III European Conference on Computational Mechanics
Solids, Structures and Coupled Problems in Engineering
C.A. Mota Soares et.al. (eds.)
Lisbon, Portugal, 5–8 June 2006

PARTICLE SWARM OPTIMIZATION: EFFICIENT GLOBALLY CONVERGENT MODIFICATIONS

Emilio F. Campana[1], **Giovanni Fasano**[1,2], **Daniele Peri**[1] **and Antonio Pinto**[1]

[1]Istituto Nazionale per Studi ed Esperienze di Architettura Navale (INSEAN)
via di Vallerano, 139 – Roma, ITALY
e-mail: {e.campana,g.fasano,d.peri,a.pinto}@insean.it

[2] Istituto di Analisi dei Sistemi ed Informatica "A. Ruberti" (IASI) - CNR
viale Manzoni, 30 – Roma, ITALY
e-mail: fasano@dis.uniroma1.it

Keywords: Global optimization, Evolutionary optimization, Particle Swarm Optimization, Global Convergence Analysis.

Abstract. *In this paper we consider the Particle Swarm Optimization (PSO) algorithm [10, 7], in the class of Evolutionary Algorithms, for the solution of global optimization problems. We analyze a couple of issues aiming at improving both the effectiveness and the efficiency of PSO. In particular, first we recognize that in accordance with the results in [5, 6], the initial points configuration required by the method, may be a crucial issue for the efficiency of PSO iteration. Therefore, a promising strategy to generate initial points is provided in the paper.*
Then, we address some very preliminary aspects of PSO global convergence towards stationary points, for some Ship Design problems. To this purpose observe that the class of Ship Design applications includes several challenging smooth problems, where expensive simulations provide information to the optimizer, and each function evaluation may require up to hours of CPU-time. In addition, the final solution provided by the optimization method is also required to be a stationary point.

1 INTRODUCTION

In this paper we use the PSO algorithm for the solution of the global optimization problem

$$\min_{x\in\mathcal{L}} f(x), \qquad \text{with} \qquad \mathcal{L} \supseteq \{x \in \mathbb{R}^n : \ \|x\|_2^2 \leq B\}, \tag{1}$$

where $f : \mathbb{R}^n \to \mathbb{R}$ is assumed *continuously differentiable*, $\mathcal{L}$ is compact and $B > 0$. PSO is a heuristic method which is used to determine an approximation of a global minimum x^* of $f(x)$ over $\mathcal{L}$, i.e. a point such that $f(x^*) \leq f(x)$ for any $x \in \mathcal{L}$. PSO may be effective on ship design problems when the function f is possibly computationally costly.
Apart from those special cases where a specific knowledge on $f(x)$ is available (e.g. $f(x)$ is convex, monotonically decreasing, etc.), the problem (1) is inherently difficult, since it requires an exhaustive exploration of $f(x)$. This usually claims for a two phase approach: on one hand, a *global* search on the set $\mathcal{L}$ attempts at identifying suitable subsets, where promising candidates of global minima are confined. Then, an efficient *local* search provides accurate approximations of each candidate, by exploring the corresponding subset of $\mathcal{L}$. Thus, the algorithms which effectively solve the global optimization problem (1) should always include the computational burden of both the *global* and the *local* phase.
Examples of problem (1) arise in several real applications, e.g. in the aerospace, automotive and naval engineering, where Navier-Stokes solvers for shape design are used.
Several iterative algorithms have been proposed in the literature to approach the solution of (1) [12]. Most of these techniques are based on the exploration of a possibly dense subset of $\mathcal{L}$; unfortunately this may require an unacceptable computational burden over ship design problems.

This drawback suggests the use of suitable iterative evolutionary heuristics to solve our problems: PSO belongs to the latter class of techniques. Its motivations arise from social behavioral interpretations of bird flocks and it is often adopted for the solution of a wide range of applications [15, 4, 8, 13, 2] (for a recent tutorial on PSO see also [3]). Some strong indications to use PSO when solving (1) are the following.

1. The solution provided by PSO is often inaccurate w.r.t. other solvers, anyway a few iterations of PSO often suffice to yield a solution with the accuracy required by the user.

2. The cost per iteration and the memory occupation are constant at each iteration.

3. The algorithm in its original form is a relatively simple heuristic, whose implementation does not require the derivatives of the objective function.

4. PSO allows a high parallelization of the computation.

By roughly speaking the rationale of the original PSO iteration is the following: a subset of P initial points (*particles*) $\{x_1^0, \ldots, x_P^0\}$ is selected in the feasible set at step $k = 0$. Then, at step $k \geq 0$ the P positions are modified according with the pair of formulae

$$\begin{aligned} d_j^k &= d_j^{k-1} + \alpha^k(x_j^{\text{best}} - x_j^k) + \beta^k(x^{\text{best}} - x_j^k), && j = 1, \ldots, P, \\ x_j^{k+1} &= x_j^k + d_j^k, && j = 1, \ldots, P. \end{aligned} \tag{2}$$

In (2) x_j^k is the current position of the j-th individual (*particle*) of a population (*swarm*), $d_j^k \in \mathbb{R}^n$ is a search direction and x_j^{k+1} is the new position at step $k + 1$. Finally α^k, β^k are suitable

random scalars and

$$x_j^{\rm best} = \underset{0\le h\le k}{\rm argmin}\,\{f(x_j^h)\}, \qquad j=1,\ldots,P,$$

$$x^{\rm best} = \underset{j=1,\ldots,P}{\rm argmin}\{f(x_j^{\rm best})\}.$$

Observe that d_j^k (the search direction from the current point x_j^k) is possibly *not* a descent direction for the objective function $f(x)$ at x_j^k. Furthermore, we remark that the direction d_j^k only depends on both the sequence $\{x_j^k\}$ and the direction d_j^{k-1} at the previous step.

We first consider in this paper the issue of determining an appropriate starting point for any particle. We can see that the latter problem may strongly affect the final results provided by PSO [5, 6]. In particular, we prove that by a direct manipulation of the formulae (2), we can explicitly figure out the role played by the initial distribution of points in PSO algorithm. We also use [5, 6] as starting references for our study. Here, the trajectory of any particle is suitably decomposed in two contributions: the first is independent of the choice of the particles starting point, while the second strongly relys on it.

On the other hand, under some additional hypotheses on $\mathcal{L}$ and $f(x)$ in (1) we introduce two globally convergent modifications of PSO iteration, based on the use of the Steepest Descent direction with constant stepsize [1]. The latter approach is still preliminary and it must be completed, since a derivative free scheme would be surely more attractive in order to preserve the property 3 at page 2. Anyway, under mild assumptions, which are largely satisfied for the class of the Ship Design problems that we consider, the schemes proved to be effective.
In order to motivate our approach, observe that the standard PSO iteration (2) in no way ensures the convergence towards a global or a local minimum. This is indeed a direct consequence of (2), where no information on first order optimality conditions is included. Consequently, the philosophy of the algorithm reduces to simply update and store one point $\hat{x}$ (corresponding to the current best computed objective function value) over the points $\{x_j^k\}$. Then, in order to improve $\hat{x}$ some attempts are considered in the literature [3]. In particular, a local minimization is performed starting from $\hat{x}$ in order to get x^*, with $f(x^*) \le f(\hat{x})$. In our proposal a globalization scheme both includes the standard PSO iteration and the local minimization step; thus, we asymptotically yield the stationary point x^* (possibly a minimum point) of the objective function, without generating the 'intermediate' point $\hat{x}$.

An extensive numerical experience on our real problems will appear in a forthcoming paper, where a derivative-free modified PSO is also included, to guarantee the global convergence towards stationary points.

In Sections 2 and 3 we describe a generalized PSO iteration, by means of a dynamic linear system, whose properties are partially used in Section 4, in order to investigate promising starting points for the particles. Section 1 describes two globally convergent schemes of PSO towards stationary points, and Section 6 contains some conclusions.

2 GENERALIZING THE PSO MODEL

According with the symbols broadly used in the literature, let us consider the following (generalized) iteration of PSO algorithm ($k = 0, 1, ...$):

$$\begin{aligned} v_j^{k+1} &= \chi\left[w^k v_j^k + c_j r_j (p_j^k - x_j^k) + c_g r_g (p_g^k - x_j^k)\right], \\ x_j^{k+1} &= x_j^k + v_j^{k+1}, \end{aligned} \tag{3}$$

where $j = 1, ..., P$ indicates the j-th particle, P is finite, v_j^k and x_j^k are the so called *speed* and *position* of particle j at step k, p_j^k and p_g^k respectively satisfy

$$\begin{aligned} p_j^k &= \text{argmin}_{l \le k} \left\{ f(x_j^l) \right\}, \\ p_g^k &= \text{argmin}_{l \le k,\ j=1,...P} \left\{ f(x_j^l) \right\}, \end{aligned} \tag{4}$$

and $\chi, w^k, c_j, r_j, c_g, r_g$ are suitable bounded coefficients. *We assume that* r_j and r_g *are uniformly distributed random parameters* with $0 \le r_j \le 1$ and $0 \le r_g \le 1$. Assuming that in (3) the speed v_j^{k+1} depends on all the terms $p_h^k - x_j^k$, $h = 1, \ldots, P$, and not only on vectors $p_j^k - x_j^k$, $p_g^k - x_j^k$, we can further generalize the PSO. Indeed we obtain the new iteration ($k = 0, 1, ...$)

$$\begin{aligned} v_j^{k+1} &= \chi_j^k \left[w_j^k v_j^k + \sum_{h=1}^{P} c_{h,j}^k r_{h,j}^k (p_h^k - x_j^k) \right], \\ x_j^{k+1} &= x_j^k + v_j^{k+1}, \end{aligned} \tag{5}$$

where $c_{h,j}^k$ and $r_{h,j}^k$ may in general depend on the step (k), the current particle (j) and the other particles (h). Let us focus on the j-th particle and omit the subscript j in the recurrence (5). Moreover, let for $k \ge 0$ in (5) $\chi_j^k = \chi$, $c_{h,j}^k = c_h$, $r_{h,j}^k = r_h$ and $w_j^k = w$, which is a common hypothesis in the PSO literature. With the latter position the iteration (5) is equivalent to the *dynamic system*

$$\begin{aligned} X(k+1) = & \begin{pmatrix} \chi w I & -\sum_{h=1}^{P} \chi c_h r_h I \\ \chi w I & \left(1 - \sum_{h=1}^{P} \chi c_h r_h\right) I \end{pmatrix} X(k) \ + \\ & \begin{pmatrix} \sum_{h=1}^{P} \chi c_h r_h p_h^k \\ \sum_{h=1}^{P} \chi c_h r_h p_h^k \end{pmatrix}, \end{aligned} \tag{6}$$

where

$$X(k) = \begin{pmatrix} v^k \\ \\ x^k \end{pmatrix} \in \mathbb{R}^{2n}, \qquad k \ge 0. \tag{7}$$

Then, (6) is used to obtain some partial indications on the assessment of the starting positions of the particles in PSO.

The sequence $\{X(k)\}$ identifies a trajectory in the state space $\mathbb{R}^{2n}$, and since (6) is a linear and stationary system, we may consider the *free response* $X_{\mathcal{L}}(k)$ and the *forced response* $X_{\mathcal{F}}(k)$ of the trajectory $\{X(k)\}$. Then, considering (6) we explicitly obtain at step $k \ge 0$ [14]

$$X(k) = X_{\mathcal{L}}(k) + X_{\mathcal{F}}(k), \tag{8}$$

where

$$X_{\mathcal{L}}(k) = \Phi(k) X(0), \qquad X_{\mathcal{F}}(k) \ = \ \sum_{\tau=0}^{k-1} H(k-\tau) U(\tau), \tag{9}$$

$$\Phi(k) = \begin{pmatrix} \chi w I & -\sum_{h=1}^{P} \chi c_h r_h I \\ \chi w I & \left(1 - \sum_{h=1}^{P} \chi c_h r_h\right) I \end{pmatrix}^k, \tag{10}$$

$$H(k-\tau) = \begin{pmatrix} \chi w I & -\sum_{h=1}^{P} \chi c_h r_h I \\ \chi w I & \left(1 - \sum_{h=1}^{P} \chi c_h r_h\right) I \end{pmatrix}^{k-\tau-1}, \tag{11}$$

$$U(\tau) = \begin{pmatrix} \sum_{h=1}^{P} \chi c_h r_h p_h^\tau \\ \sum_{h=1}^{P} \chi c_h r_h p_h^\tau \end{pmatrix}. \tag{12}$$

Observe that $X_{\mathcal{L}}(k)$ in (9) *does not depend* on the vector p_h^k, but uniquely on the initial point $X(0)$. On the contrary, $X_{\mathcal{F}}(k)$ in (9) depends on the vector p_h^k and is independent of $X(0)$. The latter practical observation allows us to compute separately the two terms. In order to carry out our conclusions, in the next sections we focus on $X_{\mathcal{L}}(k)$.

3 THE FREE RESPONSE $X_{\mathcal{L}}(k)$

From the theory of linear systems we know by definition that asymptotically

$$\lim_{k\to\infty} X(k) = \lim_{k\to\infty} X_{\mathcal{F}}(k),$$

i.e. the free response $X_{\mathcal{L}}(k)$ is effective only for finite values of k. In this section we focus on the properties of $X_{\mathcal{L}}(k)$, so that it can be used to properly define the starting point of each particle.
Let us consider the following position in (10) and (11)

$$a = \chi w, \qquad \omega = \sum_{h=1}^{P} \chi c_h r_h, \tag{13}$$

and consider the following assumptions (see also [5]).

Assumption 3.1 *We assume in (13) $a \neq 0$ and $\omega > 0$. Furthermore we assume in (13) $\omega \neq (1-\sqrt{a})^2$ and $\omega \neq (1+\sqrt{a})^2$ for any $a > 0$, so that the $2n$ eigenvectors of matrix $\Phi(1)$ are linearly independent.*

Then, it can be proved [5] that the matrices Λ and V exist such that $V^{-1}\Phi(1)V = \Lambda$, with

$$\Lambda = \begin{pmatrix} \lambda_1 I & \\ & \lambda_2 I \end{pmatrix} \in \mathbb{R}^{2n\times 2n},$$

$$V = (v_1 \cdots v_{2n}) = \begin{pmatrix} I & I \\ \frac{a-\lambda_1}{\omega} I & \frac{a-\lambda_2}{\omega} I \end{pmatrix},$$

$$V^{-1} = \frac{1}{\lambda_1 - \lambda_2} \begin{pmatrix} (a-\lambda_2)I & -\omega I \\ -(a-\lambda_1)I & \omega I \end{pmatrix}$$

and

$$\begin{aligned} \lambda_1 &= \frac{1-\omega+a-\left[(1-\omega+a)^2-4a\right]^{1/2}}{2} \\ \lambda_2 &= \frac{1-\omega+a+\left[(1-\omega+a)^2-4a\right]^{1/2}}{2}. \end{aligned} \tag{14}$$

The previous result yields (see [5])

$$\begin{aligned} X_{\mathcal{L}}(k) &= V\Lambda^k V^{-1} X(0) \\ &= \frac{1}{\lambda_1-\lambda_2} \sum_{i=1}^{n} \Big[\lambda_1^k X(0)^T \left((a-\lambda_2)e_i - \omega e_{n+i}\right) v_i - \\ &\quad \lambda_2^k X(0)^T \left((a-\lambda_1)e_i - \omega e_{n+i}\right) v_{n+i} \Big] \\ &= \frac{1}{\lambda_1-\lambda_2} \sum_{i=1}^{n} \Big[\lambda_1^k \left((a-\lambda_2)X(0)_i - \omega X(0)_{n+i}\right) v_i - \\ &\quad \lambda_2^k \left((a-\lambda_1)X(0)_i - \omega X(0)_{n+i}\right) v_{n+i} \Big], \end{aligned}$$

where $e_i \in \mathbb{R}^{2n}$ is the unit vector with 1 in position i and $X(0)_i$ is the i-th entry of $X(0)$, i.e.

$$X(0) = \begin{pmatrix} X(0)_1 \\ \vdots \\ X(0)_{2n} \end{pmatrix}.$$

Furthermore, for $i = 1, \ldots, n$

$$\begin{cases} v_i = e_i + \dfrac{a-\lambda_1}{\omega} e_{n+i} \\ v_{n+i} = e_i + \dfrac{a-\lambda_2}{\omega} e_{n+i}, \end{cases}$$

then

$$\begin{aligned} X_{\mathcal{L}}(k) = \sum_{i=1}^{n} & \left[\gamma_1(k)X(0)_i e_i - \gamma_2(k)X(0)_{n+i}e_i + \right. \\ & \left. \gamma_3(k)X(0)_i e_{n+i} - \gamma_4(k)X(0)_{n+i}e_{n+i}\right], \end{aligned} \tag{15}$$

where coefficients $\gamma_i(k)$, $i = 1, \ldots, 4$ are given in Table 1. Observe that the first value of each coefficient $\gamma_1(k), \ldots, \gamma_4(k)$ is worth for real eigenvalues λ_1 and λ_2, while the second value holds in case λ_1 and λ_2 are conjugate (i.e. $\lambda_1 = \rho e^{-j\theta}$, $\lambda_2 = \rho e^{j\theta}$, with $a = \rho^2$). In the end, from (15) we obtain

$$\begin{aligned} X_{\mathcal{L}}(k) = \sum_{i=1}^{n} & \left[\gamma_1(k)X(0)_i - \gamma_2(k)X(0)_{n+i}\right] e_i + \\ & \left[\gamma_3(k)X(0)_i - \gamma_4(k)X(0)_{n+i}\right] e_{n+i}. \end{aligned} \tag{16}$$

$$\gamma_1(k) = \begin{cases} \dfrac{\lambda_1^k(a-\lambda_2) - \lambda_2^k(a-\lambda_1)}{\lambda_1 - \lambda_2} & \lambda_1, \lambda_2 \ \text{real} \\ \rho^{k+1}\dfrac{\sin k\theta}{\sin\theta} - \rho^k \dfrac{\sin(k-1)\theta}{\sin\theta} & \lambda_1, \lambda_2 \ \text{complex}, \end{cases}$$

$$\gamma_2(k) = \begin{cases} \dfrac{\omega(\lambda_1^k - \lambda_2^k)}{\lambda_1 - \lambda_2} & \lambda_1, \lambda_2 \ \text{real} \\ \omega\rho^{k-1}\dfrac{\sin k\theta}{\sin\theta} & \lambda_1, \lambda_2 \ \text{complex}, \end{cases}$$

$$\gamma_3(k) = \begin{cases} \dfrac{(\lambda_1^k - \lambda_2^k)}{\lambda_1 - \lambda_2}\dfrac{(a-\lambda_1)(a-\lambda_2)}{\omega} & \lambda_1, \lambda_2 \ \text{real} \\ \rho^k \dfrac{\sin k\theta}{\sin\theta}\left(\dfrac{\rho^2 - 2\rho\cos\theta + 1}{\omega}\right) & \lambda_1, \lambda_2 \ \text{complex}, \end{cases}$$

$$\gamma_4(k) = \begin{cases} \dfrac{\lambda_1^k(a-\lambda_1) - \lambda_2^k(a-\lambda_2)}{\lambda_1 - \lambda_2} & \lambda_1, \lambda_2 \ \text{real} \\ \rho^{k+1}\dfrac{\sin k\theta}{\sin\theta} - \rho^k \dfrac{\sin(k+1)\theta}{\sin\theta} & \lambda_1, \lambda_2 \ \text{complex}. \end{cases}$$

Table 1: The coefficients $\gamma_i(k)$, $i = 1, \ldots, 4$ in (15).

We remark that by simply imposing $X(0)_i = X(0)_{n+i} = 0$ in (16), the free response of a particle has zero entries on the i-th and $(n+i)$-th axis. Thus, according with the free response in (16), each particle's trajectory has nonzero projection on any subspace of $\mathbb{R}^{2n}$, provided that the initial point $X(0)$ is suitably chosen. From the latter consideration and (16), a suitable choice of the starting point $X(0)$ of any particle may guarantee an improved exploration of the state space. In the next section we give numerical evidence that the latter issue plays a key role within global optimization frameworks.

4 THE STARTING POINT OF EACH PARTICLE

On the guideline of the previous section, here we give some theoretical indications about the choice of the initial point $X(0)$ of each particle, which is a crucial issue in PSO. In particular, let us consider (1) and let us indicate with $X(k)^{(j)}$ the trajectory of the j-th particle. From (8) $X(k)^{(j)}$ depends linearly on the contribution of the free response $X_{\mathcal{L}}(k)^{(j)}$ in (16), i.e. it depends on the corresponding initial point $X(0)^{(j)}$.

In this section and the next one we assume that up to n^2 particles compose the swarm. We want to assess the initial points $X(0)^{(j)}$, $j = 1, \ldots, n^2$, of the particles, in such a way that the state space $\mathbb{R}^{2n}$ is explored as widely as possible by the trajectories $\{X(k)^{(j)}\}$. This is a general issue for global optimization algorithms, where the users often resort to randomly generated sequences of initial points. Here we propose in particular a technique to generate up to n^2 starting points, such that for any fixed index k, the sequence $\{X_{\mathcal{L}}(k)^{(j)}\}$ in (16) is *scattered* in the state space $\mathbb{R}^{2n}$. Our numerical experience with PSO confirms that the latter result may be often fruitfully used.

To the latter purpose (see also [5]) we consider in (13), for any particle $j = 1, \ldots, n$, the parameters a_j, w_j in place of a, w. Accordingly, the coefficients $\gamma_1(k), \ldots, \gamma_4(k)$ in (16) are given for the j-th particle by $\gamma_1(k)^{(j)}, \ldots, \gamma_4(k)^{(j)}$.

Consider the vector $t_j \in \mathbb{R}^n$, $j = 1, \ldots, n$, such that

$$t_j = \frac{\sqrt{n}}{n} \sum_{i=1}^{n} e_i - \frac{\sqrt{n}}{2} e_j. \tag{17}$$

Then, we have immediately $t_j^T t_h = 0$, for $1 \leq j \neq h \leq n$. We also recall (see [5, 6]) that under the Assumption 3.1 the vectors (17) may be used to generate at step k the vectors $X_{\mathcal{L}}(k)^{(1)}, \ldots, X_{\mathcal{L}}(k)^{(n)}$, which form an orthogonal basis in the state subspace of the positions. We aim at extending the latter result which only relys on the orthogonality of the vectors in (17).

Let $Q \in \mathbb{R}^{n \times n}$ and suppose Q is orthogonal, i.e. $Q^T Q = I_n$. In particular, the columns $\{q_1, \ldots, q_n\}$ of Q represent an orthonormal basis of $\mathbb{R}^n$. Suppose we *rotate* each column-vector q_k of Q by a specific angle, into the vector s_k, $k = 1, \ldots, n$. Then, an orthogonal matrix $\bar{Q} \in \mathbb{R}^{n \times n}$ exists such that the vectors $s_k = \bar{Q} q_k$, $1 \leq k \leq n$, are still an orthonormal basis and $s_1^T q_1 = \cdots = s_n^T q_n$.
We introduce a measure d of the distance between the two bases $\{q_k\}$ and $\{s_k\}$, according with

$$d = \|Q - \bar{Q} Q\|_F, \tag{18}$$

where $\|\cdot\|_F$ indicates the *Frobenius norm* of a matrix, i.e. $\|A\|_F = [\mathrm{tr}(A^T A)]^{1/2}$. Then, we can consider the orthogonal matrix Q^* which satisfies

$$Q^* = \mathrm{argmax}_{\bar{Q}} \|Q - \bar{Q} Q\|_F, \qquad \text{s.t. } \bar{Q}^T \bar{Q} = I_n. \tag{19}$$

In other words, according with the definition of the distance d in (18), we say that the orthogonal matrix $Q^* Q$ is at *maximum distance* from the matrix Q. We recall that by definition we have for any $A, B \in \mathbb{R}^{n \times n}$

$$\|A - B\|_F^2 = \mathrm{tr}\left[(A - B)^T (A - B)\right] = \mathrm{tr}(A^T A) + \mathrm{tr}(B^T B) - 2\mathrm{tr}(A^T B) \tag{20}$$

thus, $Q^* = -I_n$ satisfies (19) since

$$\mathrm{tr}(Q^T Q) = \mathrm{tr}\left[(Q^* Q)^T (Q^* Q)\right] = n,$$

so that

$$\|Q - Q^* Q\|_F^2 = 2n - 2\mathrm{tr}(Q^T Q^* Q) = 4n. \tag{21}$$

From (18)-(19), the latter result yields the intuitive consideration that the bases $\{q_1, \ldots, q_n\}$ and $\{-q_1, \ldots, -q_n\}$ are maximally distant in the sense of formula (18). Therefore, since the columns of Q are orthogonal vectors, we could use the sets $\{q_1, \ldots, q_n\}$ and $\{-q_1, \ldots, -q_n\}$ in place of the vectors in (17). To sum up, if only $2n$ particles are chosen for the swarm, then from (17) and considering (7) we suggest the following set of initial particles' position ($1 \leq j \leq n$)

$$\begin{aligned} X(0)^{(j)} &= \begin{pmatrix} 0 \\ t_j \end{pmatrix}, & t_j &= 2(-1)^j \left[\frac{\sqrt{n}}{n} \sum_{i=1}^{n} e_i - \frac{\sqrt{n}}{2} e_j \right], \\ X(0)^{(n+j)} &= \begin{pmatrix} 0 \\ t_{n+j} \end{pmatrix}, & t_{n+j} &= 2(-1)^{j+1} \left[\frac{\sqrt{n}}{n} \sum_{i=1}^{n} e_i - \frac{\sqrt{n}}{2} e_j \right]. \end{aligned} \tag{22}$$

(we recall that by a short calculation the matrices $(t_1 \cdots t_n)$ and $(t_{n+1} \cdots t_{2n})$ are orthogonal).

In Table 2 we have a preliminary numerical experience, where the PSO method is applied for the solution of 35 test problems of size n ($2 \leq n \leq 30$), from the literature. We considered three scenarios with the same implementation of PSO (i.e. the coefficients χ, w^k, c_j, r_j, c_g, r_g in the scenarios were set at the same standard values), and up to 1000 function evaluations were allowed in each scenario (regardless of the size n). The column $f(x^*)_{n\_rand}$ reports the best solution obtained by a *random* choose of the initial position of n particles. The column $f(x^*)_{n\_orth}$ [$f(x^*)_{2n\_orth}$] reports the best solution obtained by choosing the initial position of n [$2n$] particles as in (22). Our proposal evidently *outperforms the random choice*; similar results were obtained when the PSO parameters were modified in a wide range.

4.1 The extension to n^2 particles

Now, let us extend the latter result and let us focus on providing the initial position of up to n^2 particles (which covers almost all the situations of practical interest). On this purpose, recalling that the columns of the matrix Q in (18) are required to be an orthonormal basis of $\mathbb{R}^n$, in (18)-(21) we can set without loss of generality $Q = I_n$. We want to determine $h > 2$ suitable orthonormal bases of $\mathbb{R}^n$, by rotating the canonical basis $\{e_1, \ldots, e_n\}$. Equivalently we aim at determining in accordance with (18)-(19) the largest positive constant $M > 0$ and a set of $h > 2$ orthogonal matrices $\{Q_1, \ldots, Q_h\}$ such that

$$M \; = \; \max \left\{ \min_{1 \leq i \neq k \leq h} \|Q_i - Q_k\|_F \right\}. \tag{23}$$

In addition, we also require that the set of column-vectors of any matrices Q_i and Q_j, $1 \leq i \neq j \leq n$, contains *exactly* $2n$ *distinct vectors*. Observe that the condition (23) substantially summarizes the fact that the orthogonal matrices $\{Q_1, \ldots, Q_h\}$ are mutually maximally distant, i.e. the h bases obtained by rotating the canonical basis $\{e_1, \ldots, e_n\}$ are maximally 'scattered' in $\mathbb{R}^n$.

The problem (23) may be quite difficult and its solution could take far from our concern. Indeed (see [9]), in (20) the quantity $\|A - B\|_F$ may be evidently unbounded when A and B range in $\mathbb{R}^{n \times n}$. Anyway, even imposing the condition $(A-B)^T(A-B) = I$, a solution which maximizes $\|A - B\|_F^2$ is $A - B = P$ (where P is any permutation matrix). Thus, no strong analytical conditions are available in order to determine $Q_1, \ldots, Q_h$. Therefore, we renounce to fully solve (23) and instead we claim for the following weaker result. We want to find out $1 \leq h \leq n$ orthogonal matrices $Q_1, \ldots, Q_h$ such that if $Q_i = (q_1^{(i)}, \ldots, q_n^{(i)})$, $1 \leq i \leq h$,

1) $q_k^{(i)} \neq q_p^{(j)}$, $1 \leq i \neq j \leq h$ and $1 \leq k \neq p \leq n$,

2) $\|Q_i - Q_j\|_F \geq m$, for any $1 \leq i \neq j \leq h$ and for some $m > 0$.

Observe that of course from (21) we have $m \leq 4n$; however, the following result holds.

Proposition 4.1 *Consider the set of h symmetric orthogonal Houseolder transformations $Q_1, \ldots, Q_h$*

$$\begin{aligned} Q_i &= I_n - 2\frac{w_i w_i^T}{\|w_i\|^2}, && 1 \leq i \leq h, \\ w_i &= \frac{\sqrt{n}}{n} \sum_{l=1}^{n} e_l - \frac{\sqrt{n}}{2} e_i, && 1 \leq i \leq h, \end{aligned} \tag{24}$$

with $n > 2$. Then, $Q_1, \ldots, Q_h$ satisfy both the requirements 1) and 2).

Proof.
As regards 1) observe that by a straightforward calculation $w_i^T w_j = 0$, $1 \le i \ne j \le h$; furthermore, let $w_i = (w_{i1} \cdots w_{in})^T$, then we have for any $1 \le i, j \le h$ and $k \ne p$,

$$q_k^{(i)} = e_k - 2\frac{w_{ik}}{\|w_i\|^2} w_i, \qquad 1 \le k \le n,$$

$$q_p^{(j)} = e_p - 2\frac{w_{jp}}{\|w_j\|^2} w_j, \qquad 1 \le p \le n.$$

Thus, by contradiction $q_k^{(i)} = q_p^{(j)}$ if and only if

$$\frac{w_{ik}}{\|w_i\|^2} w_i = \frac{e_k - e_p}{2} + \frac{w_{jp}}{\|w_j\|^2} w_j,$$

so that the orthogonality of vectors w_i and w_j yields

$$\begin{cases} w_{ik} = \dfrac{w_{ik} - w_{ip}}{2}, \\ \\ w_{jp} = -\dfrac{w_{jk} - w_{jp}}{2}, \end{cases}$$

i.e. $w_{ik} = -w_{ip}$ and $w_{jk} = -w_{jp}$. However, from (24) we have for w_{ik} and w_{ip} (similarly for w_{jk} and w_{jp})

$$w_{ik} = \begin{cases} \sqrt{n}/n - \sqrt{n}/2 & \text{if } i = k \\ \sqrt{n}/n & \text{if } i \ne k, \end{cases}$$

$$w_{ip} = \begin{cases} \sqrt{n}/n - \sqrt{n}/2 & \text{if } i = p \\ \sqrt{n}/n & \text{if } i \ne p. \end{cases}$$

Furthermore, since $k \ne p$ and $n > 2$ we always have $w_{ik} \ne -w_{ip}$ and $w_{jk} \ne -w_{jp}$. Thus, 1) holds.
As regards 2) observe that since $w_i^T w_j = 0$, $1 \le i \ne j \le h$, we have

$$\begin{aligned} \|Q_i - Q_j\|_F^2 &= 4\left\| \frac{w_i w_i^T}{\|w_i\|^2} - \frac{w_j w_j^T}{\|w_j\|^2} \right\|_F^2 = 4\left[\mathrm{tr}\left(\frac{w_i w_i^T}{\|w_i\|^2} \right) \right. \\ &\left. + \mathrm{tr}\left(\frac{w_j w_j^T}{\|w_j\|^2} \right) - 2\mathrm{tr}\left(\frac{w_i w_i^T w_j w_j^T}{\|w_i\|^2 \|w_j\|^2} \right) \right] = 8, \end{aligned} \tag{25}$$

i.e. 2) is satisfied with $m = 2\sqrt{2}$. $\mathbf{Q.E.D.}$

We also note that recalling the observations related to formula (21), for $n = 2$ the choice (24) is an optimal choice. To summarize, the formulae (24) allow us to generate up to n orthonormal bases of $\mathbb{R}^n$, which satisfy the properties 1) and 2). Consequently, since the column vectors of each matrix may be used as initial points of particles in PSO frameworks, we have a rule to suitably supply n^2 initial points. In particular, recalling that $h \le n$, by the choice of vectors

$w_1, \ldots, w_h$ in (24) and from (7) we adopt the following formulae to *explicitly* address the n^2 initial points:

$$X(0)^{(1)} = \begin{pmatrix} 0 \\ \\ e_1 - 2\dfrac{w_{11}w_1}{\|w_1\|^2} \end{pmatrix}, \qquad \cdots \qquad X(0)^{(n)} = \begin{pmatrix} 0 \\ \\ e_n - 2\dfrac{w_{1n}w_1}{\|w_1\|^2} \end{pmatrix},$$
$$X(0)^{(n+1)} = \begin{pmatrix} 0 \\ \\ e_1 - 2\dfrac{w_{n1}w_n}{\|w_n\|^2} \end{pmatrix}, \qquad \cdots \qquad X(0)^{(n^2)} = \begin{pmatrix} 0 \\ \\ e_n - 2\dfrac{w_{nn}w_n}{\|w_n\|^2} \end{pmatrix}. \tag{26}$$

Finally, the choice (24) for the vectors $\{w_1, \ldots, w_h\}$ might be suitably modified, in order to further improve the bound (25). In addition, if the feasible set $\mathcal{L}$ in (1) does not contain the points (26), then the latter could be accordingly scaled.

5 GLOBALLY CONVERGENT SCHEMES FOR PSO

In this section we address the issue of studying globally convergent modifications of PSO. The latter subject has often been disregarded in the PSO literature. Indeed, on one hand several practitioners are not concerned with sensitivity and convergence analysis. On the other hand according with Section 1, PSO is a heuristics particularly effective in the first iterations, while its progress is less significant when approaching (possibly) the solution. Nevertheless, several real applications have been described in the literature (see also [3]), where evolutionary algorithms are also required to ensure the convergence to stationary points.
Consider the PSO iteration (5) and let $|A|$ indicate the cardinality of the general set A. Let us introduce the following relevant sets for our discussion:

$$\begin{aligned} &\mathcal{K} = \{1, 2, \ldots\}, \\ &\mathcal{K}^M = \{n \in \mathbf{N} : \ n = 1 + iM, \ i \in \mathbf{N} \cup \{0\}, \ M > 0 \text{ integer}\}, \\ &\mathcal{K}^{M+} = \left\{n \in \mathcal{K} : \ n - 1 \in \mathcal{K}^M\right\}, \\ &\mathcal{K}^M \cup \mathcal{K}^{M+} \subseteq \mathcal{K}_j \subseteq \mathcal{K}, \qquad j = 1, \ldots, P. \end{aligned} \tag{27}$$

Observe that by loosely speaking the subsets $\mathcal{K}^M$ and $\mathcal{K}^{M+}$ contain "samples" of the elements in $\mathcal{K}$, and possibly we can set $\mathcal{K}_j \equiv \mathcal{K}^M \cup \mathcal{K}^{M+}$, $j = 1, \ldots, P$. Furthermore, in order to slightly generalize the position (4) we compute the quantities

$$\begin{aligned} &q_j^k = \text{argmin}_{l \le k, \ l \in \mathcal{K}_j}\{f(x_j^l)\}, \qquad j = 1, \ldots, P, \ k \ge 1, \\ &p_g^k = \text{argmin}_{1 \le j \le P}\{f(q_j^k)\}, \qquad k \ge 1. \end{aligned} \tag{28}$$

We remark that by definition $\{p_g^k\} \subseteq \{q_j^k\} \subseteq \{x_j^k\}$ and by setting $\mathcal{K}_j \equiv \mathcal{K}$ for any j, the pairs (q_j^k, p_g^k) in (28) coincide with the pairs (p_j^k, p_g^k) in (4).
We introduce the globally convergent modified PSO algorithm GPSO in Table 3. We can preliminarily observe that after a brief initialization in Step 1, in the Step k we first check for

a stopping criterion, which tests the stationarity of the current solution. If not satisfied, the new *speed* vector v_j^{k+1} is computed. Then, either if the current iteration index k belongs to the subset $\mathcal{K}^M$ or the progress is too slow (i.e. $f(p_g^{k-1}) - f(p_g^k) \leq \sigma |f(p_g^{k-1})|$), the 'best' current particle's position p_g^k is perturbed along the steepest descent direction $-\nabla f(p_g^k)$. Finally, the vectors $\{q_j^k\}$ are updated. Observe that from (27)-(28) and the instructions at Step k, we cannot have $p_g^{k+M} \equiv p_g^k$, for any $k \in \mathcal{K}^M$ and $j = 1, \ldots, P$ such that $\nabla f(x_j^k) \neq 0$. Indeed, in the latter case we surely have by Step k and definition (28) $f(p_g^{k+M}) \leq f[p_g^k - \eta_k \nabla f(p_g^k)] < f(p_g^k)$. For the algorithm in Table 3 the following general convergence result holds.

Theorem 5.1 *Let $\mathcal{L}_1 = \{x \in \mathbb{R}^n : f(x) \leq \max_{1 \leq j \leq P}\{f(x_j^1)\}\}$ be compact and $f \in C^1(\mathcal{L}_1)$. Consider (27) and the sequences $\{p_g^k\} \subseteq \{q_j^k\} \subseteq \{x_j^k\}$, $j = 1, \ldots, P$, $k \geq 1$, where p_g^k satisfies (28), and the sequences $\{q_j^k\}$, $\{x_j^k\}$, $j = 1, \ldots, P$, are generated by algorithm* GPSO. *Suppose there exists a positive constant L such that*

$$\|\nabla f(y) - \nabla f(z)\| \leq L\|y - z\|, \tag{29}$$

for any $y, z \in \mathcal{L}_1$, and in algorithm GPSO *$\eta_k = \eta < 2/L$, for any $k \in \mathcal{K}^M$. Then, assuming $M \geq 1$ in (27), either an integer $\nu \geq 1$ exists such that $\nabla f(p_g^\nu) = 0$ or the algorithm* GPSO *generates the infinite subsequence $\{p_g^k\}_{\mathcal{K}^M}$ with*

$$\lim_{k \to \infty,\ k \in \mathcal{K}^M} \left\|\nabla f(p_g^k)\right\| = 0. \tag{30}$$

Proof.

Suppose a finite index $\nu \geq 1$ does not exist such that $\nabla f\left(p_g^\nu\right) = 0$. Then, if $k \in \mathcal{K}^M$ we have also $k + M \in \mathcal{K}^M$, so that by (28) and Step k of algorithm GPSO

$$\begin{aligned} f\left(p_g^{k+M}\right) &= \min_{1 \leq j \leq P,\ l \in \mathcal{K}_j,\ l \leq k+M} \left\{f(q_j^l)\right\} \\ &= \min\left\{f\left(p_g^k\right)\ ,\ f\left(q_j^l\right)_{1 \leq j \leq P,\ l \in \mathcal{K}_j,\ k < l \leq k+M}\right\} \\ &\leq \min\left\{f\left(p_g^k\right)\ ,\ f\left[p_g^k - \eta \nabla f\left(p_g^k\right)\right]\right\}, \qquad k, k+M \in \mathcal{K}^M. \end{aligned} \tag{31}$$

Then, by setting $d^k = -\eta \nabla f(p_g^k)$ and considering the mean value theorem along with (29), we obtain for any $k, k + M \in \mathcal{K}^M$

$$\begin{aligned} f\left(p_g^{k+M}\right) &\leq f\left(p_g^k + d^k\right) \\ &= f\left(p_g^k\right) + \int_0^1 \nabla f\left(p_g^k + t d^k\right)^T d^k dt \\ &\leq f\left(p_g^k\right) + \int_0^1 \left\|\nabla f\left(p_g^k + t d^k\right) - \nabla f\left(p_g^k\right)\right\| \|d^k\| dt + \nabla f\left(p_g^k\right)^T d^k \\ &\leq f\left(p_g^k\right) + \int_0^1 t L \|d^k\|^2 dt + \nabla f\left(p_g^k\right)^T d^k \\ &= f\left(p_g^k\right) + \frac{L}{2}\|d^k\|^2 + \nabla f\left(p_g^k\right)^T d^k \\ &= f\left(p_g^k\right) + \eta\left(\frac{\eta L}{2} - 1\right)\left\|\nabla f\left(p_g^k\right)\right\|^2, \qquad k, k+M \in \mathcal{K}^M. \end{aligned} \tag{32}$$

Then, $\left\{f\left(p_g^k\right)\right\}_{\mathcal{K}^M}$ is a decreasing subsequence provided that $\eta L/2 < 1$; moreover, since $\mathcal{L}_1$ is compact, $\left\{f\left(p_g^k\right)\right\}_{\mathcal{K}^M}$ is convergent. Now, observe that by rearranging relation (32) we obtain

$$f\left(p_g^k\right) - f\left(p_g^{k+M}\right) \geq \eta\left(1 - \frac{\eta L}{2}\right)\left\|\nabla f\left(p_g^k\right)\right\|^2, \qquad k, k+M \in \mathcal{K}^M. \tag{33}$$

Thus, applying the limit $k \to \infty$, $k \in \mathcal{K}^M$, the left hand side of inequality (33) approaches zero and proves relation (30). $\mathbf{Q.E.D.}$

Observe that relation (29) often holds in ship design applications, and at least an estimation $\tilde{L} > L$ of the Lipschitz constant is available. However the explicit knowledge of at least an estimation of the parameter L, may be a too restrictive hypothesis in few cases. On this guideline, in the following result we slightly modify the statement of Theorem 5.1, so that the global convergence of algorithm GPSO is also achieved *without explicitly knowing* the constant L.

Theorem 5.2 *Let $\mathcal{L}_1 = \{x \in \mathbb{R}^n : f(x) \leq \max_{1\leq j\leq P}\{f(x_j^1)\}\}$ be compact and $f \in C^1(\mathcal{L}_1)$. Consider (27) and the sequences $\{q_j^k\} \subseteq \{x_j^k\} \subseteq \{x_j^k\}$, $j = 1, \ldots, P$, $k \geq 1$, where p_g^k satisfies (28). Let the sequences $\{q_j^k\}$, $\{x_j^k\}$, $j = 1, \ldots, P$, be generated by algorithm* GPSO. *Suppose there exists a positive constant L such that*

$$\|\nabla f(y) - \nabla f(z)\| \leq L\|y - z\|, \tag{34}$$

for any $y, z \in \mathcal{L}_1$. Let in the algorithm GPSO

$$\begin{aligned} &\textit{a)} \quad \lim_{k\to\infty,\ k\in\mathcal{K}^M} \eta_k = 0, \\ &\textit{b)} \quad \sum_{k\in\mathcal{K}^M} \eta_k = +\infty. \end{aligned} \tag{35}$$

Then, assuming $M \geq 1$ in (27), either an integer $\nu \geq 1$ exists such that $\nabla f(p_g^\nu) = 0$ or the algorithm GPSO *generates the infinite subsequence $\{p_g^k\}_{\mathcal{K}^M}$ with*

$$\lim_{k\to\infty,\ k\in\mathcal{K}^M} \left\|\nabla f(p_g^k)\right\| = 0. \tag{36}$$

Proof.
First observe that for any $y, z \in \mathcal{L}_1$ we have from the hypothesis (see also [1])

$$f(y + z) \leq f(y) + \nabla f(y)^T z + \frac{L}{2}\|z\|^2.$$

Thus, reasoning as in (31) of Theorem 5.1 we obtain for any $k, k + M \in \mathcal{K}^M$

$$\begin{aligned} f\left(p_g^{k+M}\right) &\leq \min\left\{f\left(p_g^k\right),\ f\left[p_g^k - \eta_k \nabla f\left(p_g^k\right)\right]\right\} \leq f\left[p_g^k - \eta_k \nabla f\left(p_g^k\right)\right] \\ &\leq f\left(p_g^k\right) - \eta_k \left\|\nabla f\left(p_g^k\right)\right\|^2 + \frac{\eta_k^2 L}{2}\left\|\nabla f\left(p_g^k\right)\right\|^2 \\ &= f\left(p_g^k\right) - \eta_k\left(1 - \frac{L}{2}\eta_k\right)\left\|\nabla f\left(p_g^k\right)\right\|^2. \end{aligned} \tag{37}$$

Now, from (35) $\lim_{k\to\infty, k\in\mathcal{K}^M} \eta_k = 0$; thus, the constant $0 < c \leq (1 - L\eta_k/2)$, $k \geq \bar{k}$, exists such that relation (37) eventually yields

$$f\left(p_g^{k+M}\right) \leq f\left(p_g^k\right) - \eta_k c\left\|\nabla f\left(p_g^k\right)\right\|^2, \qquad k \geq \bar{k},\ k \in \mathcal{K}^M. \tag{38}$$

Then, as long as p_g^k is not a stationary point, $\{f(p_g^k)\}_{\mathcal{K}^M}$ is monotonically decreasing and converging on the compact set $\mathcal{L}_1$. Furthermore, from (37) and the latter considerations

$$\begin{aligned} \sum_{k\geq\bar{k},\ k\in\mathcal{K}^M} \eta_k c\left\|\nabla f\left(p_g^k\right)\right\|^2 &\leq \sum_{k\geq\bar{k},\ k\in\mathcal{K}^M}\left[f\left(p_g^k\right) - f\left(p_g^{k+M}\right)\right] \\ &= f\left(p_g^{\bar{k}}\right) - \lim_{k\to\infty,\ k\in\mathcal{K}^M} f\left(p_g^k\right) < +\infty. \end{aligned} \tag{39}$$

Now, if we assume in (39) the condition $\left\|\nabla f(p_g^k)\right\| \geq \xi > 0$, $k \geq \bar{k}$, $k \in \mathcal{K}_M$, then (35) and (39) yield the contradiction

$$c\xi^2 \left(\sum_{k\geq\bar{k},\ k\in\mathcal{K}^M} \eta_k \right) < +\infty.$$

Therefore condition (36) must hold, which completes the theorem. $\mathbf{Q.E.D.}$

Remark 5.1 *Observe that in Theorem 5.2 the explicit knowledge of the Lipschitz constant L is unnecessary. Indeed, the hypothesis (35) may be straightforwardly satisfied by setting for instance $0 < \bar{\eta} < \infty$, $0 < \alpha \leq 1$ and*

$$\eta_k = \frac{\bar{\eta}}{k^\alpha} \tag{40}$$

in the algorithm GPSO.

Remark 5.2 *From (37) and (38) it is not difficult to prove that replacing the hypothesis (35) with the slightly heavier assumption*

$$\begin{aligned} &a) \qquad \lim_{k\to\infty,\ k\in\mathcal{K}^M} \eta_k = 0, \\ &b) \qquad \sum_{k\in\mathcal{K}^M} \eta_k = +\infty, \ \sum_{k\in\mathcal{K}^M} \eta_k^2 < +\infty, \end{aligned} \tag{41}$$

we obtain in place of (39)

$$\begin{aligned} \sum_{k\in\mathcal{K}^M} \eta_k \left\|\nabla f\left(p_g^k\right)\right\|^2 - \sum_{k\in\mathcal{K}^M} \frac{\eta_k^2 L}{2} \left\|\nabla f\left(p_g^k\right)\right\|^2 \ &\leq \ \sum_{k\in\mathcal{K}^M} \left[f\left(p_g^k\right) - f\left(p_g^{k+M}\right)\right] \\ = \ f\left(p_g^1\right) - \lim_{k\to\infty,\ k\in\mathcal{K}^M} f\left(p_g^k\right) \ &< \ +\infty. \end{aligned} \tag{42}$$

The hypothesis (41) may be easily satisfied by setting $0 < \bar{\eta} < \infty$, $1/2 < \alpha \leq 1$ and

$$\eta_k = \frac{\bar{\eta}}{k^\alpha} \tag{43}$$

in the algorithm GPSO*, i.e. loosely speaking (35) and (41) are substantially equivalent.*

6 CONCLUSIONS

We have described a couple of relevant issues for improving the efficiency and the effectiveness of the algorithm PSO, which is a widely used heuristics for the approximate solution of the global optimization problem (1), in ship design frameworks. First, we have given numerical evidence that a proper choice of the initial position of the particles may be crucial for the effectiveness of the method. Then, under reasonable assumptions two globally convergent modifications of PSO were proposed, in order to guarantee that at least a subsequence of iterates converges to a stationary point.

On the guideline of the previous section, a derivative-free globally convergent modification of PSO is being studied (see also [11]); extended numerical results on ship design problems will be included in a forthcoming paper.

Acknowledgements

The authors E.F.Campana and D.Peri would like to thank the “Programma Ricerca INSEAN 2005-2007” for its support. G.Fasano thanks the research program “VISIR” and A.Pinto was supported under the ONR research program NICOP “GLOBAL OPTIMIZATION” No. 000140510617.

REFERENCES

[1] D.P.Bertsekas, “Nonlinear programming - II edition,” Athena Scientific, Belmont MA, 1995.

[2] C.A.Baker, L.T.Watson, B.Grossman, R.T.Haftka, W.H.Mason, “Parallel global aircraft configuration design space exploration,” Proceedings of the 8th AIAA/NASA/ISSMO Symposium on Multidisciplinary Analysis and Optimization, Paper No. 2000–4763–CP, Long Beach, California (USA), 2000.

[3] M.Clerc, “Particle Swarm Optimization ,” ISTE Publishing Company, Lavoisier 2006.

[4] E.F.Campana, G.Liuzzi, S.Lucidi, D.Peri, V.Piccialli, “New Global Optimization Methods for Ship Design Problems,” Technical Report INSEAN 2005-01.

[5] E.F.Campana, G.Fasano, A.Pinto “Particle Swarm Optimization: dynamic system analysis for parameter selection in global optimization frameworks” Technical Report INSEAN 2005-023, Italy.

[6] E.F.Campana, G.Fasano, A.Pinto “Dynamic system analysis and initial particles position in Particle Swarm Optimization,” IEEE Swarm Intelligence Symposium (SIS2006), May 12-14, 2006, Indianapolis, Indiana, USA.

[7] M.Clerc, J.Kennedy, “The Particle Swarm - Explosion, Stability, and Convergence in a Multidimensional Complex Space,” IEEE Transactions on Evolutionary Computation, Vol. 6, No. 1, 2002.

[8] P.C.Fourie, A.A.Groenwold, “Particle Swarms in Size and Shape Optimization,” Proceedings of the International Workshop on Multidisciplinary Design Optimization, Pretoria, South Africa, August 7-10, 2000, pp. 97-106.

[9] G.H.Golub, C.F.Van Loan, “Matrix Computations - III edition,” John Hopkins University Press, USA.

[10] J.Kennedy, R.C.Eberhart, “Particle swarm optimization,” Proceedings of the 1995 IEEE International Conference on Neural Networks (Perth, Australia), IEEE Service Center, Piscataway, NJ, IV: 1942-1948, 1995.

[11] T.G.Kolda, R.M.Lewis, V.Torczon, “Optimization by direct search: new perspectives on some classical and modern methods ,” SIAM Review, Vol. 45, No. 3, pp. 385-482, 2003.

[12] J.D.Pinter, “Global Optimization in Action. Continuous and Lipschitz Optimization: Algorithms, Implementations and Applications,” Kluwer Academic Publishers, The Netherlands, 1996.

[13] A.Pinto, D.Peri, E.F.Campana, "Global Optimization Algorithms in Naval Hydrodynamics" Ship Technology Research, 51, 3, pp 123-133, 2004.

[14] P.E.Sarachik, "Principles of linear systems," Cambridge University Press, 1997.

[15] G.Venter, J.Sobieszczanski-Sobieski, "Multidisciplinary optimization of a transport aircraft wing using particle swarm optimization" Structural and Multidisciplinary Optimization, vol. 26, no. 1-2, pp. 121-131, 2004.

Problem	n	$f(x^*)_{n\_rand}$	$f(x^*)_{n\_orth}$	$f(x^*)_{2n\_orth}$
Six humps camel back	2	-0.8118	-1.0316	-1.031627
Treccani	2	0.5983	1.2125E-06	1.3723035E-04
Quartic	2	4.8516	-0.1526	-0.3523849
Schubert	2	-37.5954	-49.2691	-123.5767
Schubert pen.1	2	-30.6048	-23.9748	-123.1987
Schubert pen.2	2	-18.8594	-23.2320	-122.8207
Shekel5	4	-0.4085	-3.36416	-5.054569
Shekel7	4	-0.5047	-3.39414	-5.087214
Shekel10	4	-0.9192	-1.30725	-5.128263
Exponential	2	-1.1805E-02	-0.9999	-1.000000
	4	-4.7240E-07	-0.9999	-0.9999982
Cosine mixture	2	-4.6458E-02	-0.2000	-0.1999997
Cosine mixture	4	9.1806E-02	-0.1042	-0.3999915
hartman3	3	-0.9769	-3.6570	-3.655713
hartman6	6	-3.2795	-3.0208	-2.910185
5^n loc. minima	2	9.3625	6.6128E-03	2.8020752E-06
	5	5.1373E-02	3.6934-02	2.9364234E-02
	10	1.1255	0.1103	8.8159777E-02
	20	1.0001	0.2525	8.8145070E-02
	30	0.7282	0.1453	1.5093682E-02
10^n loc. minima	2	45.9406	2.81867E-02	1.378443
	5	5.005	1.2651	1.370563
	10	34.3466	2.3363	3.425857
	20	20.5893	4.3714	2.6454281E-05
	30	20.4258	3.3331	0.1432712
15^n loc. minima	2	0.8580	2.23199E-02	0.5647247
	5	3.9033E-02	0.2709	1.671701
	10	1.5038	0.8420	2.7893741E-08
	20	5.4427	1.5152	8.3999912E-06
	30	2.5683	2.6402	1.1098337E-05
Griewank	2	0.9505	8.1694E-07	5.7031489E-06
	5	1.96844E-03	3.7250E-05	8.5866317E-02
	10	0.5813	3.8645E-05	3.338129
	20	0.6602	1.8684E-05	3.021994
	30	0.5974	8.5806E-05	4.1590993E-06

Table 2: Numerical comparison among three different choices of the starting points' position, in a PSO framework. Our proposal (II and III column of results) outperforms the standard random choice of initial particles' position.

Global Particle Swarm Optimization (GPSO)

Data: $P \geq 1$ integer, $\sigma > 0$, $\epsilon > 0$.

Step 1: Set $x_j^1 \in \mathbb{R}^n$ and $q_j^1 = x_j^1$, $j = 1, \ldots, P$; compute p_g^1, set $k = 1$.

Step k**:** If $k \in \mathcal{K}^M$ and $\|\nabla f(p_g^k)\| < \epsilon$ STOP, else set

$$v_j^{k+1} = \chi_j^k \left[w_j^k v_j^k + \sum_{h=1}^{P} c_{h,j}^k r_{h,j}^k (q_h^k - x_j^k) \right], \qquad j = 1, \ldots, P.$$

If $\left[k \in \mathcal{K}^M \text{ or } \left(k > 1 \text{ and } f(p_g^{k-1}) - f(p_g^k) \leq \sigma |f(p_g^{k-1})| \right) \right]$ then

Do $j = 1, \ldots, P$

set $\eta_k > 0$,

$x_j^{k+1} = x_j^k - \eta_k \nabla f(x_j^k)$, if $x_j^k = p_g^k$,

$x_j^{k+1} = x_j^k + v_j^{k+1}$, if $x_j^k \neq p_g^k$,

Endo.

Else $x_j^{k+1} = x_j^k + v_j^{k+1}$, $j = 1, \ldots, P$.

Set $k = k + 1$.

Do $j = 1, \ldots, P$

If $k \in \mathcal{K}_j$ then $q_j^k = \text{argmin}_{i \in \mathcal{K}_j,\, i \leq k} \{f(x_j^i)\}$

Endo.

Calculate p_g^k and goto **Step** k.

Table 3: GPSO: a modified Globally convergent PSO algorithm.