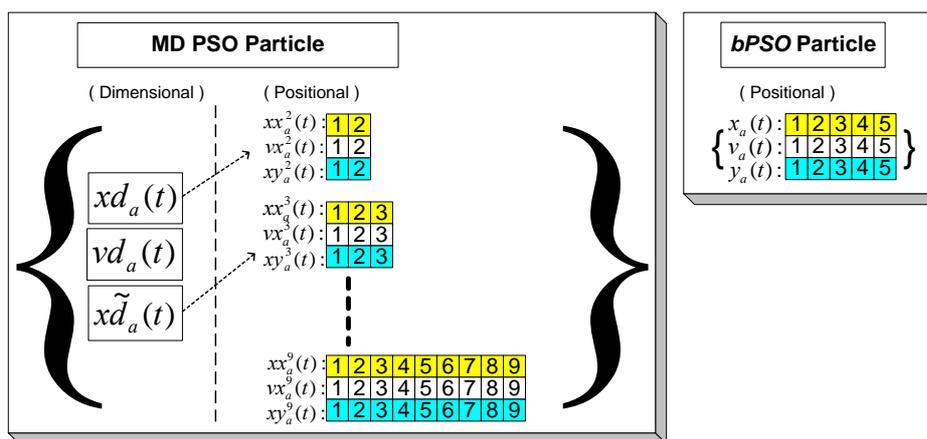
$gbest(d)$  : Global best particle index in dimension  $d$

$x\hat{y}_j^d(t)$  :  $j^{\text{th}}$  component (dimension) of the global best position of swarm, in dimension  $d$

$xd_a(t)$  : Current dimension of particle  $a$

$vd_a(t)$  : Dimensional velocity of particle  $a$

$x\tilde{d}_a(t)$  : Personal best dimension of particle  $a$



**Figure 2: Sample MD PSO (right) vs. bPSO (left) particle structures. For MD PSO  $\{D_{\min} = 2, D_{\max} = 9\}$  and at the current time  $t$ ,  $xd_a(t) = 2$  and  $x\tilde{d}_a(t) = 3$ . For bPSO  $N=5$ .**

Figure 2 shows sample MD PSO and bPSO particles with indices  $a$ . bPSO particle that is at a (fixed) dimension,  $N=5$ , contains only positional components whereas MD PSO particle contains both positional and dimensional components respectively. In the figure the dimension range for the MD PSO is given in between 2 to 9, therefore

the particle contains 8 sets of positional components. In this example the current dimension where the particle with index  $a$  resides is 2 ( $xd_a(t)=2$ ) whereas its personal best dimension is 3 ( $x\tilde{d}_a(t)=3$ ). Therefore, at time  $t$ , a positional PSO update is first performed over the positional elements,  $xx_a^2(t)$  and then the particle may move to another dimension with respect to the dimensional PSO. Recall that each positional element,  $xx_{a,j}^2(t)$ ,  $j \in \{1,2\}$ , represents a potential solution in the data space of the problem.

Let  $f$  denote the fitness function that is to be optimized within a certain dimension range,  $\{D_{\min}, D_{\max}\}$ . In accordance with the scope of the current work and without loss of generality assume that the objective is to find the maximum (position) of  $f$  in the optimum dimension within a multi-dimensional search space. Assume that the particle  $a$  visits (back) the same dimension after  $T$  iterations (i.e.  $xd_a(t)=xd_a(t+T)$ ), then the personal best position can be updated in iteration  $t+T$  as follows,

$$xy_{a,j}^{xd_a(t+T)}(t+T) = \begin{cases} xy_{a,j}^{xd_a(t)}(t) & \text{if } f(xx_a^{xd_a(t+T)}(t+T)) < f(xy_a^{xd_a(t)}(t)) \\ xx_{a,j}^{xd_a(t+T)}(t+T) & \text{else} \end{cases} \quad j=1,2,\dots,xd_a(t+T) \quad (6)$$

To overcome the necessity to remember the iteration numbers, when a particle has visited a certain dimension last time, the following updates are performed for each dimensional component except  $xd_a(t)$ , i.e.  $j \in \{1,d\}$  where  $d \in \{D_{\min}, D_{\max}\} - \{xd_a(t)\}$  and for each particle,  $a \in \{1,S\}$ ,

$$\begin{aligned} xy_{a,j}^d(t) &= xy_{a,j}^d(t-1), x\hat{y}_j^d(t) = x\hat{y}_j^d(t-1), \\ \text{for } \forall d &\in \{D_{\min}, D_{\max}\} - \{xd_a(t)\} \end{aligned} \quad (7)$$

Furthermore, the personal best dimension of particle  $a$  can be updated in iteration  $t+1$  as follows,

$$x\tilde{d}_a(t) = \begin{cases} x\tilde{d}_a(t-1) & \text{if } f(xx_a^{xd_a(t)}(t)) < f(xy_a^{x\tilde{d}_a(t-1)}(t-1)) \\ xd_a(t) & \text{else} \end{cases} \quad (8)$$

Recall that  $gbest(d)$  is the index of the global best particle in dimension  $d$  and so  $x\hat{y}^d(t) = xy_{gbest(d)}^d(t)$ . For a particular iteration  $t$ , and for a particle with index  $a \in \{1,S\}$ , first the positional components are updated in its current dimension,  $xd_a(t)$  and then the dimensional update is performed to determine its next dimension,  $xd_a(t+1)$ . The positional update is performed for each dimension component,  $j \in \{1,xd_a(t)\}$ , as follows:

$$\begin{aligned} vx_{a,j}^{xd_a(t)}(t+1) &= w(t)vx_{a,j}^{xd_a(t)}(t) + c_1r_{1,j}(t)(xy_{a,j}^{xd_a(t)}(t) - xx_{a,j}^{xd_a(t)}(t)) + c_2r_{2,j}(t)(x\hat{y}_j^{xd_a(t)}(t) - xx_{a,j}^{xd_a(t)}(t)) \\ xx_{a,j}^{xd_a(t)}(t+1) &= xx_{a,j}^{xd_a(t)}(t) + C_{vx} \left[ vx_{a,j}^{xd_a(t)}(t+1), \{V_{\min}, V_{\max}\} \right] \\ xx_{a,j}^{xd_a(t)}(t+1) &\leftarrow C_{xx} \left[ xx_{a,j}^{xd_a(t)}(t+1), \{X_{\min}, X_{\max}\} \right] \end{aligned} \quad (9)$$

where  $C_{xx}[\dots] \equiv C_{vx}[\dots]$  are the clamping operators applied over each positional component,  $xx_{a,j}^d$  and  $vx_{a,j}^d$ .  $C_{xx}[\dots]$  may or may not be applied depending on the optimization problem but  $C_{vx}[\dots]$  is basically needed to avoid exploding. Each operator can be applied in two different ways, such as,

$$C_{xx}[xx_{a,j}^d(t), \{X_{\min}, X_{\max}\}] = \left\{ \begin{array}{ll} xx_{a,j}^d(t) & \text{if } X_{\min} \leq xx_{a,j}^d(t) \leq X_{\max} \\ X_{\min} & \text{if } xx_{a,j}^d(t) < X_{\min} \\ X_{\max} & \text{if } xx_{a,j}^d(t) > X_{\max} \end{array} \right\} \quad (a) \quad (10)$$

$$C_{xx}[xx_{a,j}^d(t), \{X_{\min}, X_{\max}\}] = \left\{ \begin{array}{ll} xx_{a,j}^d(t) & \text{if } X_{\min} \leq xx_{a,j}^d(t) \leq X_{\max} \\ U(X_{\min}, X_{\max}) & \text{else} \end{array} \right\} \quad (b)$$

where the option (a) is a simple thresholding to the range limits and (b) re-initializes randomly the  $j^{\text{th}}$  positional component ( $j \leq d$ ).

Note that the particle's new position,  $xx_a^{xd_a(t)}(t+1)$ , will still be in the same dimension,  $xd_a(t)$ ; however, the particle *may* jump to another dimension afterwards. Therefore, for all the other dimensions in the dimensional range except  $xd_a(t)$ , i.e.  $\forall d \in \{D_{\min}, D_{\max}\} - \{xd_a(t)\}$ , the following updates are necessary for each dimensional component,  $j \in \{1, d\}$  and for each particle,  $a \in \{1, S\}$ :

$$\begin{aligned} vx_{a,j}^d(t+1) &= vx_{a,j}^d(t), xx_{a,j}^d(t+1) = xx_{a,j}^d(t), \\ \text{for } \forall d &\in \{D_{\min}, D_{\max}\} - \{xd_a(t)\} \end{aligned} \quad (11)$$

The dimensional updates are computed with the following equations:

$$\begin{aligned} vd_a(t+1) &= \lfloor vd_a(t) + c_1 r_1(t)(x\tilde{d}_a(t) - xd_a(t)) + c_2 r_2(t)(dbest - xd_a(t)) \rfloor \\ xd_a(t+1) &= xd_a(t) + C_{vd}[vd_a(t+1), \{VD_{\min}, VD_{\max}\}] \\ xd_a(t+1) &\leftarrow C_{xd}[xd_a(t+1), \{D_{\min}, D_{\max}\}] \end{aligned} \quad (12)$$

where  $\lfloor \cdot \rfloor$  is the *floor* operator,  $C_{xd}[\cdot, \cdot]$  and  $C_{vd}[\cdot, \cdot]$  are the clamping operators applied over dimensional components,  $xd_a(t)$  and  $vd_a(t)$ , respectively. As in (2), we employ the inertia weight for positional velocity update; however, we have witnessed no benefit of using it for dimensional PSO, and hence we left it out of (12) for the sake of simplicity.  $C_{vd}[\cdot, \cdot]$  is similar to  $C_{vx}[\cdot, \cdot]$ , which is basically applied to avoid exploding and we use the basic thresholding for this, expressed as follows:

$$C_{vd}[vd_a(t), \{VD_{\min}, VD_{\max}\}] = \left\{ \begin{array}{ll} vd_a(t) & \text{if } VD_{\min} \leq vd_a(t) \leq VD_{\max} \\ VD_{\min} & \text{if } vd_a(t) < VD_{\min} \\ VD_{\max} & \text{if } vd_a(t) > VD_{\max} \end{array} \right\} \quad (13)$$

$C_{xd}[\cdot, \cdot]$ , on the other hand, is a mandatory clamping operator, which keeps the dimensional jumps within the dimension range of the problem,  $\{D_{\min}, D_{\max}\}$ . Furthermore within  $C_{xd}[\cdot, \cdot]$ , an *optional* in-flow buffering mechanism can also be implemented. This can be a desired property, which allows only sufficient number of particles in a certain dimension and thus avoids the redundancy. Particularly *dbest* and dimensions within the close proximity have a natural attraction and without such buffering mechanism, the majority of swarm particles may be hosted within this local neighborhood and hence other dimensions might encounter a severe depletion. To prevent this, the buffering mechanism should control the in-flow of the particles (by the dimensional velocity updates) to a particular dimension. On some early *bPSO* implementations over problems with low (and fixed) dimensions, 15-20 particles were usually sufficient for a successful operation. However, in high dimensions this may not be so since more particles are usually needed as the dimension increases. Therefore, we empirically set the limit as to be

proportional with the solution space dimension and not less than 15. At a time  $t$ , let  $P_d(t)$  be the number of particles in dimension  $d$ .  $C_{xd}[\dots]$  can then be expressed with the (optional) buffering mechanism as follows:

$$C_{xd}[xd_a(t), \{D_{\min}, D_{\max}\}] = \left. \begin{cases} xd_a(t-1) & \text{if } P_{xd_a(t)}(t) \geq \max(15, xd_a(t)) \\ xd_a(t-1) & \text{if } xd_a(t) < D_{\min} \\ xd_a(t-1) & \text{if } xd_a(t) > D_{\max} \\ xd_a(t) & \text{else} \end{cases} \right\} \quad (14)$$

In short, the clamping and buffering operator,  $C_{xd}[\dots]$ , allows a dimensional jump only if the target dimension is within dimensional range and has space for a newcomer. Accordingly, the general pseudo-code of the MD PSO method can be expressed as follows:

**MD PSO** ( *termination criteria*:  $\{IterNo, \epsilon_C, \dots\}$  )

1. For  $\forall a \in \{1, S\}$  do:
  - 1.1. Randomize  $xd_a(1), vd_a(1)$
  - 1.2. Initialize  $\tilde{xd}_a(0) = xd_a(1)$
  - 1.3. For  $\forall d \in \{D_{\min}, D_{\max}\}$  do:
    - 1.3.1. Randomize  $xx_a^d(1), xv_a^d(1)$
    - 1.3.2. Initialize  $xy_a^d(0) = xx_a^d(1)$
    - 1.3.3. Initialize  $x\hat{y}^d(0) = xx_a^d(1)$
  - 1.4. End For.
2. End For.
3. For  $\forall t \in \{1, IterNo\}$  do:
  - 3.1. For  $\forall a \in \{1, S\}$  do:
    - 3.1.1. If  $(f(xx_a^{xd_a(t)}(t)) > f(xy_a^{xd_a(t)}(t-1)))$  then do:
      - 3.1.1.1  $xy_a^{xd_a(t)}(t) = xx_a^{xd_a(t)}(t)$
      - 3.1.1.2 Update  $\tilde{xd}_a(t)$  according to (8)
    - 3.1.2. Else  $xy_a^{xd_a(t)}(t) = xy_a^{xd_a(t)}(t-1)$
    - 3.1.3. End If
    - 3.1.4. If  $(f(xx_a^{xd_a(t)}(t)) > \max(f(x\hat{y}^{xd_a(t)}(t-1), \max_{1 \leq p < a} (f(xx_p^{xd_a(t)}(t)))))$  then do:
      - 3.1.4.1.  $gbest(xd_a(t)) = a$
      - 3.1.4.2. If  $(f(xx_a^{xd_a(t)}(t)) > f(x\hat{y}^{dbest}(t-1)))$  then  $dbest = xd_a(t)$
    - 3.1.5. End If.
    - 3.1.6. Do updates in other dimensions according to (7)
  - 3.2. End For.
  - 3.3. If the *termination criteria* are met, then **Stop**.
  - 3.4. For  $\forall a \in \{1, S\}$  do:
    - 3.4.1. For  $\forall j \in \{1, xd_a(t)\}$  do:
      - 3.4.1.1. Compute  $v\tilde{x}_{a,j}^{xd_a(t)}(t+1)$  and  $xx_{a,j}^{xd_a(t)}(t+1)$  using (9)
      - 3.4.1.2. Update velocities and locations in other dimensions using (11)
    - 3.4.2. End For.
    - 3.4.3. Compute  $vd_a(t+1)$  and  $xd_a(t+1)$  using (12)
  - 3.5. End For.
4. End For.

Once the MD PSO process terminates, the optimum solution will be  $\hat{x}y^{dbest}$  in the optimum dimension,  $dbest$ , achieved by the particle with the index  $gbest(dbest)$  and finally the best (fitness) score achieved will naturally be  $f(\hat{x}y^{dbest})$ . Note that MD PSO is *only* a natural extension or a generic form of the basic PSO, as it searches for both optimum dimension and solution (in the optimum dimension). Note that this is not a ‘superiority’ in terms of *convergence*, rather the ability of searching the optimum solution space dimension while searching for the positional optimum, in a simultaneous way. In other words, the basic PSO is a MD PSO process in a *fixed* dimension, which basically means that *bPSO* is identical to MD PSO for a particular (fixed) dimension. However, on contrary to *bPSO*, due to its ability to simultaneously search for both optimum dimension and solution, MD PSO can conveniently be used in many applications where the optimum dimension is unknown. For instance in [25] MD PSO (with FGBF) has been used to find out the true number of clusters in a complex data space. This is obviously a tremendous advantage since with a traditional approach such as *K-means*, the number of clusters,  $K$ , has to be given *a priori* and this may not be possible for many complex problems. In another work [21], MD PSO has been used for the automatic design of Artificial Neural Networks (ANNs) by evolving to the optimal network configuration(s) for a particular problem. Thus MD PSO can seek positional optimum in the error space and dimensional optimum in the architecture space. The optimum dimension converged at the end of a MD PSO process corresponds to the best ANN configuration (or the most appropriate ANN for that problem) where the trained network parameters (connections, weights and biases) can then be resolved from the positional optimum reached on that dimension. This presents a significant advantage over many traditional training methods such as well-known back-propagation algorithm, which can only be used to train a single ANN with a fixed configuration.

#### D. The Proposed Techniques over Multi-Dimensional MPB

For testing the proposed multi-dimensional optimization technique, we extended Branke’s MPB into a multi-dimensional version, in which there exists many search space dimensions within a dimensional range  $\{D_{\min}, D_{\max}\}$ , and the optimal dimension changes over time in addition to the dynamic nature of the conventional MPB. Peak locations in different dimensions share the common coordinates in the fitness function and thus such an extension further allows exploitation of the information gathered in other search space dimensions.

The multi-dimensional extension of the MPB is simple. The initialization and changes of peak locations must now be done in the highest possible search space dimension,  $D_{\max}$ . Locations in the other dimensions can be obtained simply by leaving out the redundant coordinates (non-existing dimensions). The optimal dimension is chosen randomly every time the environment is changed. Therefore, the fitness function with  $m$  peaks in multi-dimensional environment can be expressed as,

$$F(\vec{x}^d, t) = \max \left( B(\vec{x}^d), \max_{p=1 \dots m} P(\vec{x}^d, d, h_p(t), w_p(t), \vec{c}_p^d(t)) \right) \quad (15)$$

where  $d \in \{D_{\min}, D_{\max}\}$  is the dimension of position  $\vec{x}^d$  and  $\vec{c}_p^d(t)$  refers to first  $d$  coordinates (dimensions) of the peak center location. A *cone* peak is now expressed with the following expression:

$$P(\vec{x}^d, h_p(t), w_p(t), \vec{c}_{pd}^d(t)) = h_p(t) - s_p(t) * \left\| \vec{x}^d - \vec{c}_p^d(t) \right\| / d - (D_{opt} - d)^2 \quad \text{where} \quad (16)$$

$$s_p(t) = \frac{h_p(t)}{w_p(t)} \quad \text{and} \quad \left\| \vec{x}^d - \vec{c}_p^d(t) \right\| = \sqrt{\sum_{i=1}^d (x_i^d - c_{pi}^d)^2} \quad \forall x_i^d \in \vec{x}^d, \forall c_{pi}^d \in \vec{c}_p^d(t)$$

where  $D_{opt}$  is the current optimal dimension. If compared with expression (5), now for all non-optimal dimensions

a penalty term  $(D_{opt} - d)^2$  is subtracted from the whole environment. In addition to that the peak slopes are scaled by the term  $1/d$ . The purpose of this scaling is to prevent the benchmark from favoring the lower dimensions. Otherwise a solution, whose every coordinate differs from the optimum by 1.0 would be a lot better solution in a lower dimension as the *Euclidian* distance is used.

Similar to the uni-dimensional (PSO) case, each positional component  $xx_a^d(t)$  of MD PSO particle represents a potential solution in dimension  $d$ . Only difference is that now the dimensionality of the optimal solution is not known beforehand, but it can vary within the defined range. Even a single particle can provide potential solutions in different dimensions as it makes inter-dimensional passes as a result of MD PSO process. Our dynamic multi-dimensional optimization algorithm combines multi-swarms and FGBF with MD PSO. As in the different dimensions the common coordinates of the peak locations are the same, it does not seem purposeful for two swarms to converge to a same peak in different dimensions. Therefore, the mutual repulsion between swarms is extended to affect swarms that are in different dimensions. Obviously, only the common coordinates are considered when the swarm distance is computed.

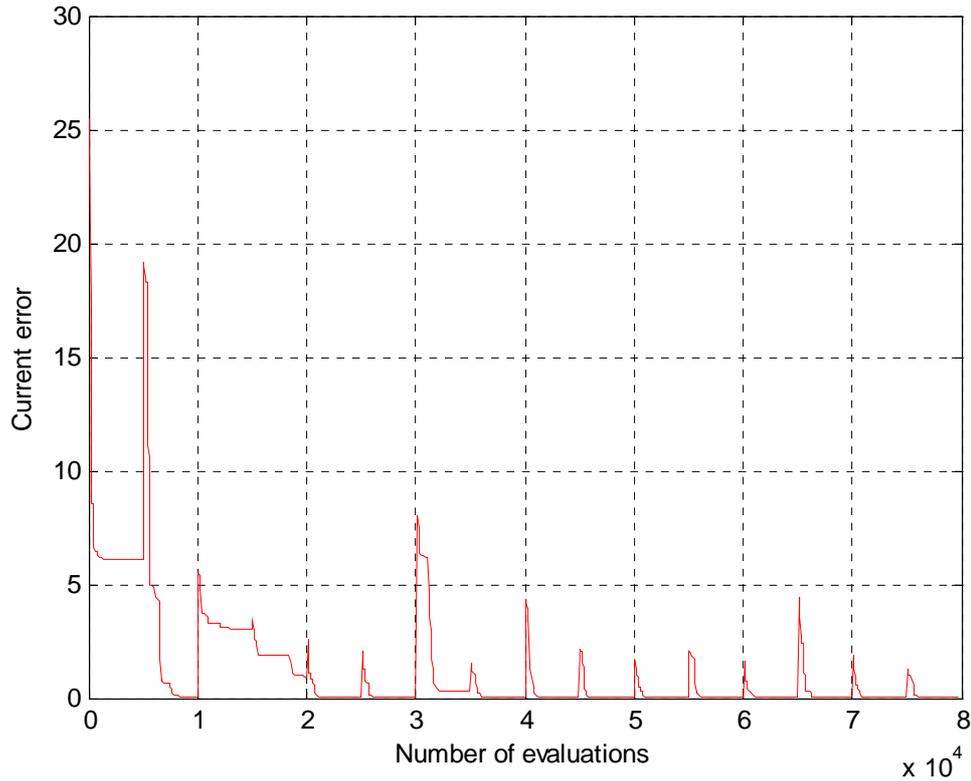
FGBF naturally exploits information gathered in other dimensions. When the *aGB* particle is created, FGBF algorithm is not limited to use dimensional components from only those particles which are in a certain dimension, but it can combine dimensional coordinates of particles in different dimensions. Note that as we still use the dimensional fitness score,  $f(a, j) = -(x_j - c_{pj})^2$ , the common coordinates of the positional components of the *aGB* particle created in different search space dimensions,  $d \in \{1, D_{max}\}$ , shall be the same. In other words, it is not necessary to create the positional components of the *aGB* particle from scratch in every search space dimension  $d \in \{1, D_{max}\}$ , instead one (new) coordinate (dimension) to the *aGB* particle is created and added. Note also that it is still possible that in some search space dimensions *aGB* beats the native *gbest* particle, while in other dimensions it does not. In multi-dimensional version also the dimension and dimensional velocity of each particle are re-initialized after an environmental change in addition to the particle velocities in each dimension.

#### IV. EXPERIMENTAL RESULTS

An extensive set of experiments was conducted over both conventional (uni-dimensional) MPB and the extended (multi-dimensional) MPB and the results will be presented in the following sub-sections.

##### A. Results on Conventional MPB

We conducted an exhaustive set of experiments over the MPB *Scenario 2* using the settings given in Section II.B. In order to investigate the effect of multi-swarm settings, we used different numbers of swarms and numbers of particles in a swarm. We applied both FGBF modes using the current and swarm peaks and to investigate how FGBF and multi-swarms individually contribute to the results, we also made experiments without using either of them.



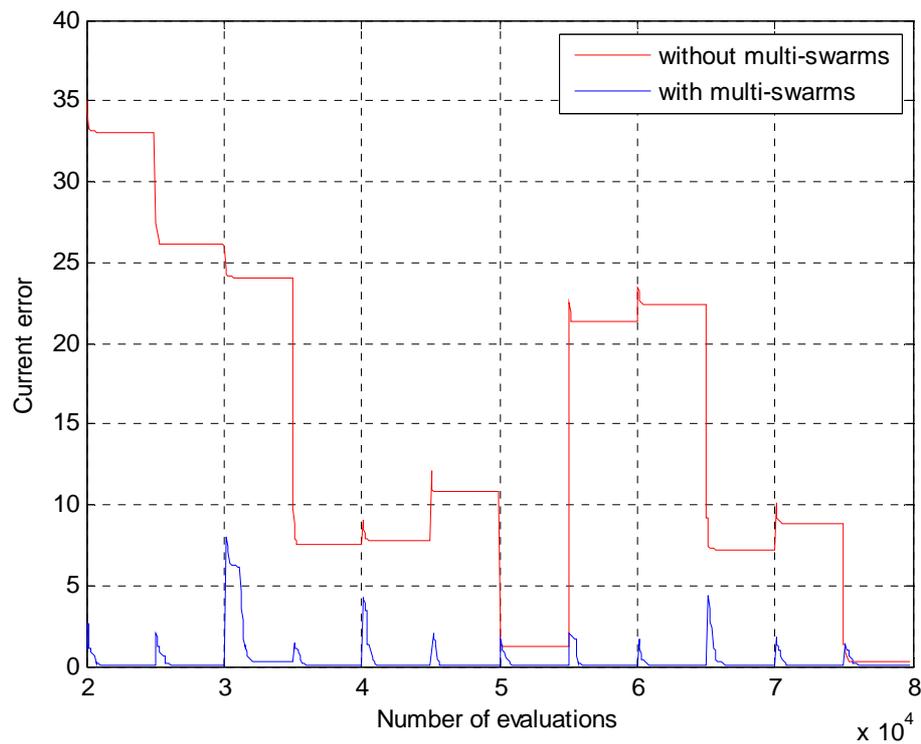
**Figure 3: Current error at the beginning of a run**

Figure 3 presents the current error plot, which shows the difference between the global maximum and the current best result during the first 80000 function evaluation, when 10 swarms each with 4 particles are used and the swarm peak mode is applied for the FGBF operation. It can be seen from the figure that as the environment changes after every 5000 evaluation, it causes results to temporarily deteriorate. However, it is clear that after environment changes the results are better than the very beginning, which shows the benefit of tracking the peaks instead of randomizing the swarm when a change occurs. The figure also reveals other typical features of algorithm behavior. First of all, after the first few environmental changes the algorithm has not yet been behaving as well as later. This is because the swarms have not yet converged to a peak. Generally, it is more difficult to initially converge to a narrow or low peak than to keep tracking a peak that becomes narrow and/or low. It can also be seen that typically the algorithm gets close to the optimal solution before the environment is changed again. In few cases, where the optimal solution is not found, the algorithm has for some reason been unable to keep a swarm tracking that peak, which is too narrow.

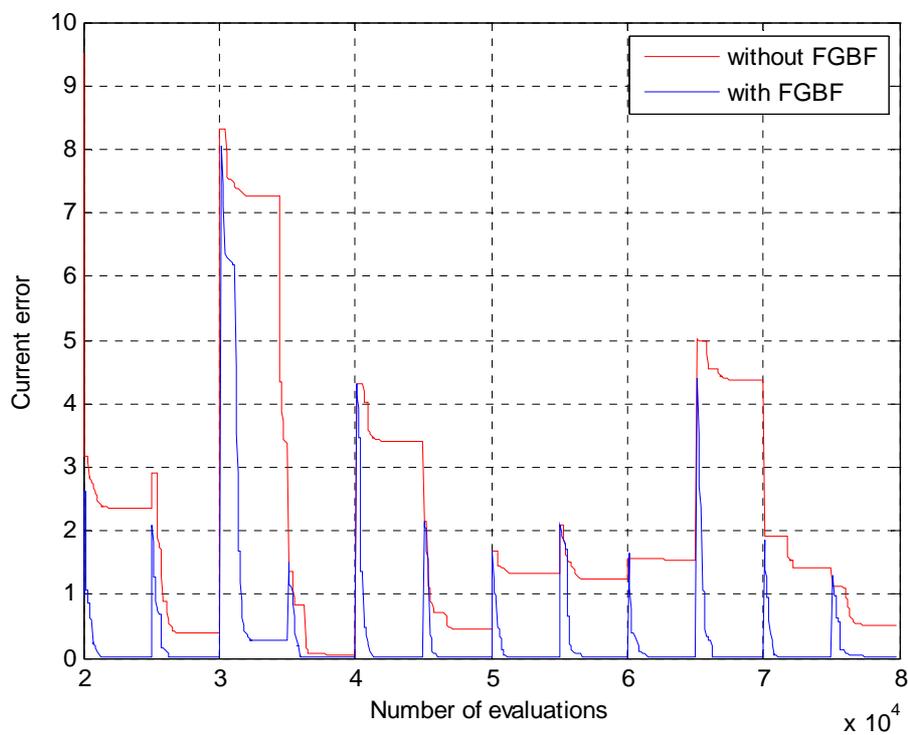
In Figure 4 and Figure 5 the contributions of multi-swarms with FGBF are demonstrated. The algorithm is run on MPB applying the same environment changes, first with both using multi-swarms and FGBF, then without multi-swarms and finally without FGBF. Same settings are used as before. Without multi-swarms the number of particles is set to 40 to keep the total number of particles unchanged.

As expected, the results without multi-swarms are significantly deteriorated due to the aforementioned reasoning. When the environment is changed, the highest point of the peak to which the swarm is converging can be found quickly, but that can provide good results only when that peak happens to be the global optimum. When multi-swarms are used, but without using the FGBF, it is clear that the algorithm can still establish some kind of follow-up of peaks as the results immediately after environment changes are only slightly worse than with FGBF.

However, if FGBF is not used, the algorithm can seldom find the global optimum. Either there is no swarm converging to the highest peak or the peak center just cannot be found fast enough.



**Figure 4: Effect of multi-swarms on results**



**Figure 5: Effect of FGBF on results**

For comparative evaluations, we selected 5 of the state-of-the-art methods, which use the same benchmark system, the MPB with the same settings. The best MPB results published so far by these competing methods are listed in Table 1.

**Table 1: Best results on MPB up to date.**

Source	Base algorithm	Offline error
Blackwell and Branke [8]	PSO	2.16±0.06
Li et. al[29]	PSO	1.93±0.06
Mendes and Mohais [33]	Differential Evolution	1.75±0.03
Blackwell and Branke [9]	PSO	1.75±0.06
Moser and Hendtlass [34]	Extremal Optimization	0.66±0.02

The overall best results have been achieved by the Extremal Optimization algorithm [34]; however, this algorithm is specially and only designed for MPB and its applicability for other practical dynamic problems is not clear. The best results by a PSO-based algorithm have been achieved by Blackwell and Branke's multi-swarm algorithm described in Section II.C.

The numerical results of the proposed technique in terms of the offline error are listed in Table 2. Each result given is the average of 50 runs, where each run consists of 500000 function evaluations. As expected the best results are achieved when 10 swarms are used. 4 particles in a swarm turned out to be the best setting. Between the two FGBF modes, better results are obtained when the swarm peak is used as the peak  $\bar{c}_p(t)$ .

**Table 2: Offline error using Scenario 2**

No. of swarms	No. of particles	Swarm peak	Current peak
10	2	1.81±0.50	2.58±0.55
10	3	1.22±0.43	1.64±0.53
10	4	<b>1.03±0.35</b>	<b>1.37±0.50</b>
10	5	1.19±0.32	1.52±0.44
10	6	1.27±0.41	1.59±0.57
10	8	1.31±0.43	1.61±0.45
10	10	1.40±0.39	1.70±0.55
8	4	1.50±0.41	1.78±0.57
9	4	1.31±0.54	1.66±0.54
11	4	1.09±0.35	1.41±0.42
12	4	1.11±0.30	1.46±0.43

### B. Results on Multi-Dimensional MPB

On the extended MPB we conducted experiments with settings similar to those used in fixed dimension except that the change frequency used is set to 15000. The search space dimension range used is  $d \in \{5, 15\}$ . Figure 6 shows how the global optimal dimension changes over time and how MD PSO is tracking these changes. Current best dimension represents the dimension, where the best solution is achieved among all swarms'  $dbest$  dimensions. 10 multi-swarms are used with 7 particles in each. FGBF is used with the swarm peak mode. It can be seen that the algorithm always finds the optimal dimension, even though the difference in peaks heights between the optimal dimension and its neighbor dimensions is quite insignificant (=1) compared to the peak heights (30-70). Figure 7 shows how the current error behaves during the first 250000 evaluations, when the same settings are used. It can be seen that the algorithm behavior is similar to the uni-dimensional case, but now the initial converging phase, when the algorithm has not been yet behaving at its best is longer. Similarly it takes longer time to regain the optimal behavior if follow-up of some peaks is lost for some reason (it is, for example, possible that higher peaks hide other lower peaks under them).

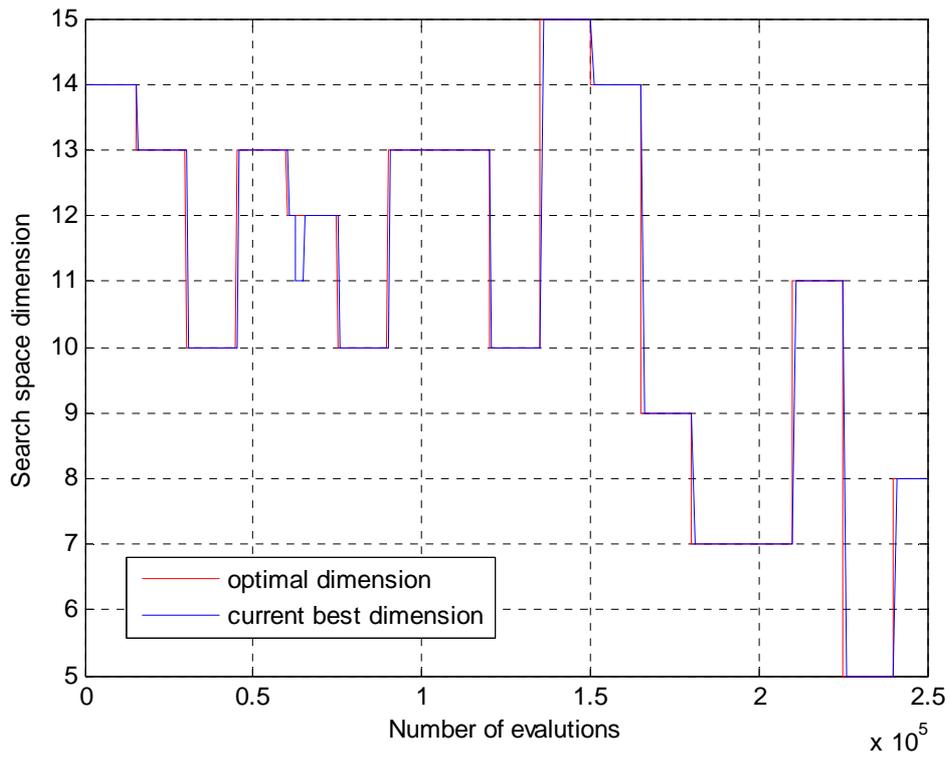


Figure 6: Optimum dimension tracking in a MD PSO run.

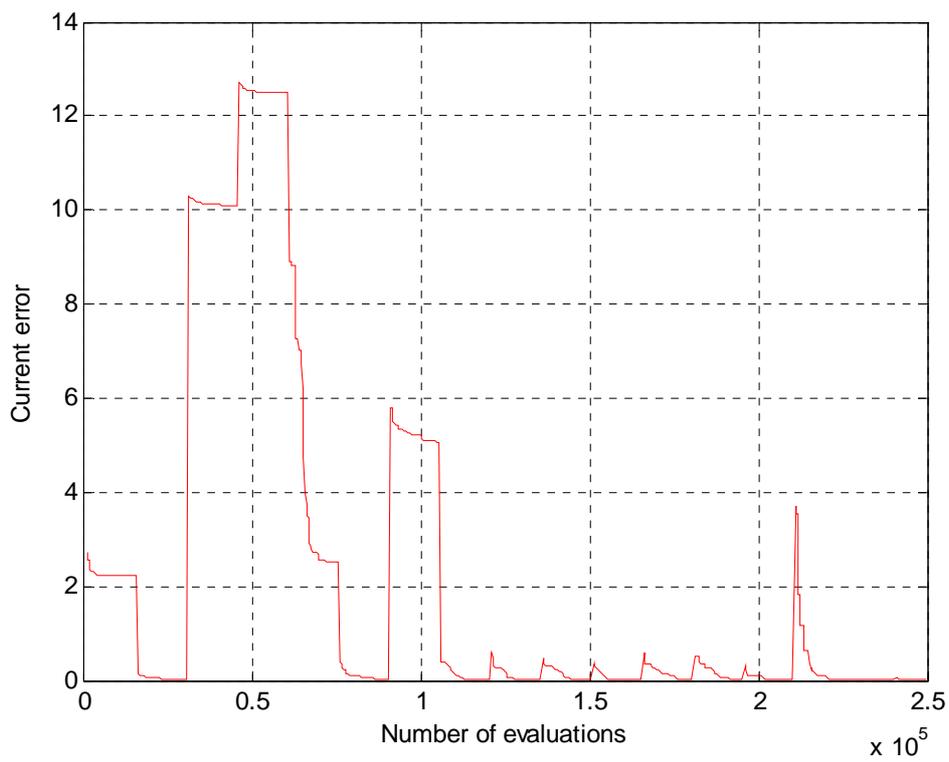
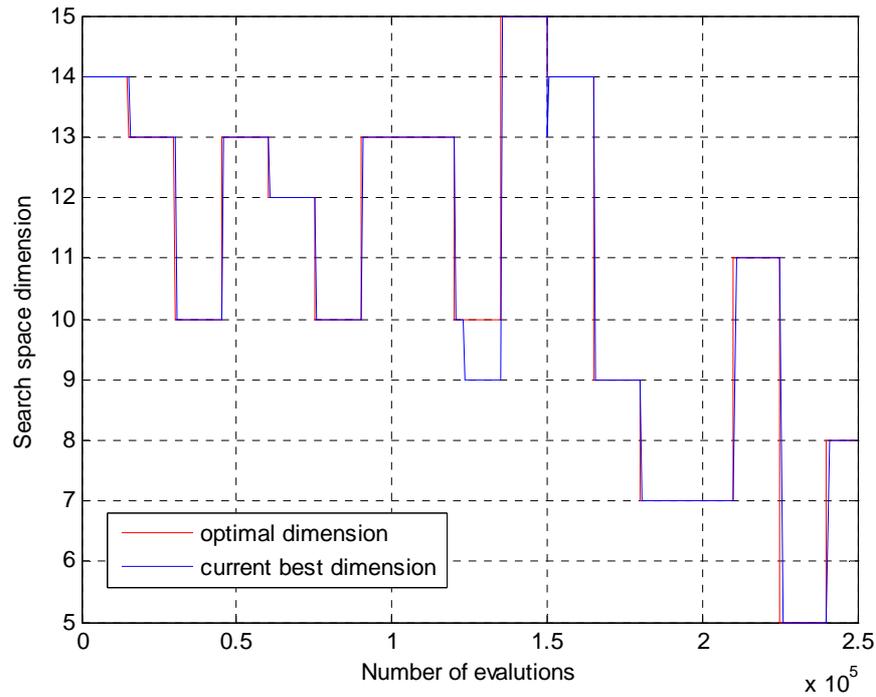
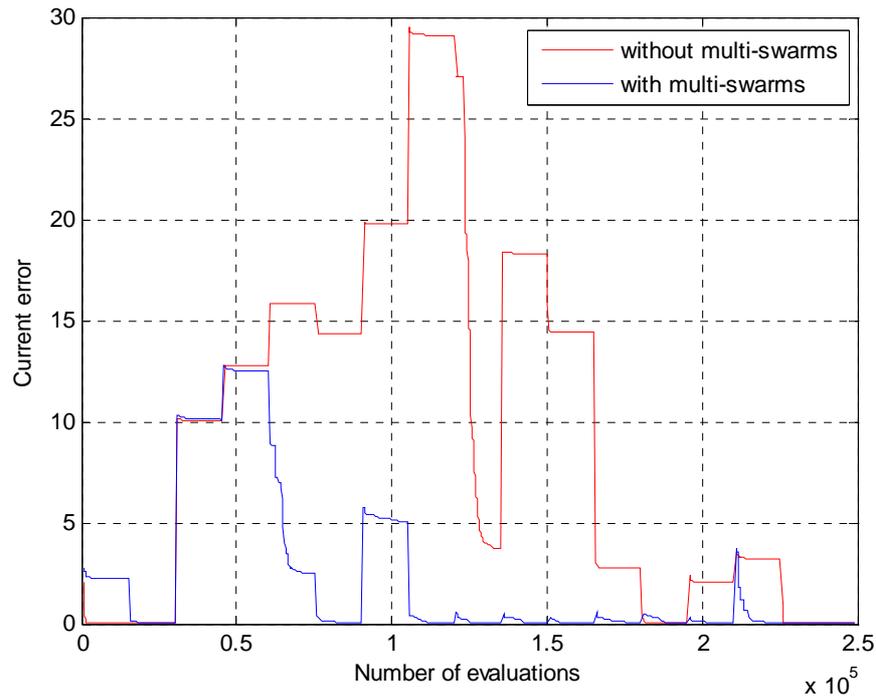


Figure 7: Current error at the beginning of a MD PSO run.

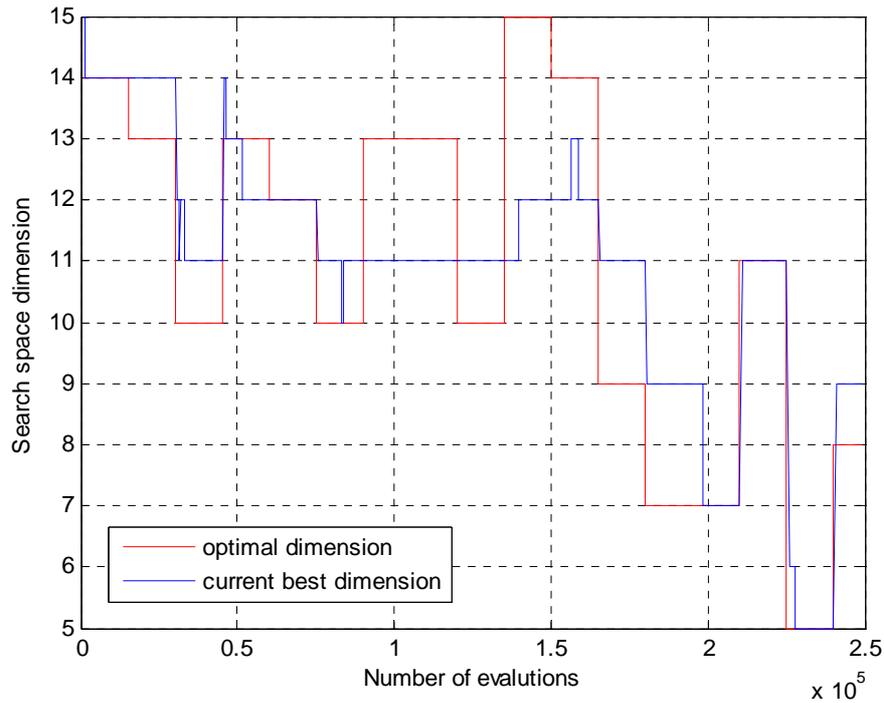


**Figure 8: Optimum dimension tracking without multi-swarms in a MD PSO run.**

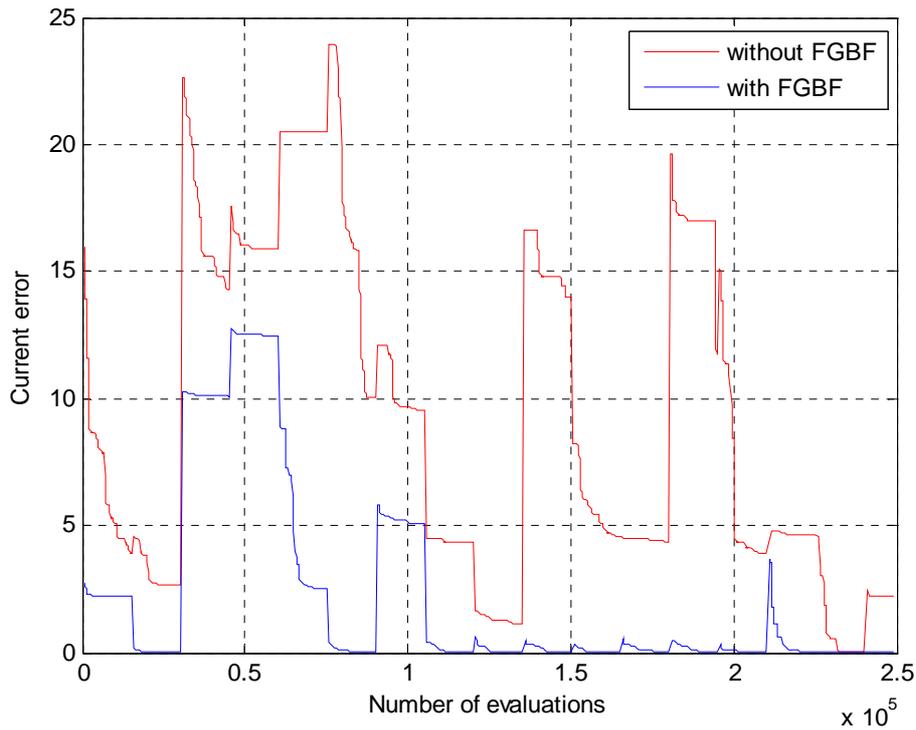


**Figure 9: Effect of multi-swarms on the performance.**

Figure 8 and Figure 9 illustrate the effect of using multi-swarms on the results. Without multi-swarms the number of particles is set to 70. Figure 8 shows that a single swarm can also find the optimal dimension easily; however, as in the uni-dimensional case, without use of multi-swarms, the optimal peak can be found only if it happens to be the peak to which the swarm is converging. This can be seen in Figure 9. During the initial converging phase of the multi-swarm algorithm results with and without multi-swarms are similar. This indicates that both algorithms initially converge to the same peak (highest) and as a result of the first few environmental changes some peaks that are not yet discovered by multi-swarms become the highest.



**Figure 10: Optimum dimension tracking *without* FGBF in a MD PSO run.**



**Figure 11: Effect of FGBF on the performance.**

Figure 10 and Figure 11 illustrate similarly the effect of FGBF on the results. In Figure 10 it can be seen that without FGBF the algorithm really has severe problems in tracking the optimal dimension. In this case, it loses the benefit of exploiting the natural diversity among the dimensional components and also it is not able to exploit information gathered in other dimensions. Therefore, even if some particles visit the optimal dimension, they cannot track the global peak *fast* enough that would hence surpasses the best results in other dimensions. Therefore, the algorithm gets trapped in some sub-optimum dimension where it happens to find the best results in

an early phase. Such reasons also cause the current error to be generally higher without FGBF, as can be seen in Figure 11.

The numerical results in terms of offline error are given in Table 3. Each result given is the average of 50 runs, where each run consists of 500000 function evaluations. As in the uni-modal case, best results are achieved when the number of swarms is equal to the number of peaks, which is 10. Interestingly when the swarm peak *mode* is used the optimal number of particles becomes 7 while with current peak mode it is still 4. Note that these results cannot be directly compared with the results on conventional MPB since the objective function of multi-dimensional MPB is somewhat different.

**Table 3: Offline error on extended MPB**

No. of swarms	No. of particles	Swarm peak	Current peak
10	4	2.01±0.98	<b>3.29±1.44</b>
10	5	1.77±0.83	3.41±1.69
10	6	1.79±0.98	3.64±1.60
10	7	<b>1.69±0.75</b>	3.71±1.74
10	8	1.84±0.97	4.21±1.83
10	10	1.96±0.94	4.20±2.03
8	7	1.79±0.91	3.72±1.86
9	7	1.83±0.84	4.30±2.15
11	7	1.75±0.91	3.52±1.40
12	7	2.03±0.97	4.01±1.97

## V. CONCLUSIONS

In this paper, we presented two PSO techniques, namely, FGBF with the multi-swarms and MD PSO for an efficient and robust optimization over the dynamic systems. Both MD PSO and FGBF have been successfully used in static optimization problems particularly as a cure to common drawbacks of the family of PSO methods such as *a priori* fixation of the search space dimension and pre-mature convergence to local optima. MD PSO efficiently addresses the former drawback by defining a new particle structure and embedding the ability of dimensional navigation into the core of the process. It basically allows particles to make inter-dimensional ‘passes’ with a dedicated PSO process whilst performing regular positional updates in every dimension they visit. Although the ability of determining the optimum dimension where the global solution exists is gained with MD PSO, its convergence performance is still limited to the same level as *bPSO*, which suffers from the lack of diversity among particles. This leads to a premature convergence to local optima especially when multi-modal problems are optimized in high dimensions. Realizing that the main problem lies in fact at the inability of using the available diversity among the dimensional components of swarm particles, the FGBF technique adapted in this paper addresses this problem by collecting the best components and fractionally creating an *aGB* particle that has the potential to be a better “guide” than the swarm’s native *gbest* particle. On MPB we do not expect to receive fractional scores with respect to the global (highest) peak, but instead we use either the peak, on which the particle is currently located (current peak) or the peak to which the swarm is converging (swarm peak). Especially swarm peak *mode* makes it possible to find and track the global highest peak quite successfully in a dynamic environment.

In order to make comparative evaluations with the current state-of-the-art, FGBF with multi-swarms is then applied over a benchmark system, the MPB. The results over the conventional MPB with common settings used (i.e. *Scenario 2*) clearly indicate the superiority of the proposed technique over other PSO-based methods. To make

the benchmark more generic for real world applications where the optimum dimension can be unknown too, MPB is extended to a multi-dimensional system in which there is a certain amount of dependency among dimensions. Note that without such dependency embedded, the benchmark would be just a bunch of independent MPBs in different dimensions and thus a distinct and independent optimization process would be sufficient for each dimension. Recall that the convergence behavior of both *bPSO* and MD PSO is same since MD PSO is only an extension of PSO for the multi-dimensional search. The performance of both methods degrades with the increasing modality and dimensionality due to the reasons mentioned earlier. When performed with FGFBF and multi-swarms, MD PSO exhibits both global convergence ability and an impressive speed gain so that their mutual performance surpasses *bPSO* by several magnitudes. The experiments conducted over the extended MPB approve that the proposed MD PSO technique with multi-swarms and FGFBF always finds and tracks the optimum dimension where the global peak resides. On both (conventional and extended) MPBs, the proposed techniques generally find and track the global peak, yet it can occasionally converge to a near-optimum peak, particularly if the height difference happens to be insignificant.

Overall, the proposed techniques fundamentally upgrade the particle structure and the swarm guidance, both of which accomplish substantial improvements in terms of speed and accuracy. Both techniques are modular and independent from each other, i.e. one can be performed without the other whilst other PSO methods/variants can also be performed conveniently with (either of) them.

#### REFERENCES

- [1] A. Abraham, S. Das and S. Roy, "Swarm Intelligence Algorithms for Data Clustering", in *Soft Computing for Knowledge Discovery and Data Mining book, Part IV*, pp. 279-313, October 25, 2007.
- [2] P.J. Angeline, "Tracking extrema in dynamic environments", *In Proc. Of the 6<sup>th</sup> Conference on Evolutionary Programming*, pp.335-345, Springer Verlag, 1997
- [3] T. Bäck, "On the behaviour of Evolutionary Algorithms in Dynamic Environments", *In Proc. of the IEEE Congress on Evolutionary Computation*, pp. 446-451, 1998.
- [4] T. Bäck and H.P. Schwefel, "An overview of evolutionary algorithms for parameter optimization", *Evolution. Comput. 1*, pp. 1–23, 1993.
- [5] T. Bäck and F. Kursawe, "Evolutionary algorithms for fuzzy logic: a brief overview", *In Fuzzy Logic and Soft Computing, World Scientific*, pp. 3–10, Singapore, 1995.
- [6] X. Chen, Y. Li, "A Modified PSO Structure Resulting in High Exploration Ability With Convergence Guaranteed", *IEEE Trans. on Systems, Man, and Cybernetics, Part B*, Vol. 37, Issue 5, pp. 1271 – 1289, Oct. 2007.
- [7] T.M. Blackwell, "Particle Swarm Optimization in Dynamic Environments", *Evolutionary Computation in Dynamic and Uncertain Environments*, Studies in Computational Intelligence, Vol. 51, pp. 29-49, Springer, 2007.
- [8] T.M. Blackwell and J. Branke, "Multi-Swarm Optimization in Dynamic Environments", *Applications of Evolutionary Computation*, vol. 3005, pp. 489-500, Springer, 2004.
- [9] T.M. Blackwell and J. Branke, "Multiswarms, Exclusion, and Anti-Convergence in Dynamic Environments", *IEEE Transactions on Evolutionary Computation*, vol. 10/4, pp. 51-58, 2004.
- [10] J. Branke, "Moving Peaks Benchmark", <http://www.aifb.uni-karlsruhe.de/~jbr/MovPeaks/>, viewed 26/06/08
- [11] Y.-P. Chen, W.-C. Peng; M.-C. Jian, "Particle Swarm Optimization With Recombination and Dynamic Linkage Discovery", in *IEEE Trans. on Systems, Man, and Cybernetics, Part B*, Vol. 37, Issue 6, pp. 1460 – 1470, Dec. 2007.
- [12] K. M. Christopher, K. D. Seppi, "The Kalman Swarm. A New Approach to Particle Motion in Swarm Optimization", *In Proc. of the Genetic and Evolutionary Computation Conf.*, GECCO, pp. 140-150, 2004.
- [13] M. Clerc, "The Swarm and the Queen: Towards a Deterministic and Adaptive Particle Swarm Optimization", *In Proc. of the IEEE Congress on Evolutionary Computation*, vol. 3, pp. 1951-1957, July 1999.
- [14] R. Eberhart, P. Simpson, and R. Dobbins, *Computational Intelligence. PC Tools*, Academic Press, Inc., Boston, MA, USA, 1996.
- [15] R. Eberhart and Y. Shi, "Tracking and Optimizing Dynamic Systems with Particle Swarms", in *Proc. of Computational Evolution Conference (CEC 2001)*, NJ, US, pp. 94-100, 2001.
- [16] A. P. Engelbrecht, *Fundamentals of Computational Swarm Intelligence*, John Wiley & Sons, 2005.
- [17] S. C. Esquivel and C. A. Coello Coello, "On the Use of Particle Swarm Optimization with Multimodal Functions", *In IEEE Trans. on Evolutionary Computation*, vol. 2, pp. 1130-1136, 2003.
- [18] U.M. Fayyad, G.P. Shapire, P. Smyth and R. Uthurusamy, *Advances in Knowledge Discovery and Data Mining*, MIT Press, Cambridge, MA, 1996.

- [19] D. Goldberg, "Genetic Algorithms in Search, Optimization and Machine Learning", Addison-Wesley, Reading, pp. 1-25. MA, 1989.
- [20] H. Higashi and H. Iba, "Particle Swarm Optimization with Gaussian Mutation", *In Proc. of the IEEE Swarm Intelligence Symposium*, pp. 72-79, 2003.
- [21] T. Ince, S. Kiranyaz, and M. Gabbouj, "A Generic and Robust System for Automated Patient-specific Classification of Electrocardiogram Signals", *IEEE Transactions on Biomedical Engineering*, in Print.
- [22] S. Janson and M. Middendorf, "A hierarchical particle swarm optimizer and its adaptive variant", *IEEE Trans. on Systems, Man, and Cybernetics, Part B*, vol. 35, Issue 6, pp. 1272 – 1282, Dec. 2005.
- [23] B. Kaewkamnerdpong, and P. J. Bentley, "Perceptive Particle Swarm Optimization: An Investigation", *In Proc. of IEEE Swarm Intelligence Symposium*, pp. 169-176, California, 8-10 June 2005.
- [24] J. Kennedy, R. Eberhart., "Particle swarm optimization", in *Proc. of IEEE Int. Conf. On Neural Networks*, vol. 4, pp. 1942–1948, Perth, Australia, 1995.
- [25] S. Kiranyaz, T. Ince, A. Yildirim and M. Gabbouj, "Fractional Particle Swarm Optimization in Multi-Dimensional Search Space", *IEEE Transactions on Systems, Man, and Cybernetics*, in Print.
- [26] J. Koza, *Genetic Programming: On the Programming of Computers by means of Natural Selection*, MIT Press, Cambridge, Massachusetts, 1992.
- [27] R. A. Krohling, L S. Coelho, "Coevolutionary Particle Swarm Optimization Using Gaussian Distribution for Solving Constrained Optimization Problems", *IEEE Trans. on Systems, Man, and Cybernetics, Part B*, Vol. 36, Issue 6, pp. 1407 – 1416, Dec. 2006.
- [28] J. J. Liang, A. K. Qin, "Comprehensive Learning Particle Swarm Optimizer for Global Optimization of Multimodal Functions", *IEEE Trans. On Evolutionary Computation*, vol. 10, no. 3, pp. 281-295, June, 2006.
- [29] X. Li, J. Branke and T. Blackwell, "Particle Swarm with Speciation and Adaptation in a Dynamic Environment", *Proc. of Genetic and Evolutionary Computation Conference*, pp. 51-58, Seattle Washington, 2006.
- [30] M. Lovberg, *Improving Particle Swarm Optimization by Hybridization of Stochastic Search Heuristics and Self-Organized Criticality*, MSc thesis, Department of Computer Science, University of Aarhus, Denmark, 2002.
- [31] M. Lovberg and T. Krink, "Extending Particle Swarm Optimisers with Self-Organized Criticality", *In Proc. of the IEEE Congress on Evolutionary Computation*, vol. 2, pp.1588-1593, 2002.
- [32] R. Mendes, J. Kennedy and J. Neves, "The fully informed particle swarm: Simpler, maybe better", *IEEE Trans. on Evolutionary Computation*, vol. 8, no 3, pp. 204–210, June 2004.
- [33] R. Mendes and A. Mohais, "DynDE: a Differential Evolution for Dynamic Optimization Problems", *IEEE Congress on Evolutionary Computation*, pp. 2808-2815, 2005.
- [34] I. Moser and T. Hendtlass, "A Simple and Efficient Multi-Component Algorithm for Solving Dynamic Function Optimisation Problems", *IEEE Congress on Evolutionary Computation*, pp. 252-259, 2007.
- [35] M. G. Omran, A. Salman, and A.P. Engelbrecht, *Particle Swarm Optimization for Pattern Recognition and Image Processing*, Springer Berlin, 2006.
- [36] B. Peng, R.G Reynolds, and J. Brewster, "Cultural Swarms", *In Proc. of the IEEE Congress on Evolutionary Computation*, vol. 3, pp. 1965-1971, 2003.
- [37] T. Peram, K. Veeramachaneni, and C.K. Mohan, "Fitness-Distance-Ratio based Particle Swarm Optimization", *In Proc. of the IEEE Swarm Intelligence Symposium*, pp. 174-181. IEEE Press, 2003.
- [38] A. C. Ratnaweera, S. K. Halgamuge, and H. C. Watson, "Particle Swarm Optimization with Self-Adaptive Acceleration Coefficients", *In Proc. of the First Int. Conf. on Fuzzy Systems and Knowledge Discovery*, pp. 264-268, 2003.
- [39] A. C. Ratnaweera, S. K. Halgamuge, and H. C. Watson, "Particle Swarm Optimiser with Time Varying Acceleration Coefficients", *In Proc. of the Int. Conf. on Soft Computing and Intelligent Systems*, pp. 240-255, 2002.
- [40] J. Riget and J. S. Vesterstrom, "A Diversity-Guided Particle Swarm Optimizer - The ARPSO", Technical report, Department of Computer Science, University of Aarhus, 2002.
- [41] M. Richards, D. Ventura, "Dynamic sociometry in particle swarm optimization," *In Proc. of the Sixth Int. Conf. on Computational Intelligence and Natural Computing*, pp. 1557-1560, North Carolina, September 2003.
- [42] Y. Shi and R.C. Eberhart, "A Modified Particle Swarm Optimizer", *In Proc. of the IEEE Congress on Evolutionary Computation*, pp. 69-73, 1998.
- [43] Y. Shi and R.C. Eberhart, "Fuzzy Adaptive Particle Swarm Optimization", *In Proc. of the IEEE Congress on Evolutionary Computation*, vol. 1, pp. 101-106, IEEE Press, 2001.
- [44] P.N. Suganthan, "Particle Swarm Optimiser with Neighborhood Operator", *In Proc. of the IEEE Congress on Evolutionary Computation*, pp. 1958-1962, IEEE Press, 1999.
- [45] F. Van den Bergh and A.P. Engelbrecht, "A New Locally Convergent Particle Swarm Optimizer", *In Proc. of the IEEE International Conference on Systems, Man, and Cybernetics*, pp. 96-101, 2002.
- [46] F. Van den Bergh and A. P. Engelbrecht, "A cooperative approach to particle swarm optimization", *IEEE Trans. on Evolutionary Computation*, vol. 3, pp. 225-239, June 2004.
- [47] F. Van den Bergh, *An Analysis of Particle Swarm Optimizers*, PhD thesis, Department of Computer Science, University of Pretoria, Pretoria, South Africa, 2002.
- [48] E.O. Wilson, *Sociobiology: The new synthesis*, Cambridge, MA: Belknap Press, 1975.
- [49] X. Xie, W. Zhang, and Z. Yang, "A Dissipative Particle Swarm Optimization", *In Proc. of the IEEE Congress on Evolutionary Computation*, vol. 2, pp. 1456-1461, 2002.
- [50] X. Xie, W. Zhang, and Z. Yang, "Adaptive Particle Swarm Optimization on Individual Level", *In Proc. of the Sixth Int. Conference on Signal Processing*, vol. 2, pp. 1215-1218, 2002.
- [51] X. Xie, W. Zhang, and Z. Yang, "Hybrid Particle Swarm Optimizer with Mass Extinction", *In Proc. of the Int. Conference on Communication, Circuits and Systems*, vol. 2, pp. 1170-1173, 2002.

- [52] K. Yasuda, A. Ide, and N. Iwasaki, "Adaptive Particle Swarm Optimization", *In Proc. of the IEEE Int. Conf. on Systems, Man, and Cybernetics*, vol. 2, pp. 1554 -1559, 2003.
- [53] W-J. Zhang and X-F. Xie, "DEPSO: Hybrid Particle Swarm with Differential Evolution Operator", *In Proc. of the IEEE International Conference on System, Man, and Cybernetics*, vol. 4, pp. 3816-3821, 2003.