J Glob Optim (2008) 41:219–244
DOI 10.1007/s10898-007-9222-5

Global optimization using a synchronization of multiple search Points autonomously driven by a chaotic dynamic model

Takashi Okamoto · Eitaro Aiyoshi

Received: 8 July 2006 / Accepted: 28 June 2007 / Published online: 28 August 2007

Abstract In the present paper, we propose a new multipoint type global optimization model using a chaotic dynamic model and a synchronization phenomenon in nonlinear dynamic systems for a continuously differentiable optimization problem. We first improve the Discrete Gradient Chaos Model (DGCM), which drives each search point's autonomous movement, based on theoretical analysis. We then derive a new coupling structure called PD type coupling in order to obtain stable synchronization of all search points with the chaotic dynamic model in a discrete time system. Finally, we propose a new multipoint type global optimization model, in which each search point moves autonomously by improved DGCM and their trajectories are synchronized to elite search points by the PD type coupling model. The proposed model properly achieves diversification and intensification, which are reported to be important strategies for global optimization in the Meta-heuristics research field. Through application to proper benchmark problems [Liang et al. Novel composition test functions for numerical global optimization. In: Proceedings of Swarm Intelligence Symposium, 2005 (SIS 2005), pp. 68–75 (2005); Liang et al. Nat. Comput. **5**(1), 83–96, 2006] (in which the drawbacks of typical benchmark problems are improved) with 100 or 1000 variables, we confirm that the proposed model is more effective than other gradient-based methods.

Keywords Global Optimization · Chaos · Multiple Points · Synchronization phenomenon · Discrete gradient model

T. Okamoto (✉)
Graduate School of Engineering, Chiba University, 1-33, Yayoi-cho, Inage-ku, Chiba-shi, Chiba 263-8522, Japan
e-mail: takashi@faculty.chiba-u.jp

E. Aiyoshi
Faculty of Science and Technology, Keio University, 3-14-1, Hiyoshi, Kouhoku-ku, Yokohama-shi, Kanagawa 223-8522, Japan
e-mail: aiyoshi@sys.appi.keio.ac.jp

1 Introduction

The development of the global optimization method, which obtains global minima without being trapped at local minima, has been investigated extensively. So-called "Physically inspired" optimization methods, which use dynamic models as computation models, have been proposed and have mainly been applied to continuously differentiable problems. The common characteristic among these models is that a global search is executed using the autonomous movement of the search point, which is driven by a vector quantity given by its dynamic system, such as a gradient vector, and the search range is then narrowed by an annealing procedure. Examples of these methods include the Discrete Gradient Chaos Model (DGCM) [1–3] and the Hamiltonian Algorithm (HA) [5].

In recent years, multipoint optimization methods, which use coupled multiple search points moving stochastically, have been proposed against the backdrop of rapid development of computer functionality. Furthermore, meta-heuristics, in which heuristics are combined based on a very good search strategy, called diversification and intensification [4], have attracted a great deal of attention. Most meta-heuristics are multipoint optimization methods. Examples of these methods include the Genetic Algorithm (GA) [6] and Particle Swarm Optimization (PSO) [7]. In meta-heuristics, diversification means that the search points should search globally without convergence to local minima. Intensification means that the search points should search intensively around the good solution. This strategy is based on the Proximate Optimality Principle (POP) (which stipulates that good solutions at one search level are likely to be found close to good solutions at an adjacent search level [4]. This often holds for practical problems). Proper combination of diversification and intensification is important for global optimization. Generally, in these methods, interaction among all search points, such as the coupling structure, is mainly used as the driven force. Therefore, these methods have a drawback in that once all of the search points are attracted to one search point, diversity is lost. In this case, either the search is terminated or a random search, which is not necessarily based on good strategy, may be executed in order to obtain diversity.

In the present paper, we propose a new multipoint type of global optimization model in which each search point is autonomously driven by a chaotic dynamic model, and the trajectories of the search points are synchronized for a continuously differentiable optimization problem with upper-lower constraints:

$$\min_{\boldsymbol{x}} \; E(\boldsymbol{x}) \tag{1a}$$

$$\text{subj. to } p_i \le x_i \le q_i, \tag{1b}$$

where $\boldsymbol{x} = [x_1, \ldots, x_N]^T \in R^N$, the objective function $E\colon R^N \to R^1$ is once continuously differentiable, the gradient of the objective function $\nabla E(\boldsymbol{x})$ is a column vector, and $i = 1, \ldots, N$, unless otherwise stated. The remainder of the present paper is organized as follows. In Chapter 2, the characteristics and drawbacks of the DGCM are discussed, based on stability analysis and one of the drawbacks is improved by the toroidalization of the constraints [8]. This model is referred to herein as the DGCMwT. The nonlinear synchronization phenomena are then discussed, and the PD type coupling model is derived in order to achieve a stable synchronization of trajectories of all search points with the chaotic dynamic models in the discrete time system in Chapter 3. In Chapter 4, we propose a new global optimization model in which the movements of multiple search points are driven by a chaotic dynamic model and in which their search trajectories are synchronized. We then confirm the effectiveness of the proposed model through application to proper benchmark problems.

In the proposed method, the DGCM is used as the driving model to implement a global search. The strong point of the DGCM is that its update equation is written using a deterministic difference equation. Therefore, we can analyze the stability of its converging solution, and we can expect the solution, to which the dynamics will converge. Hence, we choose the DGCM in order to implement a predictable global search rather than to provide random diversity. In Chapter 2, we analyze the convergence solution of the DGCM using the bifurcation diagram and the stability analysis, and we improve the DGCM based on the stability analysis. The DGCMwT, which is the improved model, converges to the widest valley, which has the widest attraction range of all solutions, by using its autonomous chaotic movement. Generally, a solution that lies on the bottom of the widest valley may be a better solution if POP holds. In particular, in the case of a problem in which the landscape is convex when viewed globally, the solution may be a global minimum or its neighbor solution. Thus, the DGCMwT has the diversification strategy due to its autonomous chaotic global search trajectory and the intensification strategy due to its convergence property. However, the DGCMwT converges to only one solution, which lies on the widest valley. If the obtained solution does not lie on the lowest valley, the DGCMwT cannot provide the global minimum. In the present paper, multiple points search using the DGCMwT and then synchronize to elite search points, the objective function values of which are smaller. Generally, it is expected that multipoint search models perform better than single-point search models, even if the original model is simply multiplexed, because the diversity is increased. In the proposed model, an increased intensification strategy, which takes the objective function value into consideration, is introduced to all search points by synchronizations to elite search points with new coupling structures, which are specialized for the optimization. This is a unique feature point of the proposed method. In the proposed method, the behavior of search points is transitioned from the "diverse search mode (the global search phase)", in which each search point autonomously implements the global search without being trapped at local minima driven by the chaotic dynamics to the "intensive search mode (the local search phase)", in which the objective function value is taken into consideration. This transition is achieved by the synchronization phenomenon in nonlinear dynamic systems. Therefore, in the intensive search mode, each search point still implements autonomous search while maintaining its chaotic search capability, i.e., while maintaining the diversification to some extent, although each search point is synchronized to elite search points and is given the intensification strategy. Thus, by introducing the synchronization to elite search points, we hope to provide a greater advantage than that given by the simple multiplexing to the proposed model.

2 Discrete gradient chaos model

Let us consider a gradient model (steepest descent method) to solve the optimization problem of Eq. 1:

$$\frac{d\boldsymbol{x}(t)}{dt} = -\nabla E(\boldsymbol{x}(t)). \tag{2}$$

The optimal solution obtained by the gradient model of Eq. 2 is the local minimum in the neighborhood of the initial point. Thus, the gradient model is not effective for global optimization. In order to provide global search ability to the gradient model, the Discrete Gradient Chaos Model (DGCM) has been proposed [1–3]. In the DGCM, the chaotic search trajectory is generated as follows. (1) The trajectory obtained by Eq. 2 is confined to the lower and upper constraints. (2) This confined trajectory is discretized by Euler's differentiation technique.

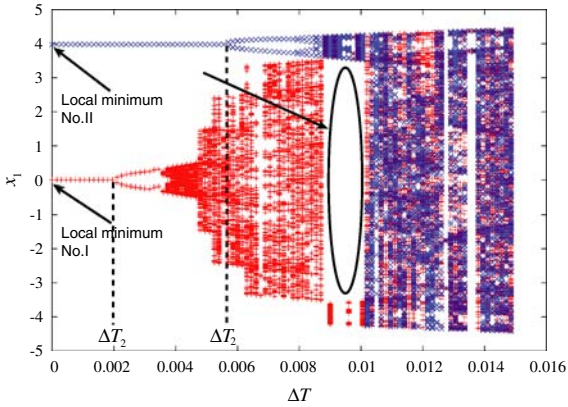


Fig. 1 Bifurcation Diagrams when the DGCM is Applied to Prob. 6. A bifurcation diagram shows the search trajectory after a sufficient time under fixing ΔT, which starts from each local minimum. In other words, a bifurcation diagram shows the search range that a chaotic search point can cover w.r.t. ΔT. In chaos annealing, the search process of the single-point DGCM corresponds to a right-to-left process. As a result of the existence of the boundary crisis, the single-point DGCM converges to local minimum No. II, which is adjacent to the boundary

(3) This trajectory is destabilized by a tuning of its sampling parameter. The study of the DGCM was started from chaos neural networks [1], and the DGCM has been investigated in several studies. The DGCM with the inner state model and linear annealing is generally given by

$$\boldsymbol{u}(k+1) = \left(1 - \frac{\Delta T(k)}{\tau}\right)\boldsymbol{u}(k) - \Delta T(k)\nabla E(\boldsymbol{x}(k)) \tag{3a}$$

$$x_i(k) = \frac{q_i + p_i \exp(-u_i(k))}{1 + \exp(-u_i(k))} \tag{3b}$$

$$\Delta T(k) = \Delta T_{\max}\left(1 - \frac{k}{k_{\max}}\right) \tag{3c}$$

where
$$\Delta T_{\max}, k_{\max} > 0, \quad \tau \gg 1, \tag{3d}$$

where $\boldsymbol{u}$ is the unconstrained inner state vector, $\Delta T_{\max}$ is the initial value of the sampling parameter, $\Delta T(k)$ is the sampling parameter at k, $k_{\max}$ is the maximum number of search steps, and τ is the parameter of a barrier function, which perturbs optimal solutions on the boundary into the interior of the lower or upper bounds.

Let us consider the optimal solution obtained by the single-point DGCM Eq. 3. First, the single-point DGCM without the annealing of Eq. 3c is applied to Prob. 6 (Original Rastrigin $N = 4$, see Appendix A). The bifurcation diagram w.r.t. ΔT is shown in Fig. 1, where bifurcation diagrams show the search trajectory after a sufficient time under fixing ΔT, which starts from local minimum No.I $(0.0, \ldots, 0.0)$ or No. II $(3.9789, \ldots, 3.9789)$.[1] In chaos annealing, ΔT is transitioned from a large value to a small value, that is, the search process of the single-point DGCM corresponds to a right-to-left process in Fig. 1. Hence, the single-point DGCM converges to local minimum No. II (which is not the global minimum). The two-bifurcation point ΔT_2, which is a boundary point of the sampling parameter between the fixed-point convergence and the two periodic oscillations, and the existence of a boundary crisis of the chaos dynamics cause this matter. In other words, the two-bifurcation point ΔT_2 of local

[1] Note that this problem has other local minima.

minimum No. II is the largest among all local minima, and the bifurcation starting from local minimum No. II does not generate a boundary crisis. Therefore, by chaos annealing, the search trajectory transitions from the chaotic trajectory to the stable trajectory converging to local minimum No. II. Even if multipoint DGCM with chaos annealing is executed in parallel, all search points converge to local minimum No.II. The convergence property of chaos annealing that is governed by this bifurcation scenario is valid in optimization problems of normal multi-peaked objective functions with no plateaus. The difference of ΔT_2 is related to the orbital stabilities of each local minimum. Let us consider the orbital stabilities of a local minimum $\boldsymbol{x}^*$. Let the inner state corresponding to $\boldsymbol{x}^*$ be $\boldsymbol{u}^*$ and its dynamics Eq. 3a be $\boldsymbol{g}(\boldsymbol{u})$, respectively, that is,

$$\boldsymbol{g}(\boldsymbol{u}) = \left(1 - \frac{\Delta T}{\tau}\right)\boldsymbol{u} - \Delta T \nabla E(\boldsymbol{x}) \tag{4a}$$

$$x_i = \frac{q_i + p_i \exp(-u_i)}{1 + \exp(-u_i)}. \tag{4b}$$

Let us consider the Jacobian of $\boldsymbol{g}$ under $\tau \gg 1$

$$D\boldsymbol{g}(\boldsymbol{u}^*) = I - \Delta T A(\boldsymbol{u}^*), \tag{5a}$$

$$\text{where } A(\boldsymbol{u}^*) = \begin{pmatrix} \dfrac{\partial x_1}{\partial u_1}\dfrac{\partial^2 E(\boldsymbol{x}^*)}{\partial x_1^2} & \cdots & \dfrac{\partial x_N}{\partial u_N}\dfrac{\partial^2 E(\boldsymbol{x}^*)}{\partial x_N \partial x_1} \\ \vdots & \ddots & \vdots \\ \dfrac{\partial x_1}{\partial u_1}\dfrac{\partial^2 E(\boldsymbol{x}^*)}{\partial x_1 \partial x_N} & \cdots & \dfrac{\partial x_N}{\partial u_N}\dfrac{\partial^2 E(\boldsymbol{x}^*)}{\partial x_N^2} \end{pmatrix}. \tag{5b}$$

Let eigenvalues of $D\boldsymbol{g}(\boldsymbol{u}^*)$ be $\lambda_1(\boldsymbol{u}^*), \ldots, \lambda_n(\boldsymbol{u}^*)$ in descending order of their absolute values. The stability of the local minimum $\boldsymbol{u}^*$ is then governed by λ_1. Next, we consider two-bifurcation point ΔT_2 using $\lambda_1(\boldsymbol{u}^*)$. Let the eigenvalues of $A(\boldsymbol{u}^*)$ be $\boldsymbol{\lambda}^a(\boldsymbol{u}^*)$. Then, the eigenvalues of $D\boldsymbol{g}(\boldsymbol{u}^*)$ are given by

$$\lambda_i(\boldsymbol{u}^*) = 1 - \Delta T \lambda_i^a(\boldsymbol{u}^*). \tag{6}$$

Let the eigenvalue with the largest absolute value in $\boldsymbol{\lambda}^a(\boldsymbol{u}^*)$ be $\lambda_1^a(\boldsymbol{u}^*)$. Then, since $A(\boldsymbol{u}^*)$ is a positive definite matrix, if $0 < \lambda_1^a(\boldsymbol{u}^*) \leq 2/\Delta T$, then $|\lambda_i(u^*)| < 1$ and the local minimum $\boldsymbol{u}^*$ is stable. Whereas, if $\lambda_1^a(\boldsymbol{u}^*) > 2/\Delta T$, then

$$\lambda_1(\boldsymbol{u}^*) = 1 - \Delta T \lambda_1^a(\boldsymbol{u}^*) < -1 \tag{7}$$

and $\boldsymbol{u}^*$ is unstable. Hence, two-bifurcation point $\Delta T_2(\boldsymbol{u}^*)$, which is the boundary between the fixed point convergence and the two periodic oscillations, is given by

$$\Delta T_2(\boldsymbol{u}^*) = \frac{2}{\lambda_1^a(\boldsymbol{u}^*)} \tag{8}$$

as the boundary value, where $\lambda_1^a(\boldsymbol{u}^*)$ changes from a stable eigenvalue, which makes the local minimum $\boldsymbol{u}^*$ stable, to an unstable eigenvalue, which makes the local minimum unstable. For example, in the case of one-variable problems, a unique eigenvalue $\lambda_1^a(u^*)$ is given by the following using Eq. 7:

$$\lambda_1^a(u^*) = \frac{dx}{du}\frac{d^2E(x^*)}{dx^2}. \tag{9}$$

Hence, $\Delta T_2(u^*)$ is given by

$$\Delta T_2(u^*) = \frac{2}{\dfrac{dx}{du}\dfrac{d^2E(x^*)}{dx^2}}. \tag{10}$$

The DGCM tends to converge to the optimal solution x^*, the ΔT_2 of which is larger, as stated above, and

$$\frac{dx}{du} = \frac{(q-p)\exp(-u)}{(1+\exp(-u))^2} > 0. \tag{11}$$

The closer x^* is to the upper or lower boundary, the smaller the dx/du. Similarly, in the case of multiple-variable problems, we expect that eigenvalue $\lambda_1(\boldsymbol{u}^*)$ is smaller at the local minimum, which is located adjacent to the boundary. Hence, the DGCM tends to converge to the optimal solution, which is located adjacent to the upper or the lower boundary. This is one of the drawbacks of the DGCM.

In order to overcome this drawback, we propose that the confining method be changed from the expression by inner states to the toroidalization of the constraint [8]. In this method, the search region is transformed into an N-dimensional torus. An image of this method is shown in [8] (Fig. 1). The search is then executed in this torus, and the trajectory is confined to the constraint. The DGCM with toroidalization of the constraint is given by

$$\boldsymbol{x}'(k+1) = \boldsymbol{x}(k) - \Delta T(k)\nabla E(\boldsymbol{x}(k)) \ (= \boldsymbol{g}(\boldsymbol{x})) \tag{12a}$$

$$x_i(k+1) = \tilde{f}(x_i'(k+1)) \tag{12b}$$

$$\tilde{f}(x_i') = \begin{cases} x_i', & p_i \le x_i' \le q_i \\ (x_i' - p_i)\mathrm{mod}(q_i - p_i) + p_i, & x_i' > q_i \\ (x_i' - q_i)\mathrm{mod}(q_i - p_i) + q_i, & x_i' < p_i \end{cases} \tag{12c}$$

$$\Delta T(k) = \Delta T_{\max}\left(1 - \frac{k}{k_{\max}}\right) \tag{12d}$$

$$\text{where } \Delta T_{\max}, k_{\max} > 0. \tag{12e}$$

This model is referred to herein as the DGCMwT. In the case of the DGCMwT, the Jacobian is given by

$$D\boldsymbol{g}(\boldsymbol{x}^*) = I - \Delta T \nabla^2 E(\boldsymbol{x}^*), \tag{13}$$

where $\nabla^2 E(\boldsymbol{x})$ is the Hessian of $E(\boldsymbol{x})$. Let the eigenvalues of $\nabla^2 E(\boldsymbol{x})$ be $\boldsymbol{\lambda}^{\nabla^2}$. Then, two-bifurcation point $\Delta T_2(\boldsymbol{x}^*)$ is given by

$$\Delta T_2(\boldsymbol{x}^*) = \frac{2}{\lambda_1^{\nabla^2}(\boldsymbol{x}^*)}, \tag{14}$$

and, in the case of one-variable problems, $\Delta T_2(x^*)$ is given by

$$\Delta T_2(x^*) = \frac{2}{\dfrac{d^2E(x^*)}{dx^2}}. \tag{15}$$

Hence, the DGCMwT tends to converge to the local minimum, where only $d^2E(x^*)/dx^2$ (or $\lambda_1^{\nabla^2}$, for multiple-variable problems) is smaller, and we expect that the abovementioned drawback of the DGCM is improved. Actually, the bifurcation diagram of the application of

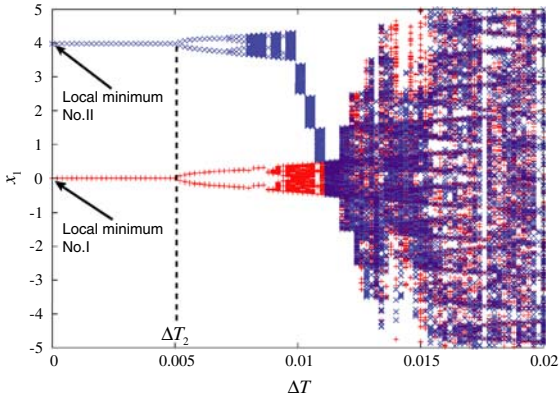


Fig. 2 Bifurcation Diagrams when the DGCMwT is Applied to Prob. 6. The stabilities of all local minima are approximately equal in Prob. 6. Therefore, ΔT_2 should be approximately equal in this problem. ΔT_2 are approximately equal, and there is no boundary crisis in this figure, in which the DGCMwT is applied, whereas ΔT_2 were not equal, and the boundary crisis existed in Fig. 1, in which the DGCM was applied

the DGCMwT to Prob. 6 (Fig. 2) confirms this improvement. Thus, the DGCMwT provides the most stable local minimum as the optimal solution without relation to its position or the constraint. Generally, such solutions lie on the bottom of the widest valley, which has the widest attraction range of all solutions. Such solutions may be better solutions if POP holds. In particular, in the case of a problem in which the landscape is convex when it is viewed globally (for example Prob. 6), the solution may be a global minimum or a neighbor solution. However, the DGCMwT converges to only one solution. If the solution does not lie on the lowest valley, then the DGCMwT cannot provide the global minimum. Hence, the DGCMwT should be improved in order to provide diversity, i.e., the possibility of the trajectories searching in multiple valleys, and at the same time, the DGCMwT should be improved to provide an increased intensification strategy, which takes the objective function value into consideration, i.e., the strategy that intensively searches the regions in which the objective function value is smaller.

3 Synchronization phenomena of coupled nonlinear oscillators

In the present paper, we achieve the abovementioned improvements with a synchronization of chaotic search trajectories. Coupled nonlinear oscillators generate synchronization phenomena. Here, the DGCMwT is a nonlinear oscillator in a discrete time system. Consequently, in this chapter, synchronization phenomena in discrete time systems are discussed with respect to continuous time systems.

3.1 Synchronization phenomena in continuous time systems

Generally, the coupled oscillators system with N variables in continuous time systems is formulated as

$$\frac{d\boldsymbol{x}^p(t)}{dt} = \boldsymbol{f}^p(\boldsymbol{x}^p(t)) + c\Gamma(\{\boldsymbol{x}^j(t)\}, \boldsymbol{x}^p(t)), \quad p = 1, \ldots, P. \tag{16}$$

Let Γ express a coupling structure, and let the pth oscillator $d\boldsymbol{x}^p(t)/dt = \boldsymbol{f}^p(\boldsymbol{x}^p(t))$ be periodic. Γ in Eq. 16 includes the following structures:

(a) global diffusive coupling

$$\Gamma(\{\boldsymbol{x}^j(t)\}, \boldsymbol{x}^p(t)) = \sum_{j=1}^{P} \left\{\boldsymbol{x}^j(t) - \boldsymbol{x}^p(t)\right\}, \tag{17}$$

(b) closest convective coupling (periodic bound)

$$\Gamma(\{\boldsymbol{x}^j(t)\}, \boldsymbol{x}^p(t)) = \boldsymbol{x}^{p+1}(t) - \boldsymbol{x}^p(t). \tag{18}$$

It is known that Eq. 16, in which coupling (a) or (b) is employed as Γ, generates synchronization phenomena if the coupling coefficient $c(>0)$ exceeds a certain value [9].

3.2 Synchronization phenomena in discrete time systems and PD type coupling model

Differential equations in continuous time systems can be transformed into corresponding discrete maps, that is, each oscillator's differential equation with no coupling:

$$\frac{d\boldsymbol{x}^p(t)}{dt} = \boldsymbol{f}^p(\boldsymbol{x}^p(t)) \tag{19}$$

can be transformed into a corresponding discrete map:

$$\boldsymbol{x}^p(k+1) = \boldsymbol{g}^p(\boldsymbol{x}^p(k)). \tag{20}$$

Let us consider a coupling model that consists of discrete maps. By replacing t in Eq. 16 by k naively, we can obtain a coupling discrete map:

$$\boldsymbol{x}^p(k+1) = g^p(\boldsymbol{x}^p(k)) + c\Gamma(\{\boldsymbol{x}^j(k)\}, \boldsymbol{x}^p(k)). \tag{21}$$

However, setting c to a large value in order to cause synchronization phenomena has been reported to result in divergences of trajectories in the system Eq. 20 and makes the system unstable [10]. In the following part of this section, another synchronization model of coupled discrete oscillators is considered for emerging synchronizations in the discrete time optimization model, such as the DGCMwT. The derivation of such a coupled discrete oscillators system includes the derivation of Fujisaka and Yamada [10] using Poincare maps. In the present paper, we introduce a simpler derivation using Proportional-Derivative (PD) type coupling and Euler's differentiation technique.

Let us introduce a coupling using a proportional term and a derivative term into Eq. 16. We then obtain

$$\frac{d\boldsymbol{x}^p(t)}{dt} = \boldsymbol{f}^p(\boldsymbol{x}^p(t)) + c\Gamma\left(\left\{\boldsymbol{x}^j(t) + \Delta T\frac{d\boldsymbol{x}^j(t)}{dt}\right\}, \boldsymbol{x}^p(t) + \Delta T\frac{d\boldsymbol{x}^p(t)}{dt}\right). \tag{22}$$

Let $\Delta T > 0$. We herein refer to the coupling structure Γ in Eq. 22 as the PD type coupling structure. The PD type coupling structure is characterized by the estimation mechanism in the sense of the coupling with the states after ΔT. Discretizing Eq. 22 by Euler's differentiation with a sampling parameter $\Delta T = t/k$ gives

$$\frac{\boldsymbol{x}^p(t+\Delta T)-\boldsymbol{x}^p(t)}{\Delta T}=\boldsymbol{f}^p(\boldsymbol{x}^p(t))+c\Gamma\left(\left\{\boldsymbol{x}^j(t)+\Delta T\frac{\boldsymbol{x}^j(t+\Delta T)-\boldsymbol{x}^j(t)}{\Delta T}\right\},\right.$$
$$\left.\boldsymbol{x}^p(t)+\Delta T\frac{\boldsymbol{x}^p(t+\Delta T)-\boldsymbol{x}^p(t)}{\Delta T}\right) \tag{23}$$
$$\boldsymbol{x}^p(t+\Delta T)=\boldsymbol{x}^p(t)+\Delta T\boldsymbol{f}^p(\boldsymbol{x}^p(t))+\Delta Tc\Gamma\left(\{\boldsymbol{x}^j(t+\Delta T)\},\boldsymbol{x}^p(t+\Delta T)\right).$$

Let us define the Euler's discrete mapping of Eq. 19

$$\boldsymbol{x}^p(t+\Delta T)=\boldsymbol{x}^p(t)+\Delta T\boldsymbol{f}^p(\boldsymbol{x}^p(t)) \tag{24}$$

as

$$\boldsymbol{x}^p(k+1)=\boldsymbol{g}^p(\boldsymbol{x}^p(k)),\ \text{ where } t=k\Delta T. \tag{25}$$

Redefining ΔTc as c in Eq. 23 gives

$$\boldsymbol{x}^p(k+1)=\boldsymbol{g}^p(\boldsymbol{x}^p(k))+c\Gamma(\{\boldsymbol{x}^j(k+1)\},\boldsymbol{x}^p(k+1)) \tag{26}$$

as a discrete mapping of Eq. 22. Furthermore, setting

$$\boldsymbol{x}_i=\begin{pmatrix}x_i^1\\ \vdots\\ x_i^P\end{pmatrix},\ X=\left[\boldsymbol{x}^1,\ldots,\boldsymbol{x}^P\right],\ \boldsymbol{g}_i(X)=\begin{pmatrix}g_i^1(\boldsymbol{x}^1)\\ \vdots\\ g_i^P(\boldsymbol{x}^P)\end{pmatrix} \tag{27}$$

and solving Eq. 26 for $X(k+1)$ gives

$$\boldsymbol{x}_i(k+1)=C^{-1}\boldsymbol{g}_i(X(k)). \tag{28}$$

For example, if the coupling structure is the global diffusive coupling Eq. 17, then matrix C is given by

$$C_{ij}=\begin{cases}(P-1)c+1 & (j=i)\\ -c & (j\neq i),\end{cases} \tag{29}$$

and if the coupling structure is the closest convective coupling Eq. 18, then matrix C is given by

$$C_{ij}=\begin{cases}c+1 & (j=i)\\ -c & (j=i+1)\\ 0 & (j\neq i,i+1).\end{cases} \tag{30}$$

In the present paper, we refer to this model as the PD type coupling model. This model uses the coupling structure (D type), which "estimates" the jth search point's position at time $k+1$, as well as the conventional coupling structure (P type). Therefore, we expect that this new coupling model will generate more stable synchronization phenomena in coupled discrete oscillations, compared to those generated by the conventional coupling model.

3.3 Simulation and discussion

In this section, the effectiveness of the PD type coupling model for the synchronization of chaotic trajectories is confirmed through a simulation and stability analysis.

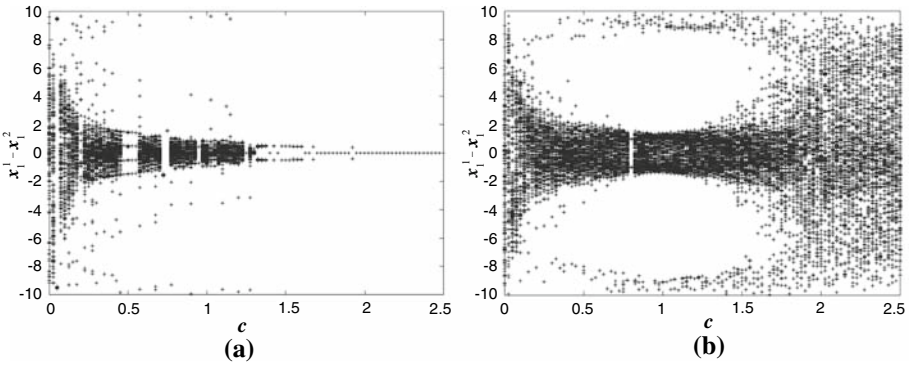


Fig. 3 Bifurcation diagrams of the differences between each oscillator. (**a**) Case of Eq. 32; (**b**) Case of Eq. 33

For simplicity, let $P = 2$ in the PD type coupling model. Then, the coefficient matrix C is

$$C^{-1} = \frac{1}{1+2c}\begin{pmatrix} 1+c & c \\ c & 1+c \end{pmatrix}, \tag{31}$$

in the cases of both the global diffusive coupling and the closet convective coupling. Hence, the DGCMwT synchronization model with the PD type coupling $P = 2$ is written as

$$\begin{pmatrix} x_i^{1\prime}(k+1) \\ x_i^{2\prime}(k+1) \end{pmatrix} = \frac{1}{1+2c}\begin{pmatrix} 1+c & c \\ c & 1+1c \end{pmatrix}\begin{pmatrix} g_i(\boldsymbol{x}^1(k)) \\ g_i(\boldsymbol{x}^2(k)) \end{pmatrix} \tag{32a}$$

$$x_i^p(k+1) = \tilde{f}(x_i^{p\prime}(k+1)) \tag{32b}$$

$$g_i(\boldsymbol{x}^p(k)) = x_i^p(k) - \Delta T \frac{\partial E(\boldsymbol{x}^p(k))}{\partial x_i}, \tag{32c}$$

where $\tilde{f}$ in Eq. 32b is the toroidalization function given by Eq. 12c. The model Eq. 32 was targeted in the simulation. For comparison, a synchronization model using Eq. 21 was also simulated.

$$\begin{pmatrix} x_i^{1\prime}(k+1) \\ x_i^{2\prime}(k+1) \end{pmatrix} = \begin{pmatrix} g_i(\boldsymbol{x}^1(k)) + c(x_i^2(k) - x_i^1(k)) \\ g_i(\boldsymbol{x}^2(k)) + c(x_i^1(k) - x_i^2(k)) \end{pmatrix} \tag{33a}$$

$$x_i^p(k+1) = \tilde{f}(x_i^{p\prime}(k+1)) \tag{33b}$$

$$g_i(\boldsymbol{x}^p(k)) = x_i^p(k) - \Delta T \frac{\partial E(\boldsymbol{x}^p(k))}{\partial x_i} \tag{33c}$$

We employed benchmark problem Prob. 6 (Original Rastrigin, $N = 4$) as $E(\boldsymbol{x})$ and set the sampling parameter $\Delta T = 0.02$, at which chaotic trajectories were generated. In order to illustrate the dependence on coupling coefficients of emerging synchronizations in these models, numerical simulations were run, changing the coupling coefficient c over the range of $c = 0.0$–2.5. Bifurcation diagrams of the differences between each oscillator $x_1^1 - x_1^2$ are shown in Fig. 3a and b. Figure 3a shows that a chaotic synchronization phenomenon emerges stably from the PD type coupling model when c is set larger. In contrast, Fig. 3b shows that a chaotic synchronization phenomenon does not emerge from the unstable coupling model as Eq. 33. Thus, the PD coupling model should be used for discrete chaotic synchronization.

Next, we consider a condition for c to generate a synchronization in the PD type coupling model with the stability analysis. In the case of $P = 2, N = 1$, we derive the condition in Appendix B. Let λ be the Lyapunov index $\lambda = \left\langle \ln \left| \frac{g(x^*)}{dx} \right| \right\rangle$. Then, the condition is given by

$$\frac{\exp(\lambda)}{1+2c} < 1. \tag{34}$$

Hence, setting c to a larger value generates the chaotic synchronization phenomenon using the PD type coupling model. In the case of $P = 2, N > 1$, a similar discussion is not easy. However, we can rewrite Eq. 50 (in Appendix B) as

$$\begin{pmatrix} \delta x_i^1(k+1) \\ \delta x_i^2(k+1) \end{pmatrix} = C^{-1} \sum_{j=1}^{N} \left\{ \frac{\partial g_i(\boldsymbol{x}^*(k))}{\partial x_j} \begin{pmatrix} \delta x_j^1(k) \\ \delta x_j^2(k) \end{pmatrix} \right\}. \tag{35}$$

Equation 35 suggests that c is a similar factor in the case of $P = 2, N > 1$. That is, we expect that setting c to a larger value generally provides the emergence of the chaotic synchronization phenomenon.

In this section, we confirm the emergence of the chaotic synchronization phenomenon in the gradient system using the PD type coupling model. In the next chapter, we propose a new optimization model by introducing this synchronization phenomenon with the PSO type advective coupling model (described in the following chapter).

4 A global optimization model using a synchronization of chaotic multiple search points

In this chapter, we propose a new multipoint type global optimization model in which each search point is autonomously driven by the DGCMwT and their search trajectories are synchronized by the PD type coupling model to their elite search points. The proposed model provides an increased diversification strategy by multiple-point autonomous chaotic search and an increased intensification strategy, which takes the objective function value into consideration, in addition to that of the DGCMwT. We illustrate the capability of the proposed model through numerical simulations, in which the model is applied to several "proper" benchmark problems, the dimension of variables of which is 100 ($N = 100$) and 1000 ($N = 1000$).

4.1 Proposed model

Let us consider that P search points are autonomously driven by the DGCMwT. This uncoupled multipoint type DGCMwT can search in multiple valleys. However, as mentioned in Chapter 2, nearly all of the trajectories are trapped at the most stable local minima, which are not always global minima, after the annealing. Therefore, we present an intensification strategy based on the objective function value and try to overcome this property. In order to provide the intensification, we introduce the PD type coupling model, in which the search trajectory of each search point synchronizes to the elite search points inspired by PSO. The elite search points are classified as follows:

1. An elite search point of "each search point" "until" the time k (personal best : pbest)

$$\boldsymbol{pb}^p(k) = \underset{\boldsymbol{x}^p(i)}{\operatorname{argmin}} \left\{ E(\boldsymbol{x}^p(i)) \mid i = 0, \ldots, k \right\} \tag{36}$$

2. An elite search point of "all search points" "until" the time k (global best : gbest)

$$\boldsymbol{pb}^{gb}(k) = \underset{\boldsymbol{pb}^p(k)}{\operatorname{argmin}} \left\{ E(\boldsymbol{pb}^p(k)) \mid p = 1, \ldots, P \right\} \tag{37}$$

3. An elite search point of "all search points" "at" the time k (current best : cbest)

$$\boldsymbol{x}^{cb}(k) = \underset{\boldsymbol{x}^p(k)}{\operatorname{argmin}} \left\{ E(\boldsymbol{x}^p(k)) \mid p = 1, \ldots, P \right\}. \tag{38}$$

In the proposed model, each search point is coupled with two elite search points, one is pbest and the other (which is represented as $\boldsymbol{gcb}$) is gbest or cbest. Let $\boldsymbol{x}^j(k+1)$ of Γ in the PD type coupling model of Eq. 26 be $\boldsymbol{gcb}$ and $\boldsymbol{pb}^p$ and $\boldsymbol{g}(\boldsymbol{x}^p(k))$ be the map of the DGCMwT. Then, the PD type coupling model with the DGCMwT and the aforementioned elite search points are given by

$$\begin{aligned} \boldsymbol{x}^{p\prime}(k+1) = {} & \boldsymbol{g}(\boldsymbol{x}^p(k)) + c_1(\boldsymbol{gcb}(k+1) - \boldsymbol{x}^{p\prime}(k+1)) \\ & + c_2(\boldsymbol{pb}^p(k+1) - \boldsymbol{x}^{p\prime}(k+1)). \end{aligned} \tag{39}$$

We refer to the coupling structure given in Eq. 39 as the PSO type advective coupling. For calculating Eq. 39, we cannot determine $\boldsymbol{gcb}(k+1)$ or $\boldsymbol{pb}^p(k+1)$ at time k. We consider the possibility of updates of elite search points at one step $(k \to k+1)$ to be small, and then let $\boldsymbol{gcb}(k+1) \approx \boldsymbol{gcb}(k)$ and $\boldsymbol{pb}^p(k+1) \approx \boldsymbol{pb}^p(k)$. Solving Eq. 39 for $\boldsymbol{x}^{p\prime}(k+1)$ with this assumption, we obtain a new optimization model:

$$\begin{aligned} \boldsymbol{x}^{p\prime}(k+1) = {} & \frac{1}{1+c_1+c_2}\boldsymbol{g}(\boldsymbol{x}^p(k)) + \frac{c_1}{1+c_1+c_2}\boldsymbol{gcb}(k) \\ & + \frac{c_2}{1+c_1+c_2}\boldsymbol{pb}^p(k) \end{aligned} \tag{40a}$$

$$x_i^p(k+1) = \tilde{f}(x_i^{p\prime}(k+1)) \tag{40b}$$

$$g_i(\boldsymbol{x}^p(k)) = x_i^p(k) - \Delta T(k)\frac{\partial E(\boldsymbol{x}^p(k))}{\partial x_i} \tag{40c}$$

$$\Delta T(k) = \Delta T_{\max}\left(1 - \frac{k}{k_{\max}}\right). \tag{40d}$$

This model achieves intensification based on the objective function value with the synchronization of chaotic search points. However, if c_1, c_2 are fixed at a certain value, then complete synchronization phenomena are generated, that is, the diversity of the search trajectory may be lost,[2] especially, in the local search phase. In order to maintain the diversity with the intensification, c_1, c_2 in Eq. 40a are replaced by $\sin^2(2\pi k/T)\bar{c}_1$, $\sin^2(2\pi k/T)\bar{c}_2$. In other words, the synchronization and non-synchronization of the search trajectories are repeated by changing c_1, c_2 periodically. Finally, we propose the following model:

$$\begin{aligned} \boldsymbol{x}^{p\prime}(k+1) = {} & \frac{1}{1+c_1(k)+c_2(k)}\boldsymbol{g}(\boldsymbol{x}^p(k)) + \frac{c_1(k)}{1+c_1(k)+c_2(k)}\boldsymbol{gcb}(k) \\ & + \frac{c_2(k)}{1+c_1(k)+c_2(k)}\boldsymbol{pb}^p(k) \end{aligned} \tag{41a}$$

$$x_i^p(k+1) = \tilde{f}(x_i^{p\prime}(k+1)) \tag{41b}$$

[2] Note however that all search points are moved with the same trajectory.

Table 1 Parameters in EC-DGCM

Parameter	Explanation
P	Number of search points
$\bar{c}_1, \bar{c}_2$	Maximum coupling coefficient
T	Period of reiteration between synchronization and non-synchronization
$k_{\max}$	Steps of the search
$\Delta T_{\max}$	Initial sampling parameter

$$\tilde{f}(x_i^{p'}) = \begin{cases} x_i^{p'}, & p_i \le x_i^{p'} \le q_i \\ (x_i^{p'} - p_i)\mathrm{mod}(q_i - p_i) + p_i, & x_i^{p'} > q_i \\ (x_i^{p'} - q_i)\mathrm{mod}(q_i - p_i) + q_i, & x_i^{p'} < p_i \end{cases} \tag{41c}$$

$$g_i(\boldsymbol{x}^p(k)) = x_i^p(k) - \Delta T(k)\frac{\partial E(\boldsymbol{x}^p(k))}{\partial x_i} \tag{41d}$$

$$\Delta T(k) = \Delta T_{\max}\left(1 - \frac{k}{k_{\max}}\right) \tag{41e}$$

$$c_1(k) = \bar{c}_1 \sin^2\left(\frac{2\pi k}{T}\right), c_2(k) = \bar{c}_2 \sin^2\left(\frac{2\pi k}{T}\right) \tag{41f}$$

$$\boldsymbol{pb}^p(k) = \underset{\boldsymbol{x}^p(i)}{\mathrm{argmin}}\left\{E(\boldsymbol{x}^p(i)) \mid i = 0, \ldots, k\right\} \tag{41g}$$

$$\boldsymbol{gcb}(k) = \begin{cases} \underset{\boldsymbol{pb}^p}{\mathrm{argmin}}\left\{E(\boldsymbol{pb}^p(k)) \mid p = 1, \ldots, P\right\} \\ \qquad \text{or} \\ \underset{\boldsymbol{x}^p(k)}{\mathrm{argmin}}\left\{E(\boldsymbol{x}^p(k)) \mid p = 1, \ldots, P\right\}. \end{cases} \tag{41h}$$

We refer to this model as the Elite Coupling type Discrete Gradient Chaos Model (EC-DGCM). The proposed optimization model is easily coded (especially, if the DGCM or PSO has been conducted). The pseudo code of the proposed model is given in Appendix C. We can change the behavior of the EC-DGCM through the choice of the elite search point of all search points. In the case of the EC-DGCM with gbest, as the population synchronizes to gbest, the search trajectories gradually shift from the autonomous global search to the concentrative search in the neighborhood of gbest. However, using only coupling with gbest causes concentration on gbest and reduces the diversity in the local search phase. Each search point's coupling with pbest prevents concentration and maintains diversity. On the other hand, in the case of the EC-DGCM with cbest, since the population synchronizes to cbest, which is updated at each step, the population searches in regions where the objective function value is smaller, while maintaining diversity to a certain extent. In this case, pbest is considered to act as "a storage search point", which stores good solutions, rather than "a fluctuation search point".

Next, we refer to the parameters in the EC-DGCM. Table 1 shows the parameters in the EC-DGCM. We should set P and $k_{\max}$ considering the trade-off between the solution precision and the calculation cost (we use a fixed value in the present paper). $\Delta T_{\max}$ must be set to a certain value, by which chaotic trajectories are generated. $\bar{c}_1$, $\bar{c}_2$ and T govern the balance between the diversification and the intensification. Taking the fact that the DGCMwT converges to any local minimum and the EC-DGCM tends to generate complete synchronization in the local search phase into consideration, we use settings that stress diversification for these parameters.

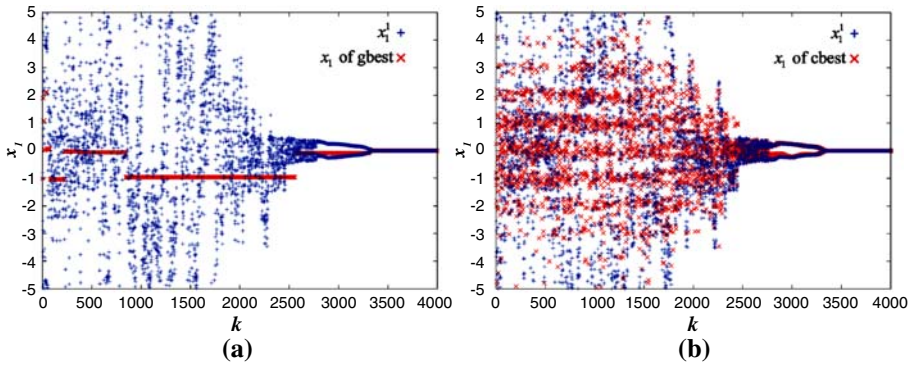


Fig. 4 Search Trajectories when the EC-DGCM is Applied to Prob. 6. (**a**) with gbest; (**b**) with cbest

4.2 Numerical simulations

In this section, we illustrate the effectiveness of the EC-DGCM through numerical simulations.

We first applied the EC-DGCM to Prob. 6 (Original Rastrigin, $N = 4$) in order to demonstrate the search trajectories. In this simulation, we used the following parameter settings:

$$P = 10,\ k_{\max} = 4000,\ \bar{c}_1 = \bar{c}_2 = 0.05,\ T = 500, \tag{42}$$

which are fixed for all simulations hereinafter, and we set $\Delta T_{\max} = 0.03$ for this simulation only. We employed both gbest and cbest on $\boldsymbol{gcb}$ and then executed simulations with these conditions. We show the search trajectories in these simulations in Fig. 4. Figure 4a shows the search trajectory of the EC-DGCM with gbest, and Fig. 4b shows the search trajectory of the EC-DGCM with cbest. Both figures show that each search point is autonomously driven by the DCGMwT and synchronizes to the trajectory that searches the neighborhood of elite search points. The search points then converge to global minima without being trapped at local minima. As mentioned earlier, the behaviors of search trajectories differ between the EC-DGCM with gbest and that with cbest. In the EC-DGCM with gbest, since gbest tends to remain constant, the population tends to search the neighborhood of gbest intensively with some loss of diversity due to the synchronized trajectory. In comparison, because cbest is updated at each step and tends to remain in regions where the objective function value is smaller, the population tends to search the regions while maintaining some degree of diversity. Note that the plots of the search trajectory of cbest become dense in the neighborhood of local minima in Fig. 4b. Thus, the behaviors differ between gbest and cbest, although both converge to the global minimum while generating the synchronization phenomena with the elite search points, and the effectiveness of the EC-DGCM is confirmed through this simulation.

Next, we confirm the global search capability of the EC-DGCM through application to five proper benchmark problems ($N = 100$). In the simulation, we gave careful consideration to the following points:

(a) Proper benchmark problems
Typical benchmark problems have several drawbacks, as reported by Liang et al. in [11,12]. One of these drawbacks is that they have the same values among all variables at the global minimum. Another is that there is no linking among all of the variables. Hence, by application only to typical benchmark problems, we cannot properly evaluate

the global optimization capability of an optimization method. In the present paper, we use proper benchmark problems, which avoids the abovementioned drawbacks. Specifically, the position at the global minimum is displaced randomly for each variable. Then, if there is no linking among all of the variables in the problem, their axes are rotated using the following equation:

$$\tilde{\boldsymbol{x}} = R\boldsymbol{x} \tag{43a}$$

$$R = T^{12} \times T^{13} \times \cdots \times T^{1N} \times T^{23} \times \cdots \times T^{N-1N} \tag{43b}$$

$$T_{kl}^{ij} = \begin{cases} \cos\alpha & k = i,\ l = i \\ -\sin\alpha & k = i,\ l = j \\ \sin\alpha & k = j,\ l = i \\ \cos\alpha & k = j,\ l = j \\ 1 & k = l \neq i, j \\ 0 & \text{not above,} \end{cases} \tag{43c}$$

where T^{ij} is $N \times N$ matrix. (43d)

This rotation procedure is necessary for Prob. 4 and Prob. 5. These proper benchmark problems are shown in Appendix A.

(b) Gradient computation

In order to compute the gradient, we use automatic differentiation (AD) with the ADOL-C tapeless forward mode [13,14]. Automatic differentiation computes the target function and its gradient simultaneously. The computation amount of this procedure is at most a constant times the computation amount of the target function. In other words, let the computation amount of a function f be $L(f)$, then

$$L(E(\boldsymbol{x})) = nL(E(\boldsymbol{x}), \nabla E(\boldsymbol{x})) \tag{44}$$

where $(E(\boldsymbol{x}), \nabla E(\boldsymbol{x}))$ is computed with AD.

Actually, we check $n \approx 5$ through numerical experiments for all benchmark problems in the case of $N = 100$.

(c) Comparison with other methods

For comparison, we also apply several other global optimization methods to the same problems. Since the EC-DGCM uses the gradient, we choose global optimization methods that also use the gradient. Specifically, we choose the following methods:

1. *TRUST (Terminal Repeller Unconstrained Subenergy Tunneling)* [15]
   This is a hybrid algorithm in which a Tunneling Algorithm and a function transformation method are used. The function transformation is used in the tunneling procedure.
2. *HDA (Hybrid Descent Algorithm)* [16]
   This is a hybrid algorithm in which Simulated Annealing and the Descent Method are used. Simulated Annealing is used in the global search phase, and the Decent Method is used in the local search phase.
3. *GRPSO (hybrid GRadient and Particle Swarm Optimization)* [17]
   This is a hybrid algorithm in which PSO and the Decent Method are used. Particle Swarm Optimization is used in the global search phase, and the Decent Method is used in the local search phase. In [17], the repulsion technique [18] is used in the PSO procedure. However, we do not use the repulsion technique, because the

Table 2 Results for Prob. 1 (Modified Levy No. 5 $N = 100$). $\Delta T_{\max} = 0.02$

Method		CR	DA	Var	Worst	Best	AD Calls	OF Calls
	no coupling	**100**	**0**	**0**	**0**	**0**	40000	0
EC-DGCM	with gbest	**100**	**0**	**0**	**0**	**0**	40000	0
	with cbest	**100**	**0**	**0**	**0**	**0**	40000	0
TRUST		0	171.900	291.754	207.867	125.279	4614	176929
HDA		**100**	**0**	**0**	**0**	**0**	737	196317
GRPSO		0	91.199	1810.950	179.324	29.505	206	198975

The result of the non-coupling type corresponds to that of running $P(=10)$ DGCMs in parallel
In the methods for comparison, we use the termination criterion as AD Calls + OF Calls$/5(= n) > 40{,}000$ (please see Eq. 45)
The bold font denotes that its evaluation is the best among all models

GRPSO without the repulsion technique provides better results than that with the repulsion technique in the benchmark problems in the present paper.

The parameters in these methods are adjusted in order to perform best for the benchmark problems. The termination criteria in these models should be set fairly with respect to the computation cost. The EC-DGCM mainly costs $P \times k_{\max}$ AD calls (which computes the target function and its gradient simultaneously). In this simulation, the number of AD calls for the EC-DGCM is 40,000 ($P = 10, k_{\max} = 4000$). Thus, we use the following termination criterion:

$$\text{AD calls} + \text{objective function calls}/n > 40{,}000. \tag{45}$$

We use $n = 5$, taking the discussion in (b) into consideration in the case of $N = 100$.

In the EC-DGCM, fixed parameters $P, \bar{c}_1, \bar{c}_2, T, k_{\max}$ were set as Eq. 42. $\Delta T_{\max}$ was set to generate chaotic trajectories for each of the problems. Both the gbest-type and the cbest-type were simulated. In addition, the non-coupling type was also simulated. Note that the non-coupling type corresponds to running $P(= 10)$ DGCMs in parallel. We then ran 100 trials, randomly resetting the initial points. The results are shown in Tables 2–6. We evaluate the results using following indices:

- CR : Convergence Rate to global minima.
- DA : Deviation Average. Deviation is calculated as

$$\begin{cases} E(\boldsymbol{x}) - E(\boldsymbol{x}^*) & |E(\boldsymbol{x}^*)| < 1 \\ \dfrac{E(\boldsymbol{x}) - E(\boldsymbol{x}^*)}{|E(\boldsymbol{x}^*)|} & |E(\boldsymbol{x}^*)| \geq 1 \end{cases} \tag{46}$$

  where $\boldsymbol{x}^*$ is the global minimum.
- Var : Variance of Deviation.
- Worst : Worst objective function value obtained as a result in all trials.
- Best : Best objective function value obtained as a result in all trials.

We also show the following indices regarding computation cost:

- AD Calls : Number of Automatic Differentiation (AD) Calls. AD computes the objective function value and its gradient simultaneously.
- OF Calls : Number of Objective Function Calls.

Note that each method is applied fairly with respect to computation cost, as mentioned above. In Tables 2–6, the bold font denotes that its evaluation is the best among all models. The average computation time of the EC-DGCM for each problem is 9.0–13.5(s) with

Table 3 Results for Prob. 2 (Modified Griewank $N = 100$). $\Delta T_{\max} = 150$

Method		CR	DA	Var	Worst	Best	AD Calls	OF Calls
	no coupling	0	0.00483	2.907e−06	**0.00898**	0.00059	40000	0
EC-DGCM	with gbest	0	0.00611	7.879e−06	0.01988	0.00129	40000	0
	with cbest	0	**0.00342**	**2.315e−06**	0.00960	0.00051	40000	0
TRUST		0	1.07882	7.789e−05	1.10375	1.05862	16629	116856
HDA		0	0.00759	1.932e−05	0.01963	**0.00030**	1101	194497
GRPSO		0	0.89598	4.507e−02	1.01053	0.15580	1354	193235

The result of the non-coupling type corresponds to that of running $P(=10)$ DGCMs in parallel
In the methods for comparison, we use the termination criterion as AD Calls + OF Calls$/5(= n) > 40{,}000$ (please see Eq. 45)
The bold font denotes that its evaluation is the best among all models

Table 4 Results for Prob. 3 (Modified Rosenbrock's Saddle $N = 100$). $\Delta T_{\max} = 0.001$

Method		CR	DA	Var	Worst	Best	AD Calls	OF Calls
	no coupling	**2**	**0.647**	**1.129**	4.344	**0**	40000	0
EC-DGCM	with gbest	0	122.806	4753.270	245.436	0.0899	40000	0
	with cbest	0	107.206	5290.080	251.401	0.1423	40000	0
TRUST		0	1.631	3.799	**3.995**	0.0002	12959	135204
HDA		0	94.301	7424.660	302.084	0.0005	1607	191964
GRPSO		0	341.983	2289.730	498.252	221.9670	101	199500

The result of the non-coupling type corresponds to that of running $P(=10)$ DGCMs in parallel
In the methods for comparison, we use the termination criterion as AD Calls + OF Calls$/5(= n) > 40{,}000$ (please see Eq. 45)
The bold font denotes that its evaluation is the best among all models

Table 5 Results for Prob. 4 (Modified2^N-minima $N = 100, \alpha = \pi/4$). $\Delta T_{\max} = 0.015$

Method		CR	DA	Var	Worst	Best	AD Calls	OF Calls
	no coupling	27	0.00385	8.723e−06	0.01083	**0**	40000	0
EC-DGCM	with gbest	**88**	**0.00047**	**1.734e−06**	**0.00722**	**0**	40000	0
	with cbest	**88**	0.00051	2.090e−06	**0.00722**	**0**	40000	0
TRUST		0	0.06918	2.135e−04	0.10327	0.04146	436	197820
HDA		0	0.02222	7.370e−05	0.04331	0.00722	64	199680
GRPSO		0	0.02387	1.013e−04	0.05053	0.00361	61	199700

The result of the non-coupling type corresponds to that of running $P(=10)$ DGCMs in parallel
In the methods for comparison, we use the termination criterion as AD Calls + OF Calls$/5(= n) > 40{,}000$ (please see Eq. 45)
The bold font denotes that its evaluation is the best among all models

"AMD Opteron™252 (2.6 GHz)". Except for Table 4, the results show that the EC-DGCM is more effective than other methods. As for the comparison of gbest and cbest, we consider cbest to be better because the EC-DGCM with cbest searches the regions in which the objective function value is smaller while maintaining diversity to some extent. Furthermore, its search trajectories do not synchronize completely, and not all of the search points concentrate to a certain search point in the local search phase. Hence, they search regions in which the objective function is smaller in the global search phase while maintaining diversity using the chaotic dynamics, and then each search point converges to a different local minimum, each of which is close to the others and lies in the lowest region of the feasible region. Thus, the

Table 6 Results for Prob. 5 (Modified Rastrigin $N = 100, \alpha = \pi/4$). $\Delta T_{\max} = 0.03$

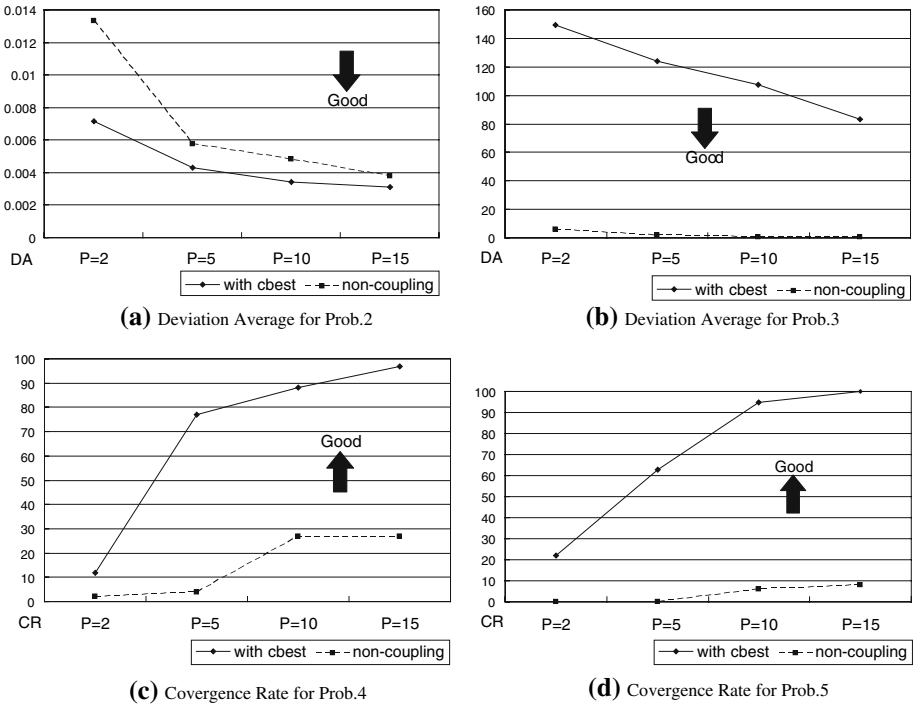
Method		CR	DA	Var	Worst	Best	AD Calls	OF Calls
	no coupling	6	13.442	292.716	81.587	**0**	40000	0
EC-DGCM	with gbest	14	2.756	5.462	9.950	**0**	40000	0
	with cbest	**95**	**0.050**	**0.047**	**0.995**	**0**	40000	0
TRUST		0	1316.820	19302.700	1566.950	963.850	40	199800
HDA		0	1254.670	22165.600	1605.830	965.101	246	198769
GRPSO		0	937.351	52782.600	1408.700	465.638	41	199802

The result of the non-coupling type corresponds to that of running $P(=10)$ DGCMs in parallel
In the methods for comparison, we use the termination criterion as AD Calls + OF Calls$/5(= n) > 40{,}000$ (please see Eq. 45)
The bold font denotes that its evaluation is the best among all models

EC-DGCM with cbest achieves diversification and intensification properly, and so is effective for the global optimization. Here, the result of Table 4 is discussed. Prob. 3 has a bent ridge structure [19] (a valley in which a global minimum lies is bent, and the variation of the objective function value along the bent direction is much smaller than that along the other directions). Hence, an optimization method, which descends gradually along the descent direction given by the gradient, tends to provide good results on Prob. 3. TRUST implements the abovementioned search suitable for Prob. 3 and performs well. The EC-DGCM without coupling also implements a search suitable for Prob. 3 in the local search phase. However, couplings with elite search points prevent the search suitable for Prob. 3. Therefore, the EC-DGCM with gbest and cbest do not perform well.

Next, we confirm the variation of the global search capability of the proposed model w.r.t. P through application to the same problems ($N = 100$). In this simulation, we employed cbest on $\boldsymbol{gcb}$ and used the same parameters, except P. We applied the EC-DGCM with cbest and without couplings, setting $P = 2, 5, 15$. The results are shown in Fig. 5. The results show that an increase of P increases the optimization capability. This increase peaks out when P is enlarged to a certain degree. In addition, the results show the effect of coupling with elite search points even in the situation of a small number of search points, except for (b). Hence, we should use coupling with elite search points and enlarge P to a degree.

Furthermore, we applied the EC-DGCM to proper benchmark problems with 1000 variables ($N = 1000$). Fixed parameters $P, \bar{c}_1, \bar{c}_2, T, k_{\max}$ were set as Eq. 42, as in the case of $N = 100$, essentially. For Prob. 5, we used $\bar{c}_1 = \bar{c}_2 = 0.1$. For Prob. 2, a local search (Armijo's linear search [20]) was implemented for a small number of steps after the termination of global search. As for the computation cost ratio between AD Calls and OF Calls, $n = 50$ was used for Prob. 1, Prob. 2, and Prob. 3, and $n = 15$ was used for Prob. 4 and Prob. 5 (please see Eq. 45). The non-coupling type EC-DGCM and the other methods for comparison were also applied. We then ran 20 trials, randomly resetting initial points. The results are shown in Tables 7–11. We also evaluated the results using the indices mentioned earlier. In Tables 7–11, the bold font denotes that its evaluation is the best among all models. In Tables 8, 10 and 11, the EC-DGCM is also more effective than other methods. Meanwhile, in Table 7, the HDA performs better (although the performance of the EC-DGCM is similar to the performance of the HDA). In Table 9, the non-coupling type EC-DGCM and HDA performs better. This trend also exists in the results for 100-variable problems. We consider this trend to increase for 1000-variable problems. On the whole, we also consider the EC-DGCM to be effective for very high dimensional problems.



(a) Deviation Average for Prob.2

(b) Deviation Average for Prob.3

(c) Covergence Rate for Prob.4

(d) Covergence Rate for Prob.5

Fig. 5 Variation of Global Search Capability w.r.t. P. The results of $P = 10$ are from Tables 2–6. For Prob. 1, the convergence rates of all results are 100%. Hence, these results are not shown

Table 7 Results for Prob. 1 (Modified Levy No. 5 $N = 1000$). $\Delta T_{\max} = 0.02$

Method		CR	DA	Var	Worst	Best	AD Calls	OF Calls
	no coupling	0	37.452	2.517	39.659	34.258	40000	0
EC-DGCM	with gbest	75	7.700e−04	1.779e−06	0.003	**0**	40000	0
	with cbest	75	7.704e−04	1.780e−06	0.003	**0**	40000	0
TRUST		0	167.331	35.639	179.360	157.241	3037	1848150
HDA		**100**	**0**	**0**	**0**	**0**	3202	1839920
GRPSO		0	160.879	16.382	169.402	154.760	101	1994960

The result of the non-coupling type corresponds to that of running $P(=10)$ DGCMs in parallel
In the methods for comparison, we use the termination criterion as AD Calls + OF Calls$/n > 40{,}000$ (see Eq. 45)
The bold font denotes that its evaluation is the best among all models

5 Conclusion

In the present paper, we propose a new global optimization model, EC-DGCM, in which multiple search points are autonomously driven by the chaotic dynamic model and synchronize to elite search points with the PD type coupling model. The proposed model has the following effective diversification strategies and intensification strategies. The diversification strategies are the autonomous search trajectory generated by the DGCM and the possibility of searches in multiple valleys by multiple search points. The intensification strategies are the convergence property of the DGCM to the bottom of the widest valley and the synchronization to elite search points taking the objective function value into consideration. Applying

Table 8 Results for Prob. 2 (Modified Griewank $N = 1000$). $\Delta T_{\max} = 150$

Method		CR	DA	Var	Worst	Best	AD Calls	OF Calls
	no coupling	0	0.00211	8.921e−07	0.00384	0.00016	40082	708
EC-DGCM	with gbest	0	0.00345	3.704e−06	0.00916	0.00145	40361	3790
	with cbest	0	**0.00198**	4.063e−07	**0.00362**	**0.00115**	40546	5846
TRUST		0	1.00323	**4.385e−09**	1.00333	1.00309	39471	26494
HDA		0	0.00861	1.295e−05	0.02074	0.00499	4509	1774550
GRPSO		0	1.15672	6.078e−05	1.16332	1.12518	167	1991680

The result of the non-coupling type corresponds to that of running $P(=10)$ DGCMs in parallel
In the methods for comparison, we use the termination criterion as AD Calls + OF Calls$/n > 40{,}000$ (see Eq. 45)
The bold font denotes that its evaluation is the best among all models

Table 9 Results for Prob. 3 (Modified Rosenbrock's Saddle $N = 1000$). $\Delta T_{\max} = 0.001$

Method		CR	DA	Var	Worst	Best	AD Calls	OF Calls
	no coupling	0	**399.735**	**6077.20**	**514.692**	232.406	40000	0
EC-DGCM	with gbest	0	1942.770	17775.60	2148.910	1591.920	40000	0
	with cbest	0	1818.420	13717.00	2030.920	1527.390	40000	0
TRUST		0	4699.960	75443.60	5167.500	3963.000	101	1994960
HDA		0	650.571	44528.70	996.007	**73.847**	4472	1776380
GRPSO		0	4699.960	75443.60	5167.500	3963.000	101	1994960

The result of the non-coupling type corresponds to that of running $P(=10)$ DGCMs in parallel
In the methods for comparison, we use the termination criterion as AD Calls + OF Calls$/n > 40{,}000$ (see Eq. 45)
The bold font denotes that its evaluation is the best among all models

Table 10 Results for Prob. 4 (Modified 2^N-minima $N = 1000, \alpha = \pi/4$). $\Delta T_{\max} = 0.015$

Method		CR	DA	Var	Worst	Best	AD Calls	OF Calls
	no coupling	0	0.01606	2.893e−06	0.01987	0.01156	40000	0
EC-DGCM	with gbest	**40**	**0.00040**	**1.288e−07**	**0.00108**	**0.00000**	40000	0
	with cbest	30	0.00041	1.730e−07	0.00144	**0.00000**	40000	0
TRUST		0	0.08046	1.454e−05	0.08811	0.07331	845	587328
HDA		0	0.03804	9.309e−06	0.04441	0.03437	3581	546286
GRPSO		0	0.04173	2.551e−05	0.05558	0.03351	101	598490

The result of the non-coupling type corresponds to that of running $P(=10)$ DGCMs in parallel
In the methods for comparison, we use the termination criterion as AD Calls + OF Calls$/n > 40{,}000$ (see Eq. 45)
The bold font denotes that its evaluation is the best among all models

proper benchmark problems with 100 variables and 1000 variables, we confirm that the EC-DGCM with cbest is more effective than the EC-DGCM with gbest or no coupling, as well as other global optimization methods that use the gradient. However, the EC-DGCM has two disadvantages. First, if the global minimum lies on the boundary, then the EC-DGCM cannot obtain it because of the property of toroidalization. We expect that this will be solved by redefining the confining region more widely. Second, the EC-DGCM is not useful for problems that have a plateau or a ridge structure, such as Prob. 3. We expect that this will be solved by taking a longer time for the local search process. In addition, a formal convergence proof for the proposed method is not given in this paper. We herein perform a quantitative analysis for the convergence solution of the DGCM. However, this solution is not necessarily

Table 11 Results for Prob. 5 (Modified Rastrigin $N = 1000$, $\alpha = \pi/4$). $\Delta T_{\max} = 0.03$

Method		CR	DA	Var	Worst	Best	AD Calls	OF Calls
	no coupling	0	5040.88	3.685e+04	5424.48	4648.42	40010	0
EC-DGCM	with gbest	0	544.64	1.697e+04	816.86	333.31	40010	0
	with cbest	0	**153.17**	**6.046e+03**	**266.65**	**51.74**	40010	0
TRUST		0	14007.00	2.982e+05	14980.30	13031.20	62	599066
HDA		0	8671.17	1.305e+06	10701.60	6259.24	68	598978
GRPSO		0	8484.06	7.781e+05	11030.20	7235.28	70	598960

The result of the non-coupling type corresponds to that of running $P(=10)$ DGCMs in parallel
In the methods for comparison, we use the termination criterion as AD Calls + OF Calls$/n > 40{,}000$ (see Eq. 45)
The bold font denotes that its evaluation is the best among all models

one of global minima, and there is no global convergence proof for the DGCM. In addition, multiple points driven by the DGCM are synchronized in the proposed method. This makes its dynamics more complex. Hence, it is not necessarily easy to prove mathematically that the proposed method converges to global minima with certainty. The proposed method is heuristic in that no global convergence proof is given. In the future, we will investigate solutions to the abovementioned problems. In addition, the framework of the proposed method, that is, the synchronization of the nonlinear dynamic optimization method to elite search points with PD type coupling, will be expanded so as to be applicable with other methods.

Acknowledgements This research was supported in part by the Japan Society for the Promotion of Science through a Grant-in-Aid for JSPS Fellows (No.18-5977) and a Grant-in-Aid for Scientific Research (C) (No.19500196).

Appendix A: Benchmark problems

This appendix lists the benchmark problems that are used in the present paper. The rotation Matrix R is given by Eq. 43. In the following, $\boldsymbol{x}^*$ denotes the coordinates of the global minimum (given randomly), and $U(a, b)$ is the uniform distribution between a and b.

Prob. 1 Modified Levy No.5. The original version is given in [21].

$$\begin{cases} \min\limits_{\boldsymbol{x}} E(\boldsymbol{y}(\boldsymbol{x})) = \dfrac{\pi}{N}\Big[B\sin^2(\pi y_1) + \displaystyle\sum_{i=1}^{N-1}\{(y_i - A)^2(1 + B\sin^2(\pi y_{i+1}))\} + (y_N - A)^2\Big] \\ y_i = A + 10.0 z_i, \;\; \boldsymbol{z} = \boldsymbol{x} - \boldsymbol{x}^*, \;\; x_i^* = U(-0.8, 0.8) \\ \text{subj. to } |x_i| < 1.0, \;\; i = 1, \ldots, N \\ \text{where } A = 1.0, \;\; B = 5.0 \\ \text{minima: } E(\boldsymbol{y}(\boldsymbol{x}^*)) = 0.0 \end{cases}$$

Prob. 2 Modified Griewank function. The original version is given in [22].

$$\begin{cases} \min\limits_{\boldsymbol{x}} E(\boldsymbol{z}(\boldsymbol{x})) = \dfrac{1}{ND}\displaystyle\sum_{i=1}^{N}\{z_i^2\} - \prod_{i=1}^{N}\left\{\cos\left(\frac{z_i}{\sqrt{i+1}}\right)\right\} + 1.0 \\ \boldsymbol{z} = \boldsymbol{x} - \boldsymbol{x}^*, \;\; x_i^* = U(-20.0, 20.0) \\ \text{subj. to } |x_i| < 25.0, \;\; i = 1, \ldots, N \\ \text{where } D = 2000.0 \\ \text{minima: } E(\boldsymbol{z}(\boldsymbol{x}^*)) = 0.0 \end{cases}$$

Prob. 3 Modified Rosenbrock's Saddle function. The original version is given in [23].

$$\begin{cases} \min\limits_{\boldsymbol{x}} E(\boldsymbol{z}(\boldsymbol{x})) = \sum\limits_{i=1}^{N-1} \left\{100.0(z_{i+1} - z_i^2)^2 + (z_i - 1)^2\right\} \\ \boldsymbol{z} = \boldsymbol{x} - \boldsymbol{x}^* + \mathbf{1.0}, \;\; x_i^* = U(-2.4, 0.4) \\ \text{subj. to } -2.0 - 1.0 < x_i < 2.0 - 1.0, \;\; i = 1, \ldots, N \\ \text{minima } E(\boldsymbol{z}(\boldsymbol{x}^*)) = 0.0 \end{cases}$$

Prob. 4 Modified 2^N-minima function. The original version is given in [24].

$$\begin{cases} \min\limits_{\boldsymbol{x}} E(\boldsymbol{z}(\boldsymbol{x})) = \sum\limits_{i=1}^{N} \left\{z_i^4 - 16.0z_i^2 + 5.0z_i\right\} \\ \boldsymbol{z} = R(\alpha)(\boldsymbol{x} - \boldsymbol{x}^*) - \mathbf{2.9035}, \;\; x_i^* = U(-1.0, 7.0), \;\; \alpha = \pi/4 \\ \text{subj. to } -5.0 + 2.9035 < x_i < 5.0 + 2.9305, \;\; i = 1, \ldots, N \\ \text{minima: } E(\boldsymbol{z}(\boldsymbol{x}^*)) = -78.319N \end{cases}$$

Prob. 5 Modified Rastrigin function. The original version is given in [25].

$$\begin{cases} \min\limits_{\boldsymbol{x}} E(\boldsymbol{z}(\boldsymbol{x})) = AN + \sum\limits_{i=1}^{N} z_i^2 - A\cos(B\pi z_i) \\ \boldsymbol{z} = R(\alpha)(\boldsymbol{x} - \boldsymbol{x}^*), \;\; x_i^* = U(-4.0, 4.0), \;\; \alpha = \pi/4 \\ \text{subj. to } |x_i| < 5.0, \;\; i = 1, \ldots, N \\ \text{where } A = 10.0, \;\; B = 2.0 \\ \text{minima: } E(\boldsymbol{z}(\boldsymbol{x}^*)) = 0.0 \end{cases}$$

Prob. 6 Original Rastrigin function [25].

$$\begin{cases} \min\limits_{\boldsymbol{x}} E(\boldsymbol{x}) = AN + \sum\limits_{i=1}^{N} x_i^2 - A\cos(B\pi x_i) \\ \text{subj. to } |x_i| < 5.0, \;\; i = 1, \ldots, N \\ \text{where } A = 10.0, \;\; B = 2.0 \\ \text{minima: } E(\boldsymbol{x}^*) = 0.0 \end{cases}$$

Appendix B: Derivation of the emergence condition of the chaotic synchronization for c

For simplicity, let $N = 1$ and $P = 2$, and let $x^*(k) = x^1(k) = x^2(k)$ denote the synchronization state at time k. Let us consider $x^p(k) = x^*(k) + \delta x^p(k)$, which is a neighborhood point of $x^*(k)$. Let

$$\begin{pmatrix} x^*(k+1) \\ x^*(k+1) \end{pmatrix} = C^{-1} \begin{pmatrix} g(x^*(k)) \\ g(x^*(k)) \end{pmatrix}. \tag{47}$$

Then, $x^p(k+1) = x^*(k+1) + \delta x^p(k+1)$, which is a neighborhood point of $x^*(k+1)$ at time $k+1$, is written as

$$\begin{pmatrix} x^*(k+1) + \delta x^1(k+1) \\ x^*(k+1) + \delta x^2(k+1) \end{pmatrix} = C^{-1} \begin{pmatrix} g(x^*(k) + \delta x^1(k)) \\ g(x^*(k) + \delta x^2(k)) \end{pmatrix}, \tag{48}$$

where $g(x^*(k) + \delta x^p(k))$ is written as

$$g(x^*(k) + \delta x^p(k)) \approx g(x^*(k)) + \frac{dg(x^*(k))}{dx}\delta x^p(k). \tag{49}$$

Substituting Eq. 49 into Eq. 48 gives the following difference equation:

$$\begin{pmatrix}\delta x^1(k+1)\\ \delta x^2(k+1)\end{pmatrix} = \frac{dg(x^*(k))}{dx} C^{-1} \begin{pmatrix}\delta x^1(k)\\ \delta x^2(k)\end{pmatrix}. \tag{50}$$

Hence, the perturbation from the synchronization state at time $k + K$ is written as

$$\begin{pmatrix}\delta x^1(k+K)\\ \delta x^2(k+K)\end{pmatrix} = \prod_{l=k}^{k+K} \left\{\frac{g(x^*(l))}{dx}\right\} \left(C^{-1}\right)^K \begin{pmatrix}\delta x^1(k)\\ \delta x^2(k)\end{pmatrix}. \tag{51}$$

Let the eigenvalues of C^{-1} be ν_1, ν_2 and their eigenvectors be $\boldsymbol{c}_1$, $\boldsymbol{c}_2$. Then, $\nu_1 = \dfrac{1}{1+2c}$, $\nu_2 = 1$, $\boldsymbol{c}_1 = (1, -1)^T$, $\boldsymbol{c}_2 = (1, 1)^T$ under Eq. 31. Hence, we obtain

$$\begin{pmatrix}\delta x^1(k+K)\\ \delta x^2(k+K)\end{pmatrix} = \prod_{l=k}^{k+K} \left\{\frac{g(x^*(l))}{dx}\right\} (d_1 \nu_1^K \boldsymbol{c}_1 + d_2 \nu_2^K \boldsymbol{c}_2), \tag{52}$$

where d_1 and d_2 are non-zero coefficients. Let λ be the Lyapunov index $\lambda = \left\langle \ln \left| \dfrac{g(x^*)}{dx} \right| \right\rangle$. Then, we obtain

$$\lambda + \ln|\nu_1| < 0 \tag{53a}$$

$$\frac{\exp(\lambda)}{1+2c} < 1 \tag{53b}$$

as a condition of the synchronization stability.

Appendix C Pseudo code of the proposed model

- **Variables**
  $\boldsymbol{x}^p$, $\boldsymbol{g}^p$, $\boldsymbol{pb}^p$: N dimension vectors ($p = 1, \ldots, P$)
  $\boldsymbol{gb}$, $\boldsymbol{cb}$, $\boldsymbol{gcb}$: N dimension vectors
  E^p, E^p_{pb} : scalar values ($p = 1, \ldots, P$)
  k, c_1, c_2, elite, ΔT : scalar values
- **Step1 Initialization**
  Set the following parameters: $\bar{c}_1$, $\bar{c}_2$, $\Delta T_{\max}$, $k_{\max}$, T, P
  Initialize the search points of all search points randomly:
  for $p \leftarrow 1$ **to** P
  for $i \leftarrow 1$ **to** N
  $x_i^p \leftarrow U(p_i, q_i)$
  endfor
  endfor
  where $U(a, b)$ generates a uniform distribution between a and b.
  Initialize the time: $k = 0$.
  Initialize the objective function value of the pbest:
  for $p \leftarrow 1$ **to** P $E^p_{pb} \leftarrow \infty$ **endfor**.

Choose the elite search point type: elite ← gbest or cbest.
goto Step2.

- **Step2 Evaluation for All Search Points**
Calculate the gradients and objective function values of all search points:
for $p \leftarrow 1$ **to** P
$E^p \leftarrow E(\boldsymbol{x}^p),\ \boldsymbol{g}^p \leftarrow \nabla E(\boldsymbol{x}^p)$
endfor.
Set the pbest for each search point:
for $p \leftarrow 1$ **to** P
if $E^p < E^p_{pb}$ **then** $E^p_{pb} \leftarrow E^p,\ \boldsymbol{pb}^p \leftarrow \boldsymbol{x}^p$ **endif**
endfor.
Set the gbest:
$E^{gb} \leftarrow E^1_{pb},\ \boldsymbol{gb} \leftarrow \boldsymbol{pb}^1$
for $p \leftarrow 2$ **to** P
if $E^p_{pb} < E^{gb}$ **then** $E^{gb} \leftarrow E^p_{pb},\ \boldsymbol{gb} \leftarrow \boldsymbol{pb}^p$ **endif**
endfor.
Set the cbest (if cbest is needed):
if elite is cbest **then**
$E^{cb} \leftarrow E^1,\ \boldsymbol{cb} \leftarrow \boldsymbol{x}^1$
for $p \leftarrow 2$ **to** P
if $E^p < E^{cb}$ **then** $E^{cb} \leftarrow E^p,\ \boldsymbol{cb} \leftarrow \boldsymbol{x}^p$ **endif**
endfor
endif.
Set the elite search point of all search points at k:
if elite is gbest **then**
$\boldsymbol{gcb} \leftarrow \boldsymbol{gb}$
else if elite is cbest **then**
$\boldsymbol{gcb} \leftarrow \boldsymbol{cb}$
endif.
goto Step3.

- **Step3 Check the Termination Criterion**
if $k > k_{\max}$ **then**
return E^{gb}, $\boldsymbol{gb}$ *and* **exit**.
endif.
goto Step4.

- **Step4 Main Search**
Set the sampling parameter at k: $\Delta T \leftarrow \Delta T_{\max}\left(1 - \dfrac{k}{k_{\max}}\right)$.
Set the coupling coefficients at k: $c_1 \leftarrow \bar{c}_1 \sin^2\left(\dfrac{2\pi k}{T}\right), c_2 \leftarrow \bar{c}_2 \sin^2\left(\dfrac{2\pi k}{T}\right)$.
Execute Chaotic Search:
for $p \leftarrow 1$ **to** P $\boldsymbol{x}^p \leftarrow \boldsymbol{x}^p - \Delta T \boldsymbol{g}^p$ **endfor**.
Execute PD type Coupling:
$C \leftarrow 1 + c_1 + c_2$
for $p \leftarrow 1$ **to** P $\boldsymbol{x}^p \leftarrow \dfrac{1}{C}\boldsymbol{x}^p + \dfrac{c_1}{C}\boldsymbol{gcb} + \dfrac{c_2}{C}\boldsymbol{pb}^p$ **endfor**.
Execute Toroidalization:
for $p \leftarrow 1$ **to** P
for $i \leftarrow 1$ **to** N

if $x_i^p > q_i$ **then**
$x_i^p \leftarrow (x_i^p - p_i)\mathrm{fmod}(q_i - p_i) + p_i$
else if $x_i^p < p_i$ **then**
$x_i^p \leftarrow (x_i^p - q_i)\mathrm{fmod}(q_i - p_i) + q_i$
endif
endfor
endfor.
Increment the time: $k \leftarrow k + 1$.
goto Step2.

References

1. Aihara, K., Takabe, T., Toyoda, M.: Chaotic neural networks. Phys. Lett. A. **144**(6–7), 333–340 (1990)
2. Tokuda, I., Onodera, K., Tokunaga, R., Aihara, K., Nagashima, T.: Global bifurcation scenario for chaotic dynamic systems that solve optimization problems and analysis of their optimization capability. Electron. Commun. Japan (Part III: Fund. Electron. Sci.) **81**(2), 1–12 (1998)
3. Masuda, K., Aiyoshi, E.: Global optimization method using chaos of discrete gradient dynamics. In: Proceedings of IFAC Workshops ALCOSP and PSYCO 2004, pp. 825–830 (2004)
4. Mizukami, M., Hirano, M., Shinjo, K.: Simultaneous alignment of multiple optical axes in a multistage optical system using Hamiltonian algorithm. Opt. Eng. **40**(3), 448–454 (2001)
5. Glover, F., Laguna, M.: Tabu search. Kluwer Academic Publishers, The Netherlands (1997)
6. Goldberg, D.E.: Genetic algorithms in search, optimization, and machine learning. Addison-Wesley, USA (1989)
7. Kennedy, J., Eberhart, R.C.: Particle swarm optimization. In: Proceedings of IEEE Int. Conf. Neural Networks, pp. 1942–1948 (1995)
8. Chao, J., Ratanaswan, W., Tsujii, S.: A new global optimization method and supervised learning of multilayer neural networks. IEICE Trans. E **73**(11), 1796–1799 (1990)
9. Fujisaka, H., Yamada, T.: Stability theory of synchronized motion in coupled-oscillator systems. Progr. Theor. Phys. **69**(1), 32–47 (1983)
10. Fujisaka, H., Yamada, T.: Stability theory of synchronized motion in coupled-oscillator systems. II. Progr. Theor. Phys. **70**(5), 1240–1248 (1983)
11. Liang, J.J., Suganthan, P.N., Deb, K.: Novel composition test functions for numerical global optimization. In: Proceedings of Swarm Intelligence Symposium, 2005 (SIS 2005), pp. 68–75 (2005)
12. Liang, J., Baskar, S., Suganthan, P., Qin, A.: Performance evaluation of multiagent genetic algorithm. Nat. Comput. **5**(1), 83–96 (2006)
13. Griewank, A., Juedes, D., Utke, J.: Algorithm 755; ADOL-C: a package for the automatic differentiation of algorithms written in C/C++. ACM Trans. Mathemat. Software. **22**(2), 131–167 (1996)
14. Walther, A., Griewank, A.: ADOL-C 1.10.1. http://www.math.tu-dresden.de/∼adol-c/. Cited 16 Apr 2006 (Last Modified in 2005)
15. Barhen, J.: TRUST: A deterministic algorithm for global optimization. Science **276**(5315), 1094–1097 (1997)
16. Yiu, K.F.C., Liu, Y., Teo, K.L.: A hybrid descent method for global optimization. J. Global Optimiz. **28**(2), 229–238 (2004)
17. Wang, Y.J., Zhang, J.S., Zhang, Y.F.: A fast hybrid algorithm for global optimization. In: Proceedings of Machine Learning and Cybernetics, 2005, pp. 3030–3035 (2005)
18. Parsopoulos, K.E., Vrahatis, M.N.: On the computation of all global minimizers through particle swarm optimization. IEEE Trans. Evol. Comput. **8**(3), 211–224 (2004)
19. Sakuma, J., Kobayashi, S.: Extrapolation-directed crossover for real-coded GA: overcomingdeceptive phenomena by extrapolative search. In: Proceedings of IEEE Congress on Evolutionary Computation, 2001 (CEC2001), pp. 655–662 (2001)
20. Armijo, L.: Minimization of functions having Lipschitz continuous first partial derivatives. Pacific J. Mathemat. **16**(1), 1–3 (1966)
21. Levy, A.V., Montalvo, A.: The tunneling algorithm for the global minimization of functions. SIAM J. Sci. Stat. Comput. **6**(1), 15–29 (1985)
22. Griewank, A.O.: Generalized descent for global optimization. JOTA. **34**(1), 11–39 (1981)

23. Rosenbrock, H.H.: An automatic method for finding the greatest or least value of a function. Comp. J. **3**(3), 175–184 (1960)
24. Yasuda, K., Iwasaki, N.: Adaptive particle swarm optimization using velocity information of swarm. In: Proceedings of IEEE Int. Conf. on Systems, Man, and Cybernetics, 2004, pp. 3475–3481 (2004)
25. Rastrigin, L.A.: Systems of extremal control. Nauka, Moscow (1974)