Theoretical Derivation, Analysis and Empirical Evaluation of a Simpler Particle Swarm Optimiser

Riccardo Poli Dan Bratton Tim Blackwell Jim Kennedy

***Abstract*— In this paper we derive a simpler form of Particle Swarm Optimiser (PSO) which still retains the key properties of the original model. We do so by progressively altering the original model via mathematical transformations which have clear and understandable probabilistic effects. Then we study this new PSO mathematically and compare it to the original PSO. In particular we analyse the stability and the sampling distribution for the new PSO during stagnation. Finally, we test it on 14 standard benchmark problems and with two different communication topologies, with very encouraging results.**

I. Introduction

Despite its apparent simplicity, the Particle Swarm Optimiser (PSO) presents formidable challenges to those interested in understanding swarm intelligence. More and more sophisticated analysis tools have become available in the last few years to study and understand the classical PSO (e.g., [7], [11], [10], [1], [4], [2], [5], [9]). However, to date neither a fully comprehensive mathematical model nor a universally accepted qualitative explanation of the PSO are available. In this context then one may wonder whether progress can also be made by simplifying the PSO itself to its bare bones, as was done for example in [6]. The challenge then becomes finding simpler forms of PSO which still retain the key properties of the original model.

In this paper we derive a new such PSO (Section II). We do so by progressively altering the original model via mathematical transformations which have clear and understandable probabilistic effects. Then, in Section III, we study this new PSO mathematically and compare it to the original PSO during stagnation (i.e., while particles are looking for an improved personal or neighbourhood best). In particular we analyse the stability of the new PSO. This allows us to determine good parameter settings for it. In Section IV we study the sampling distribution for the new PSO. We then test it on 14 standard benchmark problems and with two different communication topologies (Section V), with very encouraging results. We draw some conclusions in Section VI.

II. Derivation of the new PSO

Let us consider the 1–D case of the standard update of a traditional PSO with constriction [3]:

$$v_{t+1} = \chi\left(v_t + \frac{\phi}{2}u_1(p - x_t) + \frac{\phi}{2}u_2(g - x_t)\right) \quad (1)$$

Riccardo Poli (rpoli@essex.ac.uk) is with the Department of Computer Science, University of Essex, UK. Tim Blackwell (tim.blackwell@gold.ac.uk) and Dan Bratton (d.bratton@gold.ac.uk) are with the Department of Computing, Goldsmiths College, London, UK. Jim Kennedy (Kennedy.Jim@bls.gov) is with the US Bureau of Labor Statistics, Washington, USA.

where u_1 and u_2 are two random variables uniformly distributed in the interval $[0, 1]$, i.e., $u_1, u_2 \sim U(0, 1)$, p is the particle's best position and g is the neighbourhood best. With a little algebra we can rewrite this equation as

$$v_{t+1} = \chi\left(v_t + \frac{\phi}{2}(u_1 + u_2)(\hat{x} - x_t)\right) \quad (2)$$

where

$$\hat{x} = \frac{u_1 p + u_2 g}{u_1 + u_2}. \quad (3)$$

Note that $\hat{x}$ is a random variable representing a random point along the line connecting p and g. Note, however, that this variable is not uniformly distributed along the segment. Instead it has a "tent" like profile with a peak at $(p + g)/2$, as shown in Figure 1 where we report the density function of $\hat{x}$ for the case $p = 0$ and $g = 1$ estimated by sampling it 1,000,000 times and computing histograms.

How important is the precise shape of this distribution? A simple way to answer this is to explore what happens if we replace such a distribution with a different one. Given its simplicity, a rectangular distribution $U(p, g)$ is particularly appealing in this sense. This distribution presents the same mean as the original, but it has an increased variance. For example, the variance of the distribution shown in Figure 1 is 0.056997, while the variance of the corresponding uniform distribution, $U(0, 1)$, is $\frac{1}{12} \approx 0.083333$. So, by picking $\hat{x} \sim U(p, g)$ we are effectively adding more noise to the particle's movement, so we should expect to find that the search is more explorative than with the original model. Also, note that while $\hat{x}$ is a function of u_1 and u_2 in the original formulation, in the new formulation, $\hat{x} \sim U(p, g)$, we have that $\hat{x}$, u_1

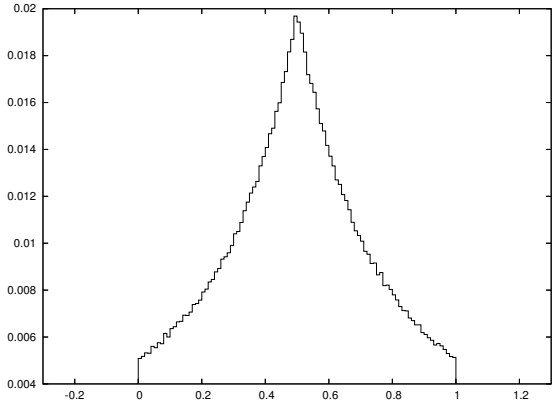

Fig. 1. Tent-like distribution of $\hat{x}$ (see Equation (3)) for $p = 0$ and $g = 1$.

and u_2 are uncorrelated. Finally, note that we can express $U(p,g) = p + (g-p)U(0,1)$.

Could we simplify things even further? An idea would be to replace the factor (u_1+u_2) in Equation (2) with something else. Note that (u_1+u_2) has a triangular distribution (the convolution of two identical uniform distributions) with mean $E[u_1+u_2]=1$. A radical way to replace (u_1+u_2) would then be to use its mean instead.

These modifications lead to the following new PSO:

$$v_{t+1} = \chi\left(v_t + \frac{\phi}{2}((g-p)u + p - x_t)\right) \quad (4)$$

where $u \sim U(0,1)$. In the following we will refer to this as **SPSO**, which stands for "Simpler PSO".

The main benefit of SPSO is that it is simpler to understand and analyse. Because this PSO was derived by mean preserving transformations of the original, one might expect that Clerc and Kennedy's formula for choosing χ and ϕ, namely

$$\chi = \frac{2}{2-\phi-\sqrt{\phi^2-4\phi}} \quad (5)$$

(see [3]), might work just fine for SPSO. We study Equation (4) and verify this conjecture in the next section.

III. Stability analysis

To analyse our new PSO we will use a recently developed technique which allows one to determine the moments of the sampling distribution of a PSO and its stability [8].

Let us start by rewriting Equation (4) making use of the fact that $x_{t+1} = x_t + v_{t+1}$ and that $v_t = x_t - x_{t-1}$

$$x_{t+1} = x_t + \chi\left(x_t - x_{t-1} + \frac{\phi}{2}((g-p)u + p - x_t)\right). \quad (6)$$

Collecting terms we obtain:

$$x_{t+1} = \omega x_t - \chi x_{t-1} + \theta u + \tau \quad (7)$$

where $\omega = 1+\chi-\frac{\chi\phi}{2}$, $\theta = \frac{\chi\phi}{2}(g-p)$ and $\tau = \frac{\chi\phi p}{2}$.

If we now take the expected value of both sides of Equation (7) we obtain

$$E[x_{t+1}] = \omega E[x_t] - \chi E[x_{t-1}] + \gamma \quad (8)$$

where $\gamma = \frac{\theta}{2} + \tau$, since $E[u] = 1/2$. If, instead, we first multiply both sides of Equation (7) by x_t and then take expectations we obtain:

$$E[x_{t+1}x_t] = \omega E[x_t^2] - \chi E[x_t x_{t-1}] + \gamma E[x_t] \quad (9)$$

where we used the fact $E[ux_t] = E[u]E[x_t]$ because of the statistical independence of u and x_t. Finally, if we first square both sides of Equation (7) and apply expectations we obtain

$$\begin{aligned} E[x_{t+1}^2] &= E[(\omega x_t - \chi x_{t-1} + \theta u + \tau)^2] \\ &= \omega^2 E[x_t^2] - 2\omega\chi E[x_t x_{t-1}] + \chi^2 E[x_{t-1}^2] \\ &\quad +2\omega\gamma E[x_t] - 2\chi\gamma E[x_{t-1}] + \delta \end{aligned} \quad (10)$$

where $\delta = \frac{\theta^2}{3} + \theta\tau + \tau^2$. Given some appropriate initial conditions Equations (8), (9) and (10) allow one to compute the first and second order moments of the sampling distribution produced by SPSO during stagnation for any t.

Here, we are interested, however, in studying their asymptotic behaviour and stability. It is easy to see that Equation (8) has fixed point

$$E[x]^* = \frac{\gamma}{1+\chi-\omega} = \frac{g+p}{2}.$$

With this in hand, we can then compute fixed points of Equations (9) and (10). The most important of the two is, of course, the fixed point, $E[x^2]^*$, for the second order moment $E[x_t^2]$. This is because the variance of the sampling distribution is given by $V[x_t] = E[x_t^2] - E[x_t]^2$, and, so, its fixed-point, $V[x]^* = E[x^2]^* - (E[x]^*)^2$. We report the formula for $E[x^2]^*$ in Equation (11) and the formula for the fixed point of $E[x_t x_{t-1}]$, which we will call $E[xx]^*$, in Equation (12) (see Figure 2).

Inspection of these equations reveals that there are two situations where the fixed points might not exist. One is when $1-\omega^2-\chi^2 = 0$, being this expression at the denominator. Solving this equation for ϕ leads to the following relation

$$\phi = \frac{2}{\chi} \pm \frac{2\sqrt{1-\chi^2}}{\chi} + 2$$

As we will see below, this situation actually does not lead to instability. The other situation is when

$$1 + \chi + \frac{2\omega\chi\omega}{1-\chi^2-\omega^2} = 0$$

Equality is obtained when $\omega = 1+\chi$. But $\omega = 1+\chi-\frac{\chi\phi}{2}$ by definition. So, this implies $\phi = 0$. It is also obtained when

$$\phi \quad = \quad 4 + \frac{4}{\chi}. \quad (13)$$

This is the real stability boundary as we will show below. It is then easy to see that the recommended (χ,ϕ) pairs for the standard PSO, obtained by using Equation (5), are always within the stability region for SPSO, and, so, can freely be used in SPSO.

Equations (8), (9) and (10) form a linear system of second-order difference equations. A more precise stability analysis can be performed by turning this into the following system of first order equations

$$\mathbf{z}(t+1) = M\mathbf{z}(t) + \mathbf{b} \quad (14)$$

where

$$\mathbf{z}(t) = \left(E[x_t] \;\; E[x_{t-1}] \;\; E[x_t^2] \;\; E[x_t x_{t-1}] \;\; E[x_{t-1}^2]\right)^T$$

while

$$M = \begin{pmatrix} \omega & -\chi & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \\ 2\omega\gamma & -2\chi\gamma & \omega^2 & -2\omega\chi & \chi^2 \\ \gamma & 0 & \omega & -\chi & 0 \\ 0 & 0 & 1 & 0 & 0 \end{pmatrix} \quad \text{and} \quad \mathbf{b} = \begin{pmatrix} \gamma \\ 0 \\ \delta \\ 0 \\ 0 \end{pmatrix}.$$

It is then trivial to verify under what conditions $E[x_t]$, $E[x_t^2]$ and $E[x_t x_{t-1}]$ (thereby also $StdDev[x_t]$) will converge to stable fixed-points. All eigenvalues of M must be

$$E[x^2]^* = \frac{\delta + 2E[x]^*\gamma\omega - 2E[x]^*\chi\gamma - \frac{2\chi\omega\left(E[x]^*\left(\frac{\theta}{2}+\tau\right)+\frac{\omega(\delta+2E[x]^*\gamma\omega-2E[x]^*\chi\gamma)}{1-\chi^2-\omega^2}\right)}{\chi+\frac{2\chi\omega^2}{1-\chi^2-\omega^2}+1}}{1-\omega^2-\chi^2} \tag{11}$$

$$E[xx]^* = \frac{E[x]^*\left(\frac{\theta}{2}+\tau\right) + \frac{\omega(\delta+2E[x]^*\gamma\omega-2E[x]^*\chi\gamma)}{1-\chi^2-\omega^2}}{\chi+\frac{2\chi\omega^2}{1-\chi^2-\omega^2}+1} \tag{12}$$

Fig. 2. Fixed points for the moments $E[x_t^2]$ and $E[x_t x_{t-1}]$ of the sampling distribution of SPSO.

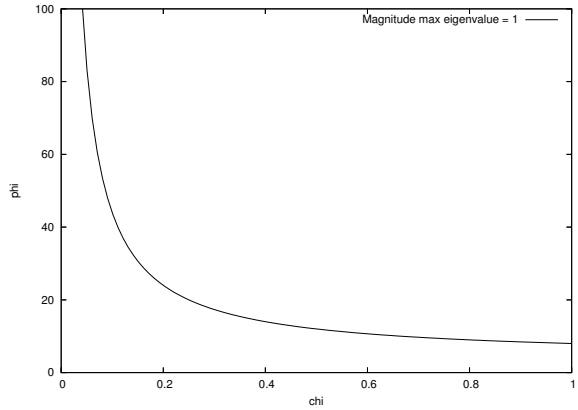

Fig. 3. Region of stability for SPSO.

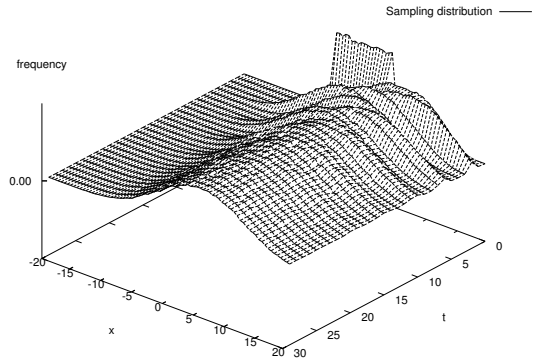

Fig. 4. Empirical sampling distribution for SPSO, when $g = 10$, $p = 0$, $\phi = 4$ and $\chi = 0.7298$.

within the unit circle, i.e. $\Lambda_m = \max_i |\lambda_i| < 1$. When this happens we will say that SPSO is *order-2 stable*.

Naturally, when $\Lambda_m < 1$, in principle we could symbolically derive the fixed-point for the system, which we will denote as $\mathbf{z}^*$. This would be simply given by

$$\mathbf{z}^* = (I - M)^{-1}\mathbf{b}$$

When the system is order-2 stable, by the simple change of variables $\tilde{\mathbf{z}}(t) = \mathbf{z}(t) - \mathbf{z}^*$ we can then represent the dynamics of the system via following linear homogeneous equation

$$\tilde{\mathbf{z}}(t+1) = M\tilde{\mathbf{z}}(t)$$

which can trivially be integrated to obtain the explicit solution

$$\tilde{\mathbf{z}}(t) = M^t\tilde{\mathbf{z}}(0).$$

Naturally, all these operations can be performed numerically once χ and ϕ are fixed. An empirical analysis of eigenvalues reveals that the region of order-2 stability (i.e, where $\Lambda_m < 1$) for SPSO is actually limited by Equation (13). This region is shown in Figure 3. Interestingly, this region coincides with the region of order-1 stability, that is with the region where Equation (8) is stable, while in the canonical PSO these two regions overlap but are not identical [8]. Note that the region of stability is much bigger than the region of stability for the canonical PSO, where rarely can one use values of ϕ bigger than 4. For SPSO, on the contrary all values of $\phi < 8$ produce PSOs, irrespective of the value of χ.

Following the same approach (as we did in [8] for the canonical PSO), we can study also higher order moments of the sampling distribution of SPSO and determine whether these converge to a stable fixed point or not. The analysis of higher order moments still leads to linear systems, and so standard eigenvalue analysis can be used to determine their stability.

We computed the M matrix for order-3 and order-4 moments of SPSO. Surprisingly, the regions of stability of skewness and kurtosis also coincide with the region depicted in Figure 3. That is, for SPSO either the mean, variance, skewness and kurtosis converge to stable fixed points, or none does. This is not the case for the canonical PSO, where the regions of convergence for skewness and kurtosis are smaller than the order-2 and -1 stable regions. For example, the canonical PSO with $\phi = 4.1$ and $\chi = 0.7298$ has stable mean, variance and skewness, but an unstable kurtosis that grows indefinitely with t [8].

IV. Sampling distribution

The sampling distribution for SPSO empirically estimated in 100,000 runs is shown in Figure 4 for the case $g = 10$, $p = 0$, $\phi = 4$ and $\chi = 0.7298$. Starting from a uniform initial distribution (we chose x_0 and v_0 uniformly at random in the range $[-5, +5]$), within one iteration the distribution becomes bell-shaped (this is the reason for the wedge shaped bump in the background of the figure). It then oscillates around the fixed-point $E[x]^* = \frac{g+p}{2} = 5$ for around 20

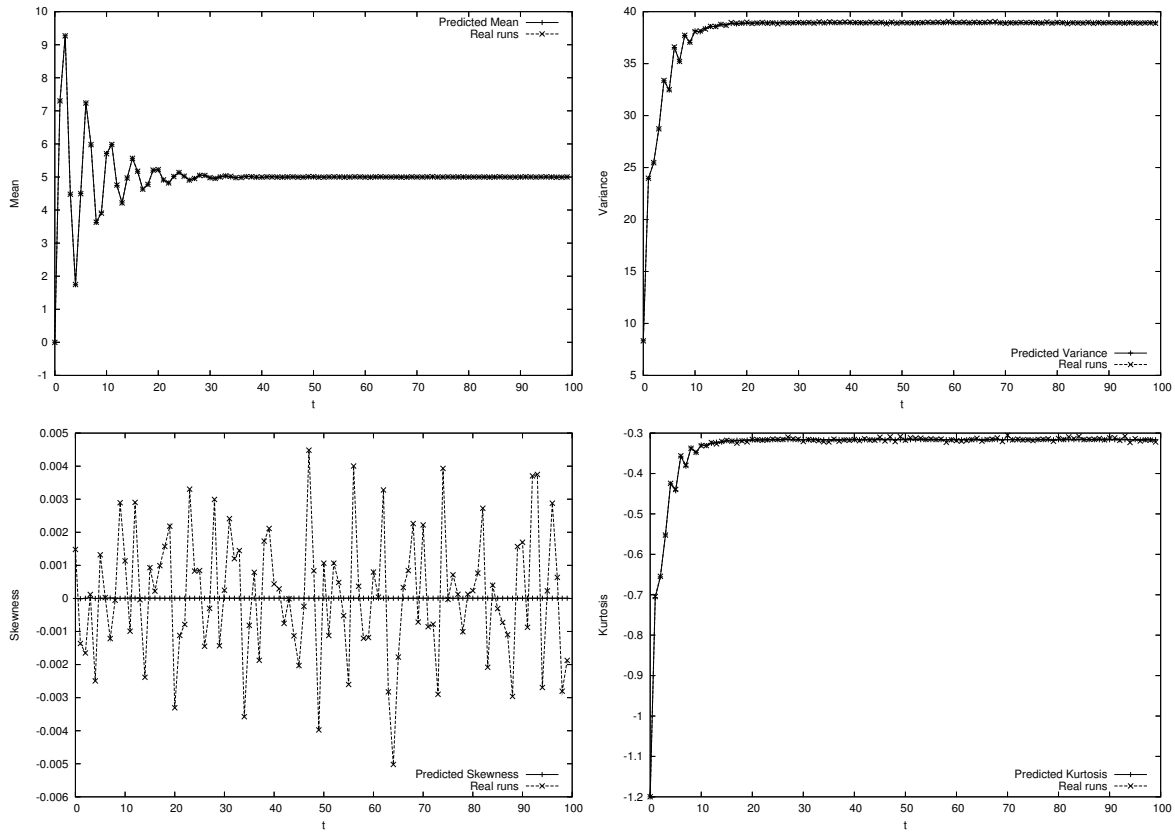

Fig. 5. Mean, variance, skewness and excess kurtosis predicted by the model and corresponding quantities estimated in 1,000,000 independent runs for $g = 10$, $p = 0$, $\phi = 4$ and $\chi = 0.7298$.

generations, after which it stabilises.

The variations of the moments of the distribution for this and other parameter settings are exactly predicted by the models described in the previous section. This is illustrated, for example, in Figure 5 where we report the mean, variance, skewness and excess kurtosis predicted by the model and compare them to the corresponding quantities estimated in 1,000,000 (one million) independent runs. As one can see the agreement between runs and theory is striking, confirming the exactness of our models.[1] Interestingly, while the oscillations of the mean continue for 20 or so generations, the variance and excess kurtosis effectively reach their stable states within around 10 generations. Also, the distribution remains symmetric (skewness ≈ 0) throughout the runs. Furthermore, the stable states for skewness and kurtosis suggest that the distribution is almost Gaussian.[2]

One may wonder what relevance the oscillations of the mean may have on the behaviour and performance of a PSO. These oscillations are present also in the canonical PSO [8]. So, despite it being simpler, SPSO preserves this feature. We believe this is very important, for the following reasons.

At the first oscillation, the PSO effectively overshoots its target, the middle point between g and p. This can be seen as an attempt to perform an extrapolation (similar to that performed in certain forms of crossover in real-valued genetic algorithms) as to what might be a good region to sample, based on direction of movement and two "good" points the PSO knows already, g and p. If an improved point is found (which happens with relatively high probability if the fitness function is somehow smooth) then the strategy pays off. However, the overshot is temporary. So, if an improvement is not found, the distribution will then swing back and effectively the PSO will attempt to extrapolate in the opposite direction. This swinging process continues for some time, but eventually a stable PSO settles onto a sample distribution that favours the region between g and p. This strategy seems very reasonable. The PSO will give a few "shots" to the extrapolation strategy, and turn to a more reliable, but perhaps slower strategy if the first fails.

Clearly, the duration of the oscillatory phase as well as the characteristics of the sampling distribution, particularly its variance, will influence performance. Whether it is better to try to extrapolate for a long time before giving up or *vice versa* depends on the fitness function. Likewise, whether it is better to have a wide or a narrow sampling distribution depends on the fitness function.

A nice property of SPSO is a direct correlation between the desired degrees of extrapolation and exploration and the

[1]The apparent disagreement between predicted and actual skewness is only an artifact of the scale used and of finite sample effects.

[2]A Gaussian distribution has zero excess kurtosis, so it has slightly fatter tails than our.

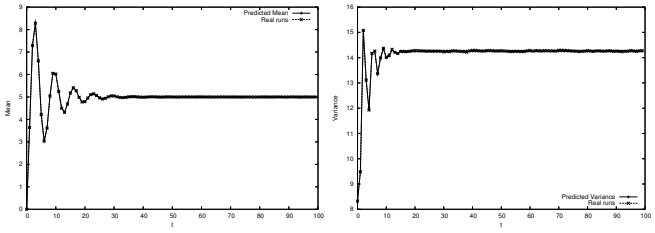

Fig. 6. Mean and variance predicted by the model and corresponding quantities estimated in 1,000,000 independent runs for $g = 10$, $p = 0$, $\phi = 2$ and $\chi = 0.7298$.

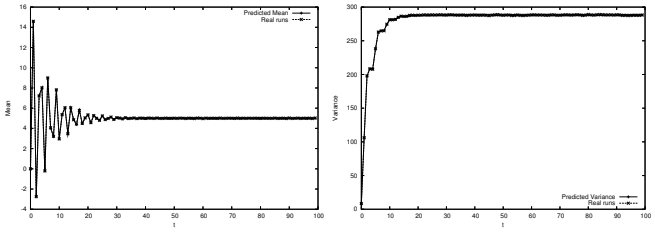

Fig. 7. As in Figure 6 but for $\phi = 8$ and $\chi = 0.7298$.

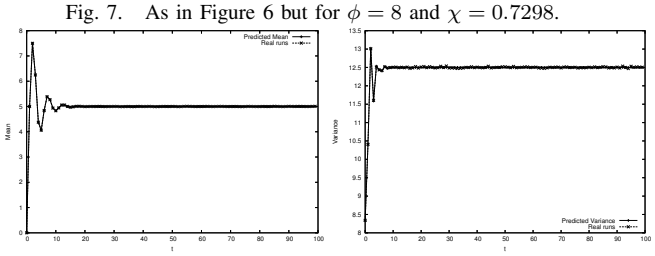

Fig. 8. As in Figure 6 but for $\phi = 4$ and $\chi = 0.5$.

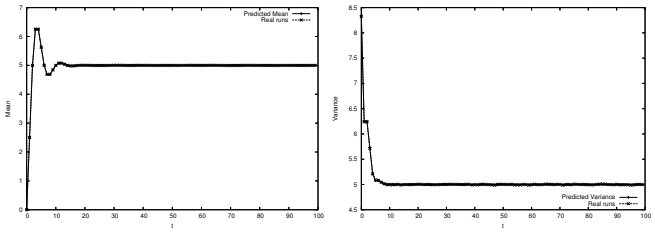

Fig. 9. As in Figure 6 but for $\phi = 2$ and $\chi = 0.5$.

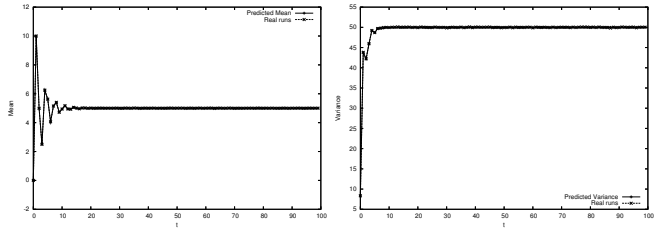

Fig. 10. Mean and variance predicted by the model and corresponding quantities estimated in 1,000,000 independent runs for $g = 10$, $p = 0$, $\phi = 8$ and $\chi = 0.5$.

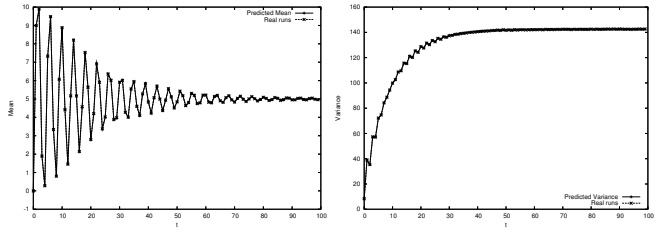

Fig. 11. As in Figure 10 but for $\phi = 4$ and $\chi = 0.9$.

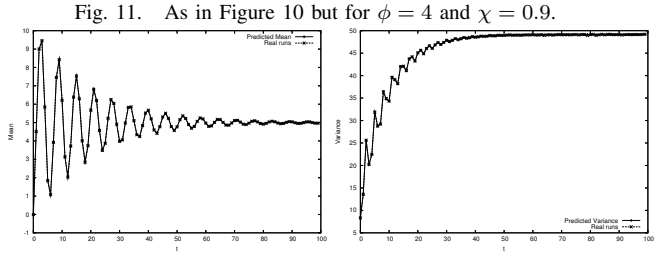

Fig. 12. As in Figure 10 but for $\phi = 2$ and $\chi = 0.9$.

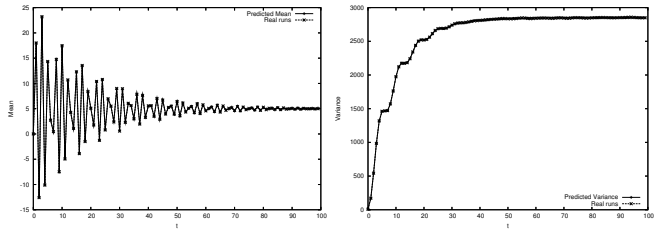

Fig. 13. As in Figure 10 but for $\phi = 8$ and $\chi = 0.9$.

parameters χ and ϕ. As shown in Figure 5 and Figures 6–13, the larger ϕ the bigger the variance of the sampling distribution, while the bigger χ the longer the duration of the transitory oscillations. So, it is easy for a designer to shape the dynamics and fixed-points of the search distribution of SPSO to better suit the problem at hand.

V. Experimental results

In the previous sections we have seen that SPSO is simpler to define, understand and control than the canonical PSO. However, this PSO would be of little interest if it could not perform well in practical experimentation.

To evaluate the performance of SPSO, we compared it with the standard constricted PSO (Equation (1)) on 14 standard benchmark problems and with two different communication topologies, lbest and gbest.

Each algorithm was run on the common benchmarks shown in Table I. Functions $f_1 - f_3$ are simple unimodal problems, $f_4 - f_9$ are highly complex multimodal problems with many local minima, and $f_{10} - f_{14}$ are multimodal problems with few local minima. This assortment of test functions should be diverse enough to give an indication of the general performance of the algorithms. In each run we performed 300,000 fitness evaluations. Both PSOs used standard values $\phi = 4.1$ and $\chi = 0.7298$ defined in [3]. The performance values given in Table II represent the mean error $|f(x) - f(x^\star)|$ found over 30 runs for a problem, where $f(x^\star)$ is the optimum fitness for the problem. Values $< 10^{-8}$ were rounded to 0.0 to allow for reproduction using programming languages that may not include floating point precision at smaller values.

In these comparisons we used the standard strategy of not evaluating particles when they leave the search space. That is, particles were allowed to cross the boundary of the feasible search space unaltered, but in these circumstances the process of fitness evaluation was skipped. This avoids the situation of having an infeasible position being set as a personal and/or global best. This method of "letting the

TABLE I

BENCHMARK FUNCTIONS.

Equation	Name	D	Feasible Bounds
$f_1 = \sum_{i=1}^{D} x_i^2$	Sphere/Parabola	30	$(-100, 100)^D$
$f_2 = \sum_{i=1}^{D} (\sum_{j=1}^{i} x_j)^2$	Schwefel 1.2	30	$(-100, 100)^D$
$f_3 = \sum_{i=1}^{D-1} \left\{ 100 \left(x_{i+1} - x_i^2 \right)^2 + (x_i - 1)^2 \right\}$	Generalized Rosenbrock	30	$(-30, 30)^D$
$f_4 = -\sum_{i=1}^{D} x_i \sin\left(\sqrt{x_i}\right)$	Generalized Schwefel 2.6	30	$(-500, 500)^D$
$f_5 = \sum_{i=1}^{D} \left\{ x_i^2 - 10\cos\left(2\pi x_i\right) + 10 \right\}$	Generalized Rastrigin	30	$(-5.12, 5.12)^D$
$f_6 = -20\exp\left\{ -0.2\sqrt{\frac{1}{D}\sum_{i=1}^{D} x_i^2} \right\} - \exp\left\{ \frac{1}{D}\sum_{i=1}^{D} \cos\left(2\pi x_i\right) \right\} + 20 + e$	Ackley	30	$(-32, 32)^D$
$f_7 = \frac{1}{4000}\sum_{i=1}^{D} x_i^2 - \prod_{i=1}^{D} \cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$	Generalized Griewank	30	$(-600, 600)^D$
$f_8 = \frac{\pi}{D}\left\{ 10\sin^2\left(\pi y_i\right) + \sum_{i=1}^{D-1} \left(y_i - 1\right)^2 \left\{ 1 + 10\sin^2\left(\pi y_{i+1}\right) \right\} + \left(y_D - 1\right)^2 \right\}$ $+ \sum_{i=1}^{D} \mu\left(x_i, 10, 100, 4\right)$ $y_i = 1 + \frac{1}{4}\left(x_i + 1\right)$ $\mu\left(x_i, a, k, m\right) = \begin{cases} k\left(x_i - a\right)^m & x_i > a \\ 0 & -a \le x_i \le a \\ k\left(-x_i - a\right)^m & x_i < -a \end{cases}$	Penalized Function P8	30	$(-50, 50)^D$
$f_9 = 0.1\left\{ \sin^2\left(3\pi x_i\right) + \sum_{i=1}^{D-1} \left(x_i - 1\right)^2 \left\{ 1 + \sin^2\left(3\pi x_{i+1}\right) \right\} + \left(x_D - 1\right)^2 \times \right.$ $\left. \left\{ 1 + \sin^2\left(2\pi x_D\right) \right\} \right\} + \sum_{i=1}^{D} \mu\left(x_i, 5, 100, 4\right)$	Penalized Function P16	30	$(-50, 50)^D$
$f_{10} = 4x_1^2 - 2.1x_1^4 + \frac{1}{3}x_1^6 + x_1x_2 - 4x_2^2 + 4x_2^4$	Six-hump Camel-back	2	$(-5, 5)^D$
$f_{11} = \left\{ 1 + \left(x_1 + x_2 + 1\right)^2 \left(19 - 14x_1 + 3x_1^2 - 14x_2 + 6x_1x_2 + 3x_2^2\right) \right\} \times$ $\left\{ 30 + \left(2x_1 - 3x_2\right)^2 \left(18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2\right) \right\}$	Goldstein-Price	2	$(-2, 2)^D$
$f_{12} = -\sum_{i=1}^{5} \left\{ \sum_{j=1}^{4} \left(x_j - a_{ij}\right)^2 + c_i \right\}^{-1}$	Shekel 5	4	$(0, 10)^D$
$f_{13} = -\sum_{i=1}^{7} \left\{ \sum_{j=1}^{4} \left(x_j - a_{ij}\right)^2 + c_i \right\}^{-1}$	Shekel 7	4	$(0, 10)^D$
$f_{14} = -\sum_{i=1}^{10} \left\{ \sum_{j=1}^{4} \left(x_j - a_{ij}\right)^2 + c_i \right\}^{-1}$	Shekel 10	4	$(0, 10)^D$

TABLE II

MEAN FITNESS AFTER 30 INDEPENDENT RUNS OF 300,000 EVALUATIONS.

Algorithm	f_1	f_2	f_3	f_4	f_5	f_6	f_7
Constricted PSO (gbest)	**0.0**	**0.0**	**8.1579**	3508	**140.4876**	**17.6628**	0.0308
Std Dev	**0.0**	**0.0**	**15.2458**	183	**26.5856**	**5.6042**	0.0346
Constricted PSO (lbest)	**0.0**	0.0855	10.1823	**3322**	163.4585	19.5116	**0.0010**
Std Dev	**0.0**	0.0614	8.1502	**126**	23.6479	0.3200	**0.0031**
SPSO (gbest)	**0.0**	**0.0**	12.8403	3882	270.0388	19.7822	0.0206
Std Dev	**0.0**	**0.0**	17.7081	192	49.365	0.0630	0.0197
SPSO (lbest)	**0.0**	0.00043	16.0938	3430	227.0989	19.7157	0.0012
Std Dev	**0.0**	0.00038	10.6249	95.9	35.3190	0.0613	0.0052
Algorithm	f_8	f_9	f_{10}	f_{11}	f_{12}	f_{13}	f_{14}
Constricted PSO (gbest)	0.1627	0.0040	**0.0**	**0.0**	4.5882	4.4747	3.8286
Std Dev	0.2987	0.0089	**0.0**	**0.0**	1.5555	2.0504	2.5600
Constricted PSO (lbest)	**0.0035**	**0.0**	**0.0**	**0.0**	2.5500	1.8899	0.5408
Std Dev	**0.0189**	**0.0**	**0.0**	**0.0**	2.5047	2.5456	1.6501
SPSO (gbest)	1.8223	1.0414	**0.0**	**0.0**	4.0784	3.5118	3.9657
Std Dev	2.6303	2.7076	**0.0**	**0.0**	2.0741	2.4713	2.4323
SPSO (lbest)	0.0484	**0.0**	**0.0**	**0.0**	**2.2203**	**0.3543**	**0.0**
Std Dev	0.1381	**0.0**	**0.0**	**0.0**	**2.5573**	**1.3485**	**0.0**

particles fly" allows infeasible particles to be drawn back into feasible space by the influence of their personal and neighborhood bests without the need for artificial means.

Swarm initialization took place in an area of the search space known to not contain the optimum. See Table III for details. If the optimum was at the origin it was independently randomly shifted in each dimension, for each separate run.

From these results it is apparent that SPSO is competitive

TABLE III

OPTIMA AND INITIALIZATION RANGES FOR ALL FUNCTIONS.

Function	Feasible Bounds	Optimum	Initialization
f_1	$(-100, 100)^D$	0.0^D	$(50, 100)^D$
f_2	$(-100, 100)^D$	0.0^D	$(50, 100)^D$
f_3	$(-30, 30)^D$	1.0^D	$(15, 30)^D$
f_4	$(-500, 500)^D$	420.9687^D	$(-500, -250)^D$
f_5	$(-5.12, 5.12)^D$	0.0^D	$(2.56, 5.12)^D$
f_6	$(-32, 32)^D$	0.0^D	$(16, 32)^D$
f_7	$(-600, 600)^D$	0.0^D	$(300, 600)^D$
f_8	$(-50, 50)^D$	-1.0^D	$(25, 50)^D$
f_9	$(-50, 50)^D$	1.0^D	$(25, 50)^D$
f_{10}	$(-5, 5)^D$	$(-0.0898, 0.7126)$, $(0.0898, -0.7126)$	$(2.5, 5)^D$
f_{11}	$(-2, 2)^D$	$(0, -1)$	$(1, 2)^D$
f_{12}	$(0, 10)^D$	4.0^D	$(7.5, 10)^D$
f_{13}	$(0, 10)^D$	4.0^D	$(7.5, 10)^D$
f_{14}	$(0, 10)^D$	4.0^D	$(7.5, 10)^D$

on many functions. It is especially good on the 3 Shekel problems (f_{12}–f_{14}).

Performance graphs for selected functions are shown in Figure 14. These show how the performance of SPSO is affected by the choice of parameters χ and ϕ. Note how the harder function requires longer periods of extrapolation (χ) and larger variances (ϕ) for good performance than the unimodal function. Diversity of SPSO population in the final generation of runs on f_1 as a function of χ and ϕ is shown in Figure 15.

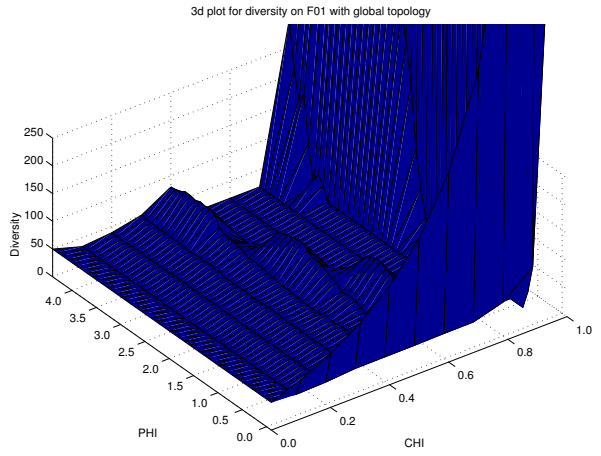

Fig. 15. Diversity of SPSO populations in the final generation of runs on f_1 as a function of χ and ϕ.

VI. CONCLUSIONS

In this paper we derived a simpler form of particle swarm optimiser, SPSO, which still retain the key properties of the original model. We did so by progressively altering the original model via mathematical transformations which have clear and understandable probabilistic effects. We studied the stability and sampling distribution during stagnation of SPSO mathematically and compared them to those of the original PSO. This showed that SPSO is more stable than the original and clarified effect of the parameters on the search strategy of SPSO. We tested SPSO on 14 standard benchmark problems and with two different communication topologies, comparing it to the canonical PSO, with very encouraging results.

We obtained SPSO via transformations that preserved key features (e.g., the mean) of the original PSO's components. However, we should note that these choices are not unique, and other transformations may make sense. So, following our approach other simplified forms of PSO could be derived. In future research we will study benefits and drawbacks of other such forms.

Finally, we should emphasise that characterisations of the search biases of an optimiser, such as our analysis of the dynamics of the sampling distribution of SPSO, are prerequisites for being able to match algorithms to problems and go beyond NFL [12]. In the future, we will attempt to use this knowledge to establish on which classes of functions the search behaviour of SPSO is expected to be beneficial.

ACKNOWLEDGEMENTS

The authors would like to acknowledge support by EPSRC XPS project (GR/T11234/01).

REFERENCES

[1] Blackwell, T. M. (2005) Particle swarms and population diversity. *Soft Computing*, 9:793–802.

[2] Clerc, M. (2006) Stagnation analysis in particle swarm optimisation or what happens when nothing happens. Technical Report CSM-460, Department of Computer Science, University of Essex, August 2006. Edited by Riccardo Poli. Also available from http://clerc.maurice.free.fr/pso/

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

[4] Campana, E.F., Fasano, G., and Pinto, A. (2006) Dynamic system analysis and initial particles position in particle swarm optimization. In *IEEE Swarm Intelligence Symposium*, Indianapolis.

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

[6] Kennedy, J. (2003) Bare bones particle swarms. *Proceedings of the IEEE Swarm Intelligence Symposium*, 80–87. Indianapolis, IN.

[7] Ozcan, E., and Mohan, C. (1999). Particle swarm optimization: surfing the waves. *Proc. 1999 Congress on Evolutionary Computation*, 1939–1944. Piscataway, NJ: IEEE Service Center.

[8] Poli, R. (2007), On the Moments of the Sampling Distribution of Particle Swarm Optimisers, Proceedings of the Workshop on Particle Swarm Optimisation: the second decade of the Genetic and Evolutionary Computation Conference (GECCO), London, July 2007, ACM Press (forthcoming).

[9] Poli, R., Langdon, W. B., Clerc, M., and Stephens, C. R. (2007) Continuous Optimisation Theory Made Easy? Finite-element Models of Evolutionary Strategies, Genetic Algorithms and Particle Swarm Optimizers, Proceedings of the 9th Foundations of Genetic Algorithms Workshop (FOGA), Mexico City, Springer, LNCS 4436 (forthcoming).

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

[11] van den Bergh, F. (2002) *An Analysis of Particle Swarm Optimizers*. PhD thesis, Department of Computer Science, University of Pretoria, Pretoria, South Africa.

[12] Wolpert, D. H., and Macready, W. G. (1997) No free lunch theorems for optimization. *IEEE Transactions on Evolutionary Computation*, 1(1):67–82.

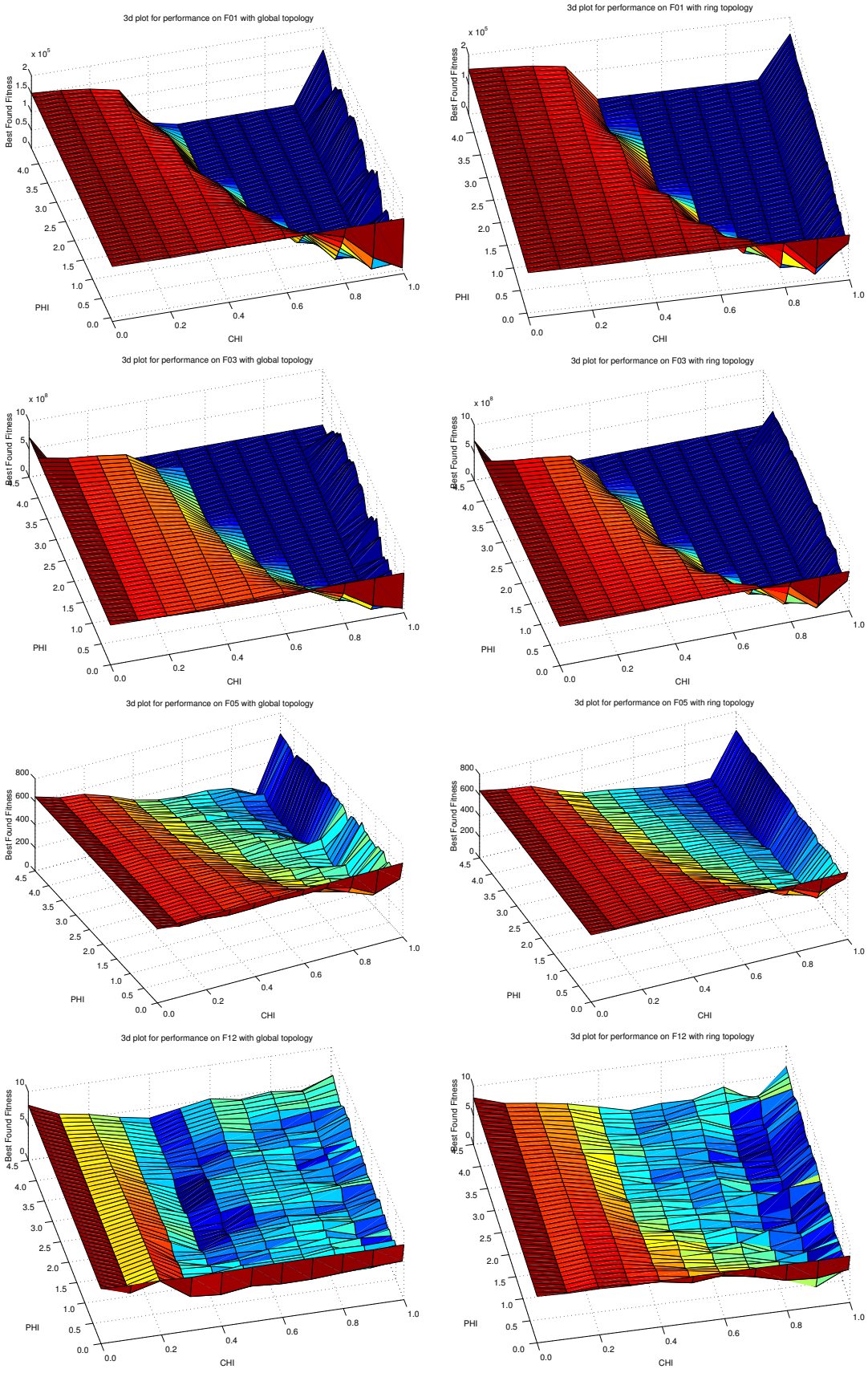

Fig. 14. Performance of SPSO on f_1, f_3, f_5 and f_{12} as a function of χ and ϕ.